GRAPE

Group for Research in APplied Economics

GRAPE Working Paper # 22

Within occupation wage dispersion and the task content of jobs

Lucas van der Velde

FAME | GRAPE, 2017

Foundation of Admirers and Mavens of Economics
Group for Research in Applied Economics

Within occupation wage dispersion and the task content of jobs

Lucas van der Velde
GRAPE | FAME
Warsaw School of Economics

Abstract
The relation between income inequality and technological progress has many chapters, of which the most recent corresponds to the task content of jobs. Proponents of this theory suggest that falling prices of computational power coupled with the increasing power of computers leads to an increasing substitution of workers with computers and a hollowing of the middle of the income distribution. While empirical analysis on task content of jobs explain inequality between occupations, we test whether the framework can also foster our understanding of wage dispersion within occupations. Using European data, we obtain estimates of wage dispersion and residual wage dispersion for each occupation and relate it to the task content. The results suggest that non-routine intensive occupations presented greater wage dispersion, even after controlling for a variety of factors.

Keywords:
wage inequality, occupation, task content, routinization

JEL Classification
J31, J24

Corresponding author
Lucas van der Velde, lvelde@grape.org.pl

Acknowledgements
Earlier versions of this paper received valuable comments from J. Svejnar, A. Szulc, J. Tyrowicz, K. Staehr, W. Hardy, S. Estrin and participants of WLE in Trier (2017) , and the IEA 2017 and GRAPE economic seminars. This research was supported by a grant from the National Science Centre, UMO-2016/21/N/HS4/02108. Remaining errors are ours.

Published by: FAME | GRAPE
ISSN: 2544-2473

Foundation of Admirers and Mavens of Economics
ull. Mazowiecka 11/14
00-052 Warszaw
Poland

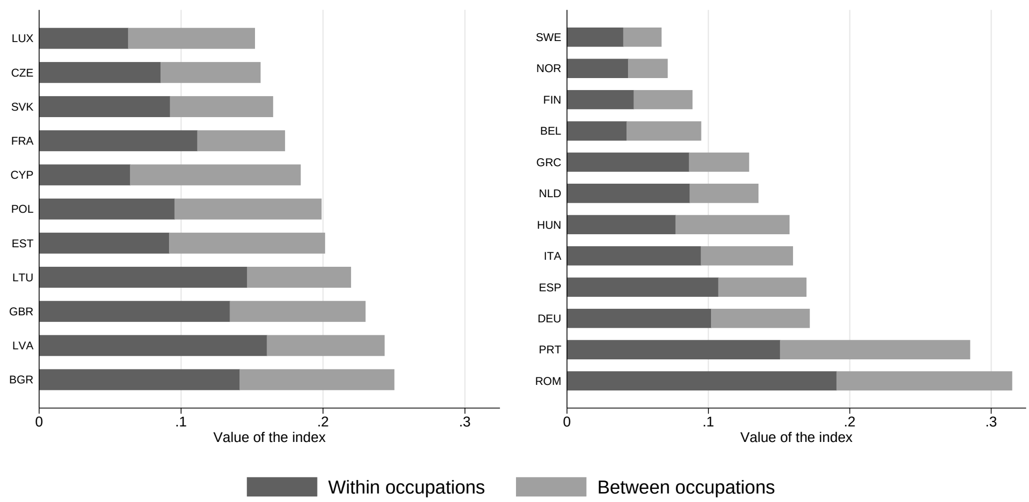
W | grape.org.pl
E | grape@grape.org.pl
TT | GRAPE_ORG
FB | GRAPE.ORG
PH | +48 799 012 202

1 Introduction

The relation between technological progress and inequality has been a central topic in economics at least since the Kuznets curve. In recent years, one brunch of research focused on the how new technologies (i.e. automation) affected the labor market. The authors point to at least two changes. First, a polarization of the demand for labor: new job opportunitties were developed at the top and at the bottom of the income distribution, whereas workers at the middle struggle to compete with machines Autor et al. (2003, 2006). Second, a change in the relative wages: workers at the top of the income distribution beneffited from greater complementarity with new technologies, which drove their wages further up Autor et al. (2003), Acemoglu and Autor (2011), Goos et al. (2014).

These contributions focus help to explain the growth in income inequality across occupations. Yet, changes in relative wages and demand across occupations represents only a part of the rise in labor market inequality, and not necessarily the most important. According to Kim and Sakamoto (2008), most of the increase in wage inequality in the United States was associated with a higher dispersion of wages *within* occupations. Differences across occupations, by contrast, appeared to be relatively stable and reflect business cycle fluctuations. This trend is not unique to the US. Figure 1 shows that in Europe within occupation wage inequality can represent as much as 50% of total wage inequality. These values are close to the figures reported in Kim and Sakamoto (2008) for the US and in Akerman et al. (2013) for Sweden.

Figure 1: Between and within components of Theil index



Notes: Figure displays the value of the Theil index for several European countries. Data were taken from EU-SES 2010. In countries on the left panel, the within component was calculated on ISCO 08 codes at the three digit levels. For countries on the right panel, only information at the two digit level was available.

In this article, we focus on within occupation wage inequality under the light of the task content of jobs. This literature starts from the premise that in order to understand how new techologies' impact on the labor market, one needs to understand what workers actually do on their jobs, i.e. what tasks workers perform. The task model allows for several possible links between task content of the job and within occupation wage inequality. A first explanation refers to task seggregation: workers in nominally identical occupations perform different tasks, which explains wage differences (see Deming 2015, Garicano and Hubbard 2016, for a theoretical treatment). Empirically, Visintin et al. (2015) who found that around 5% of the variance in wage inequality could be related to within occupation task dispersion. The second explanation focuses on the relation between the nature of task and workers productivity. This strand has not been analyzed empirically and constitutes the main focus of our paper. In particular, we test whether Occupations with higher non-routine intensity will present higher dispersion in wages.

This hypothesis is based on a common assumption behind theoretical models: abstract (cognitive and interpersonal) tasks are more sensitive to differences in workers' skill level than routine cognitive and manual tasks. Jung and Mercenier (2014) makes this assumption explicit, but it is much more widespread. In fact, literature characterizes routine tasks as leaving workers little autonomy (Oldenski 2012, Marcolin et al. 2016) and not exploiting creativity of the workers (Frey and Osborne 2013). Moreover, some of the indicators These descriptions indicate that workers who perform routine tasks are better substitutes for each other, i.e. that differences in productivity among them have a lower impact on output. Moreover, on the face of automations, workers performing might have less bargaining power, which should not only drive wage level down, but also the dispersion.

An alternative explanation of the link between the nature of tasks and wage dispersion refers to the reallocation process itself. As labor demand shifted towards non-routine intensive occupations, the number of mismatches in these occupations increases, even if the percentage remains constant. Workers who were initially well-matched to routine tasks might need to relocate to non-routine jobs, were they might perform poorly. Reallocation of workers can lead to higher wage dispersion in the new occupation by extending the left tail of the distribution. The loss of wages following a change in occupation has received some theoretical treatment in Carrillo-Tudela and Visschers (2013), who develop a model, where workers lose occupation-specific human capital as a result of change in occupations. Gathmann and Schonberg (2010) provided evidence that workers switching between occupations with different task content experience a wage reduction. To the best of our knowledge, no research has focused on the effect of newcomers on the wage distribution for the occupation.

Our analysis is not the first to link within group wage dispersion to task content. As mentioned, Visintin et al. (2015) explores variation in tasks performed and wage inequality, concluding that there is a positive, but small correlation between the two variables. Our analysis differs from theirs in that we keep the tasks performed by workers constant within the occupation. This arises partly as a response to data limitations, but also as a way to focus on alternative mechanisms. The current analysis is complementary and not substitute to Visintin et al. (2015). From a method perspective, our analysis is closer to Altonji et al. (2014). Their research shows that variance in task content of workers from different majors can explain wage dispersion. But, their unit of analysis is major, so the hypotheses they test is that if workers from a given major can perform very different occupations, this dispersion of occupation might correlate with variance in wages. Our analysis presents a different definition of groups and explores different sources of variation.

Evaluating the hypothesis requires reliable data on workers' occupation and their earnings. These data should minimize problems associated with misclassification of occupations or misreporting wages. Second, sample sizes should be large enough to allow for a meaningful calculation of dispersion measures within occupations. To the best of our knowledge, only data from the Structure of Earnings Survey of the European Union (EU-SES) meet these criteria for Europe. Consequently, we employ three editions of the EU-SES: 2002, 2006 and 2010. Since data are collected from employers, we can expect this database to be more accurate in terms of occupations and earnings than alternative sources.

We find that, consistent with our expectations, occupations that were more non-routine intensive presented greater dispersion of wages, with similar coefficients at the top and at the bottom of the within occupation income distribution. This relation appears robust to the introduction of workers' characteristics and other occupation traits, such as the change in employment or the presence of winner-takes-all markets. The size of the effect varies across specifications: an standard deviation increase in the routine content of the job (roughly the difference between) is associated with increases in wage dispersion of 6 to 10 percentage points. Such results indicate that policies to improve matches in non-routine occupations, as well as those to improve workers skills in those occupations, could be used to reduce wage dispersion. However, it is unlikely to bring large changes.

The remaining of the article is organized as follows. In the next section we review the literature on the task content of jobs, focusing on the interaction between task type and workers' skills. These section reviews both theoretical models and previous empirical results. The section provides further details on the databases that we employ: EU-SES, for measures of wage dispersion within occupations; and O*NET, for task content of jobs. Section presents the main results relating wage inequality and task content, whereas in Section we discuss alternative channels. Section concludes.

2 Different channels for increase in inequality

Two alternative arguments could explain raising inequality as a result of technological change: worker heterogeneity and firm heterogeneity. Worker heterogeneity is present from the very beginning of the literature on the task content (Autor et al. 2003), yet the effects of reallocation on wage inequality has not been explored before, to the extent of our knowledge. So, in the first part of this section we review the conditions under which (de)routinization might increase wage inequality, depending on how wages are formed and the relation between skills and tasks. The second channel states that firms with differing characteristics might have different incentives to invest in both in the acquisition of new capital (e.g. state-of-the-art software, computers, etc.) and the improvement of human capital within the firms, e.g. financing training for the workers. If only a share of firms could afford this investments, then workers in those firms might be more productive, even if *a priori* they had the same characteristics.

The two mechanisms might be at play at any point in time. In fact, through assortative matching, they might reinforce each other: more productive workers sort into firms that offer better development opportunities. Notwithstanding, the distinction between these channels is relevant from a policy perspective. If constraints appear at the firm level, for example due to lack of investment in new technologies, the policy recommendations would be different than if the problem lies at the worker level.

2.1 Worker heterogeneity

Literature treats the relation between tasks and wages in at least two different ways, depending on whether papers are theoretical or empirical. In the case of theoretical papers, workers specialize in the production of the single task in which they are relatively more productive. This result echoes the early insights from Roy's model (Roy 1951). Wages in these models are formulated as the product of two elements: a constant, competitive wage rate per unit produced, and the number of units produced, that is productivity. Since the wage rate is constant, wage inequality should be proportional to skill inequality.[1] In empirical analysis, such as Firpo et al.

[1] The exact relation between the wage and skill inequality depends on whether the measure of dispersion used is scale invariant or not. Ratios of percentiles are invariant, and thus wage inequality would be equal to skill inequality. Variance, on the contrary, depends on the unit of measurement, and therefore the relation between

(2011) or Autor and Handel (2013), it is recognized that jobs involved the realization of several types of tasks, i.e. there are no purely routine manual jobs. Wages in this case are related to both the type of tasks performed by the worker and the relative productivity in performing those tasks. Regardless of which model one chooses, it is possible to establish some fairly general conditions under which shifts in the demand for labor induced by deroutinization will increase wage inequality.

Jung and Mercenier (2014) develop a model that is empirically tested by Cortes (2016). Workers are born with an innate ability that is drawn from an unspecified probability distribution. This innate ability, which might be called z, can later be applied to performing manual, routine and non-routine tasks alike. The function $\phi_t(z)$ maps the relation between the distribution of innate ability and the productivity of the worker in performing the task t, where productivity can be understood as the number of units of tasks that the worker can supply in a given amount of time. The function $\phi_t(z)$ is bounded from below, it can never result in a negative productivity.

Jung and Mercenier (2014) further characterize the productivity functions, such that the following criteria are met:

$$0 \quad < \quad \frac{\partial \, ln\phi_m(z)}{\partial z} \quad < \quad \frac{\partial \, ln\phi_r(z)}{\partial z} \quad < \quad \frac{\partial \, ln\phi_{nr}(z)}{\partial z} \quad , \tag{1}$$

where m, r, nr stand for (non-routine) manual, routine (cognitive and manual) and non-routine (cognitive and interpersonal) jobs respectively. These criteria have three implications. First, and trivially, productivity is an increasing function of innate ability: more capable workers are more productive than the rest. Second, difference in productivity between workers of different skill levels depends on the tasks that they perform: differences in productivity are larger for workers in non-routine tasks than in routine tasks, where in turn are larger than in manual tasks. This result reflects the intuition that a lawyer can become an excellent cab driver, while the converse needs not to hold. Finally, productivity in different tasks is positively correlated, e.g. best lawyers also make the best cab-drivers, hairdressers, etc.[2]

As in Autor et al. (2003), Acemoglu and Autor (2011), Jung and Mercenier (2014) define the marginal worker as indifferent between working in each type of tasks. It is possible to find two ability levels that indicate the cut-off productivities, such that a worker supplies only one type of tasks. In consequence, it is possible to determine wages as follows, where C_i denote the marginal wage rate for each type of tasks:

$$W(z) = \begin{cases} C_m * \phi_m(z), & \text{if } z_i < z^* \\ C_r * \phi_r(z), & \text{if } z^* < z_i < z^{**} \\ C_{nr} * \phi_{nr}(z), & \text{if } z^{**} < z_i \end{cases}$$

The cut-off values z^* and z^{**} are endogenously determined by the model, and they represent the productivity of the marginal worker, who is indifferent between working in manual or routine jobs (z^*); and between routine and non-routine jobs (z^{**}). The process is efficient, which implies that more skilled workers will supply non-routine tasks, medium skilled routine, and low skilled,

the two dispersions is proportional. See Cowell (2011) for a detailed analysis of the properties of inequality measures.

[2]Given these criteria, one might wonder why workers supply different types of labor. To answer this question one should consider the entire model, which includes agents with Cobb-Douglas preferences over two goods. One of them is produced with low skill labor only (e.g. low-skill services) while the second (manufactured good) requires a combination of routine (supplied by workers and computers) and non-routine tasks. This setup guarantees the demand and supply of workers for each of the tasks.

manual. We can already observe that the relation between within occupation wage inequality and the task content of the job. To the extent that the distribution of productivity is more spread in non-routine tasks, we can also expect it to present higher inequality than in the remaining cases, *ceteris paribus.*[3]

Jung and Mercenier (2014) model routinization as an increase in computer capital, that lowers the demand for workers in middle-skilled occupations. As a result, there is also a movement of the cut-off values, such that the new cut-off values are closer to the center of the skill distributions. In short, the following inequalities hold: $z' > z^*$ and $z'' < z^{**}$. The change in the cut-offs has clear implications for wage inequality. By increasing the range of skills within the manual and non-routine occupations, we expect that wage inequality in these tasks will raise; while for workers in routine tasks, it might decrease due to the exit of the most and least productive workers within the skill level.

Jung and Mercenier (2014) present the two mechanisms that can lead to an increase in wage inequality based on differences across workers. First, the spread of the productivity distribution in non-routine tasks would by itself generate more inequality, even if workers have the same innate ability. In short, wage inequality might arise solely as the result of condition stated in equation 1. Second, changes in the demand for tasks result in movement of workers that could also raise within occupation wage inequality. These movements increase the range of ability in non-routine tasks at the expense of more routine tasks. This second mechanism could be interpreted using a matching setup, where workers are first perfectly matched with their skills and as a result of changes in the demand, each time new, worse matches are formed. This process decreases the average skill level in abstract tasks, and raises the average level in manual tasks.

Some of the criteria employed by Jung and Mercenier (2014) might be considered too strict; in particular, the positive correlation between skill levels in different tasks might raise some concerns. However, similar results can be found in other setups. Autor et al. (2003) present a model where workers have different, not necessarily related, productivity levels: one for routine tasks and another for non-routine tasks.[4] Similarly, we can define the marginal workers as that who is indifferent between working in each sector. In short, the following equality should hold for her: $w_r r = w_{nr} nr$, where w_i indicates the wage rate for performing i tasks.

The ratio of wages (w_r/w_{nr}) determines the cutoff productivity ratio ($\eta^* = nr/r$) such that workers with a higher productivity ratio will provide non-routine tasks. The routinization process envisioned by Autor et al. (2003) involves a decrease in w_r, which is translated into a movement of the cutoff productivity ratio, which in turn initiates a process of reallocation of workers. However, to determine whether these would lead to changes in wage inequality is harder. Given the assumption of a constant wage rate, inequality will raise only if the range of the productivity distribution increases. Following the assumption that skills are distributed uniformly, we can write the 90/10 ratio as follows:

$$\frac{p90(w_{nr})}{p10(w_{nr})} = \frac{0.9 * (1-a) + a}{0.1 * (1-a) + a},$$

where $a = w_r/w_{nr}r$ represents the non-routine skill level of the marginal worker, and thus it represents the minimum level of skills such that individuals move to the new sector. One can

[3]Clearly, this argument holds only if the number of workers in each task is comparable. The result might be reversed if the difference in the proportion of workers is large enough. In the extreme case, when only one worker serves the entire demand for non-routine tasks, there is no inequality.

[4]Though the authors do not specify the distribution from which these draws are made, we proceed in the analysis as if they were from a uniform distribution.

verify that $\partial(p90/p10)/\partial a < 0$, which implies that skill inequality will raise whenever there is an increase in the range of the skill distribution. Unlike Jung and Mercenier (2014), we can observe that if there is negative correlation between productivity in the two sectors, the distribution of skills will broaden and the inequality in non-routine tasks will increase with each new change. The same movement of workers would simultaneously reduce the ratio among low skill workers. The same diverging trends are then observed, under much different assumptions.

If samples are independent, then the effect on wage inequality depends crucially on the original cut-off value. For initial cut-off values higher than one, a fall in the cut-off values would bring only small decreases in income inequality. However, if the value is equal or smaller than one, then the result is reverse. Wage inequality increases strongly with the fall in the cut-off value. Given that the cut-off value can be interpreted as the inverse skill-premium, one can expect that the second case ($\eta * < 1$) will be more common. Figure A1 in the appendix illustrates changes at the top of the income distribution against cutoff values.

The revision of the model presented in Autor et al. (2006), Autor and Dorn (2013) has different implications for within occupation wage inequality. In their model, only high skill workers can perform abstract tasks, so mobility across worker types and the consequent spread of underlying skill distribution is not possible. Unit invariant measures of wage inequality should be constant for this group. Second, non-college workers have the same productivity in manual tasks, though their productivity in routine task differs. Workers with initial high productivity found jobs in routine tasks. Technological change leads to a progressive fall in employment of routine workers that affects first those who were less productive. Given a constant wage rate, wage inequality should decline in routine tasks, while no effects should be visible among workers in manual tasks.

To conclude this review, we can also analyze the empirical specifications used in Firpo et al. (2011) and Autor and Handel (2013). Both analyses assume that the process described above for individual tasks can be extended to situations where workers do not specialize in a task each. In this scenario, wages (or rather their logarithm) are the sum of the tasks performed by the worker and the wage rate assigned to them, plus an additional random term. In short, the specification resembles the following function $ln(wage_j) = \sum_{i=1}^{5} \beta_i task_{i,j} + e_j$, where j represents the worker and i the type of tasks. To derive a formal expression of the ratio of wages is much more complex in this case, even if the productivity distributions are assumed to follow the uniform distribution. Hence, we consider the variance, as a measure of dispersion.

We can observe that variance in (log) wages is a positive function of the variance in each of the tasks. If these tasks are assumed to be independent, the variance corresponds to a sum of the variances weighted by the squared value of the coefficients. Under this empirical setup, we can expect wage inequality to be larger in non-routine occupations if the following conditions are met. First, there is a greater dispersion in productivity in performing non-routine tasks, which is not a far-fledged assumption. Second, workers in more non-routine jobs spend more time in non-routine activities. If the dispersion in productivity in non-routine tasks is large, but workers devote little time to these tasks, then the impact on wage dispersion might be negligible. Third, the returns to these non-routine activities are higher in non-routine jobs. Were it not the case, differences in productivity would not be observable. Of these only the first condition is truly problematic, as productivity is seldom observable, let alone comparable across occupations. Indeed, the fact that workers spend more time on non-routine activities is part of the definition of non-routine intensive occupations. Similarly, for workers to spend more time on non-routine tasks, they should be compensated accordingly, which indicates that β_{nr} should be higher in non-routine occupations. The claims on time spent on working activities and rewards were empirically explored in Autor and Handel (2013), who concluded that these two propositions hold.

The revision presented here is indicative that there might be a relation between wage inequality and the task content of occupations. In particular, we argued that in more non-routine occupations wage inequality would be larger, and described two plausible explanations: a) that dispersion in productivity in the performance of such tasks is larger; and b) that wages the increase in the demand for these tasks leads to the formation of progressively worse matches, where the productivity of the marginal workers falls with each expansion. Only in the latter case, we allowed rewards for non-routine tasks to vary as a result of technological progress,[5] and this would exacerbate inequality under the condition that the rewards to non-routine tasks increase with the share of non-routine tasks performed by the worker. These insights lead to the elaboration of two hypotheses. First, we would expect that workers that switched occupations not only had lower salaries, but also that their exclusion leads to a reduction in wage inequality. Second, we expect that once we control for human capital variables, inequality in productivity will be smaller in routine intensive occupations.

3 Data

To test our hypotheses, we work mainly with data from the Structure of Earnings Survey of the European Union, provided by EUROSTAT (EU-SES). This database is collected in Member States of the European Union[6] every four years. So far, three waves were made available to the public: 2002, 2006 and 2010. Unlike LFS, primary sampling units are not workers, but firms. All non-agricultural firms with over ten employees are eligible to be included in the sample. This criterion represents a minimum that Member States must fulfill, but it is not binding. Member States can also include smaller firms in their samples. Similarly, the inclusion of workers in public administration is optional. Once firms are included in the sample, five employees are selected at random to participate in the survey. Workers' data are obtained directly from firms' records.

The EU-SES has two distinctive advantages over alternative data sources that make it suitable for this analysis. First, given that data come from firms' records, one can expect wages and hours worked (including overtime and bonus) to have an accuracy level that other databases can hardly match. The data collection process implies also that the coding of occupations is detailed and accurate. The second advantage comes from a larger sample size. For example, the 2010 EU-SES database for Great Britain has nearly 180 thousand observations, while the LFS from the same year had around 90 thousand, of which only half reported to be wage employed. The size of the databases, plus its accuracy allows calculating meaningful measures of within occupation income inequality that would not be possible with other sources.

Notwithstanding, there are also some limitations regarding the use of EU-SES. First, the sample does comprise small enterprises, public sector workers or self-employed workers, so it is not representative of the whole economy. Second, by construction, the survey leaves out questions concerning individual characteristics related to the situation in their household (marital status, number of children), and to previous employment experiences. Hence, we only observe the current occupation of a worker. The problem is somehow mitigated by the available information on tenure, which together with age and education help to identify new workers.

[5] In Autor et al. (2003) relative premium increases, but mainly as a result of the fall in the rewards of those performing routine tasks.

[6] The sample does not cover all Member States as data from some countries, e.g. Ireland, are not anonymized and only available for consultation *in situ*. Cyprus and Luxembourg are excluded as they do not present data on firm size.

The most important variable in this study is hourly wages. We derive this variable dividing monthly earnings, excluding bonuses and additional payments related to overtime and shift work, by the average number of hours worked. Since we are mostly interested in measures of wage inequality that a) are unit invariant and b) are calculated within country-years, we do not convert them to a common currency to avoid introducing additional errors in the data.

Besides hourly wages, we obtain measures of task content of occupations. The literature offers several alternatives to measure the task content of jobs, each of them with their own shortcomings. Given the lack of representative data for most countries in the EU, we derive our measures from O*NET, following the same procedure as Acemoglu and Autor (2011). This combination of American and European data is not new in the analysis of task content of job, as it has been already used in Goos and Manning (2007), Goos et al. (2014). We take as a reference the 2010 release of O*NET, which corresponds to the last year in which we have information from the EU-SES. The weights used for the estimation of the task content are derived independently for each country using the EU-LFS, in order to minimize issues on the representativeness of EU-SES sample[7]. The decision to estimate separate task content at country level represents also a normative standpoint. Obtaining a single measure for the EU, as it was done in Goos et al. (2014), corresponds to treating the EU effectively as a single labor market.

We use several cross-walks to match occupations in O*NET (which use US codes) to the ISCO codes from EU-SES. Two features of occupation classification deserve thorough treatment. First, while Member States had to report occupations using ISCO codes, countries could choose the level of aggregation (two or three digit codes). As reported in Table A1 in the appendix, data at the two-digit level are available for all countries. As employing different levels of aggregation has a direct impact on measures of within occupation wage inequality, we collapsed three digit codes into two digits for all countries that report three digit occupations.

Second, in 2008, ILO released new standardized codes (ISCO-08), which were incorporated in the 2010 edition of EU-SES. ISCO-08 was introduced as a minor revision of the previous version (ISCO-88); however, it rearranges occupations across major groups, which means that categories are not strictly comparable. This does not represent an issue for the first part of the analysis, as the task content of occupations can be obtained independently for both codes; but it becomes problematic when we explore the relation between wage inequality within occupations and changes in employment. Available correspondence tables, or cross-walks, include many-to-many matches. The problem is more severe in countries that presented occupations at the two digit level. In the appendix, we detailed the problems for matching occupation codes and some indicators of the type and size of errors that one can expect to observe.

In addition to the variables presented above, at several points we control for workers characteristics. These characteristics include gender, age, educational level, tenure, whether the worker has a full time or a part time contract and firm characteristics: plant size (a dummy for establishments with over 250 employees) and industry of the firm. Descriptive statistics for these variables are presented in Table 1. The Table presents information on the sample sizes in the upper part, and then average values of these variables across occupations.

Table 1 presents clearly one of the main advantages of using the EU-SES: sample size. In total, we have access to over 25 million observations. Sample sizes differ by country due to sample design. As a consequence, sample moments need not be representative of the EU wide economy. The average number of observations in each occupation might be misguiding, as the distribution of workers in each occupation highly skewed to the right. Medians, reported

[7] In fact, for specific countries and years, some of the occupations in EU-SES did not have any observation.

Table 1: Sample characteristics

Year	2002		2006		2010	
N	5979339		10967088		10404855	
Avg. $N_{country}$	298967		522242		495469	
Avg. N_{occ}	11747	(18421)	18876	(51498)	12491	(23744)
Median N_{occ}	5120		6627		4926	
- Wage inequality						
p90/p10	3.73	(1.13)	3.72	(1.14)	3.61	(0.92)
p90/p50	2.10	(0.41)	2.07	(0.35)	2.06	(0.33)
p50/p10	1.75	(0.3)	1.77	(0.3)	1.73	(0.26)
- Personal char.						
% Female	0.42	(0.25)	0.44	(0.26)	0.44	(0.27)
% Age 20-60	0.95	(0.04)	0.94	(0.06)	0.93	(0.06)
% Univ. Studies	0.30	(0.33)	0.24	(0.31)	0.24	(0.29)
% Firms >250	0.48	(0.26)	0.48	(0.25)	0.50	(0.22)

Notes Data obtained from the EU-SES waves 2002-2010. Cyprus and Luxembourg excluded as no data on firm size was made available for these countries. Germany only presents data as of 2006. Standard deviations in parentheses.

under the mean, are a better indicator of central tendency and suggest that there are enough observations in each occupation to have meaningful calculations of dispersion measures.

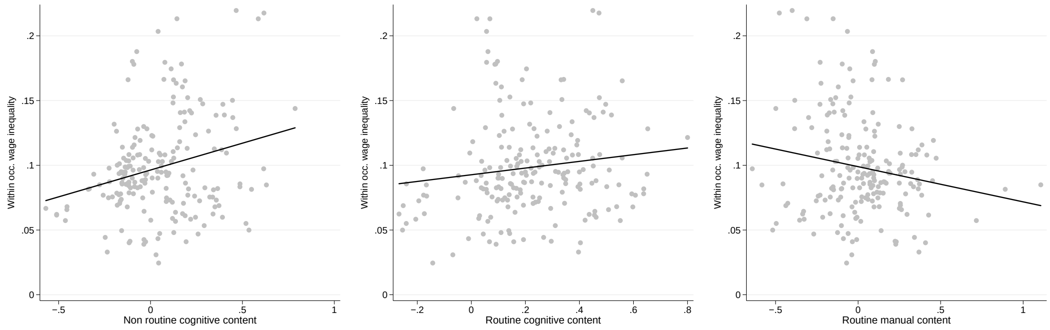
The remaining rows of Table 1 characterize occupations, as the averages were obtained by assigning the same weight to all observations. A few trends are worth noticing. First, wage inequality appears to be decreasing over time, both at the top and at the bottom, though neither change is statistically significant. The bottom rows show composition of the workforce across years. A value of .42 for female indicates that on average 42% of respondents within an occupation in 2002 were women. Table 1 presents some intuition of dispersion across countries and occupations, yet to obtain a better picture of the dispersion of characteristics across occupations, we added Figure A2 in the appendix. This figure plots the average values of these characteristics within occupations. In all sub figures, the rightmost plot includes more observations, which chiefly reflects the change in the classification codes described above.

4 From tasks to wage inequality

In Figure 1, we use the properties of the Theil index to decompose wage inequality into a part that corresponds to differences between occupations and a part that corresponds to differences within occupations. A correlation of these measures and task content of jobs provides an initial analysis of the relation between the two concepts. This exercise is performed in Figure 2. We estimate the contribution of occupations to total wage inequality at the regional level to increase the number of units, from country to regional level. Full evidence, with a wide array of controls and different measures of the importance of within wage inequality are presented in Table A2 in the appendix.

Regions with more non-routine cognitive and non-routine abstract tasks present greater within occupation wage inequality, even after including a large set of controls, country and year interactions. Manual task content, routine or not, appears to be related to lower within occupation wage inequality. A possible explanation of this result is that these occupations are usually found at the bottom of the income distribution, where minimum wages are binding and there is not much room for wage dispersion. Routine cognitive tasks are only weakly associated

Figure 2: Within components of Theil index and task content of jobs in European regions



Notes: Figures display the correlation between three measures of task content of jobs and the contribution of within occupation wage inequality to total wage inequality in NUTS-1 European regions. Based on data from EU-SES 2002, 2006 and 2010. Cyprus and Luxembourg excluded from the sample.

to within occupation wage inequality, as only in some specifications the coefficient is significant. Changing the dependent variable to the contribution of within occupation inequality to overall wage inequality in the region does not affect the results.

The analysis presented above has some shortcomings, though. First, and most obvious, EU-SES presents regional information only at a very aggregated level, NUTS-1, which for many countries, particularly in Central and Eastern Europe, corresponds to the entire country, which is much broader than commuting zones presented in Autor and Dorn (2009, 2013). Consequently, it provides only a rough approximation to local labor markets. If some occupations are more spread across cities regions than others, this could also affect the contribution of occupation to overall wage inequality. This example illustrates one difficulty with this approach. Second, if other workers' characteristics are correlated with occupations, then this approach cannot distinguish the contribution of these characteristics and occupations. Estimating measures of dispersion within cells of workers characteristics might not solve the problem, as for some cells there might not be enough observations.

In order to deal with these limitations, we provide an alternative approach to estimate the contribution of task content to within occupation wage inequality. Instead of analyzing the contribution of occupations to wage inequality within a region, we obtain measures of wage inequality for each occupation. Specifically, we will work with the logarithm of the 90/10, 90/50 and 50/10 percentiles of the income distribution within occupations. This approach provides more flexibility to control for differences in workforce characteristics within occupations.[8] After obtaining the ratio of percentiles, we proceed to analyze the relation between wage inequality at the occupation level and task content by estimating regressions of the following form:

$$log(y) = \alpha_0 + \beta RTI + D\gamma' + X\psi' + \epsilon\,,$$

where y represents different measures of inequality. β is the coefficient of interest and represents the effect of an increase in one standard deviation of RTI on wage inequality. D refers to a set of fixed effects that control for country and year. Finally, X represents additional controls for

[8] An alternative approach to control for workers' characteristics is to employ the Shapley decomposition. In this procedure, it is possible to estimate the contribution of occupations after other sources of heterogeneity have been removed. After all possible combinations of factors were considered, one obtains an average of the results. The procedure employed in this Chapter provides a more flexible approach to control for other factors and it is preferred.

workers characteristics.[9]

Table 2 presents results from such regression. Rows indicate the dependent variable and columns correspond to the set of controls included. Column 1 presents the base specification, where only fixed effects are included. The coefficient on RTI appears as significant and indicates a similar relation at the top and bottom of the income distribution. The negative sign is consistent with previous expectations. Routine intensive occupations are characterized by lower wage inequality. The size, however, is not large. An increase by one standard deviation in the task content results in a fall by between 7 and 10 percent in the ratio, depending on the specification. This effect comes solely from variation in occupations within country-year units.

In column 2, we expand the list of the controls by including average hourly wage, which acts as a proxy for the skill level. In columns 3 to 5, we add controls for workers' characteristics within the occupations. In column 3 includes the proportions of female, prime age, and tertiary educated workers, and the proportion of those working in large firms. Since wage inequality might not be related to levels, but with dispersion, we include the variance of these proportions in column 4. In column 5, we introduce both, proportions and variance.

Columns 6 and 7 provide an alternative approach to control for workers' characteristics. In these columns the dependent variables are the (90/50) and (50/10) ratio of residuals from a wage regression. This regression is run at the country-year level and includes additional control variables: age, gender, education, size, type of contract (part or full time) and industry (1 digit NACE). Unlike previous regressions, age, education and size present more than two categories. In column 7, we include also hourly wages as an additional control.

The relation between task content of occupations and wage dispersion is, in general, what we expected: More routine jobs exhibit lower wage dispersion. The effect is somehow weaker once we include controls for workers' mean characteristics, but still with a p-value smaller than .15. Dispersion in characteristics, though statistically significant, has only small effects on the variable of interest. If we control for characteristics with the help of a Mincerian wage regression, then all coefficients remain negative, significant and around the same size than the base specification.

In the middle and lower panel, we explore heterogeneous effects at the top and bottom of the distribution. Regressions reveal that the relation between wage dispersion and task content was mainly driven by the ratio between the 90^{th} and 50^{th} percentiles of the within occupation wage distribution. At the bottom, effects are slightly less robust, as they even change sign when mean worker characteristics are included as covariates, but not when we control for these characteristics in a first stage regression.

In the appendix, we tackle a number of potential concerns with this specification. First, we test for non-linear effects of RTI by including a quadratic term on RTI. Coefficient from this term is positive and significant, which implies a U-shape relation between RTI and within occupation wage inequality. Yet, the minimum value implied by coefficients falls outside of the range of RTI. If anything, the inclusion of a quadratic term emphasizes the higher wage inequality in non-routine intensive occupations. Table A3 displays coefficients from regressions, whereas a comparison of predicted values is presented in Figure A3. A second concern relates to the definition of occupations. In Table 2, we defined occupations on the basis of two digit ISCO-88 codes. In Tables A4, we estimate the same regressions, but classifying occupations

[9] One might be tempted to add occupation fixed effects as well. But, if we do so the coefficient on RTI would lack any significant interpretation, as there are no sources of variation left. Notice that excluding time and country effects might not be feasible. Time dummies capture the effect of business cycle, and there are reasons to believe that these effects might have affected wage inequality, but also the variation that is due to changes in the classification of occupations. Country dummies, on the other hand, act as proxies for institutional settings that might affect wage inequality.

Table 2: Relation between RTI and within occupation wage inequality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				log(p90/p10)			
β	-0.10***	-0.09***	-0.02**	-0.05***	-0.01	-0.08***	-0.07***
Se	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
N	1,862	1,862	1,862	1,862	1,862	1,862	1,862
R^2	0.45	0.48	0.54	0.56	0.62	0.52	0.54
				log(p90/p50)			
β	-0.05***	-0.04***	-0.02***	-0.03***	-0.02***	-0.05***	-0.04***
Se	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
N	1,862	1,862	1,862	1,862	1,862	1,862	1,862
R^2	0.41	0.42	0.44	0.49	0.54	0.46	0.48
				log(p50/p10)			
β	-0.06***	-0.04***	0.01	-0.03***	0.01*	-0.04***	-0.03***
Se	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
N	1,862	1,862	1,862	1,862	1,862	1,862	1,862
R^2	0.35	0.39	0.48	0.46	0.52	0.45	0.47
Year FE	Y	Y	Y	Y	Y	Y	Y
Country FE		Y	Y	Y	Y	Y	Y
Log(hwage)		Y	Y	Y	Y		Y
Ind. Charact.			Y		Y		
Var. Charact.				Y	Y		

Notes Table divided in three panels. In the upper, the dependent variable is the p90/p10 ratio of wages within the occupation. In the medium and lower panels, the dependent variables are the p90/p50 and the p50/p10 ratio respectively. In column 6 and 7, the ratio was calculated on the residuals from a Mincerian wage regression. Each column adds more controls to the specification. "Ind. charact." refers to the proportion of female, prime age and tertiary educated workers, and workers from large firms in each occupation. "Var. charact." indicates the variance of those variables. Standard errors in parentheses. *,**,*** indicate significance at the 10%, 5% and 1%.

on the basis of 3 digit ISCO codes. Since only some countries presented such disaggregated classification, this procedure reduced our sample size. Yet, results appear to be consistent. Finally, in Table A5, we test whether results are driven by a particular country or year. We proceed to exclude one country (or year) at a time and reestimate the relation. The coefficient on RTI proves to be quite stable.

In principle, patterns described in Table 2 could relate to both an increasing wage compression among routine jobs, for instance due to competition with machines lowering task returns, and to higher dispersion of productivity within non-routine employment. In Table 3 we begin to tackle this issue by regressing wage inequality on the different components of the RTI index. In column 1, each cell represents a different regression of within occupation wage inequality measures on each task. In columns 2 and 3 of Table 3 all measures of tasks content are included on a "horse race". Except for the measures of tasks, specifications in columns 2 and 3 of Table 3 are identical to those presented in columns 5 and 7 of Table 2, respectively.

Table 3 shows two points. First, coefficients from column 1 suggest that the most important divide is not between routine and non-routine, but rather between manual and non-manual. These regressions suggest that manual intensive jobs have a more compress wage distribution. One could expect this result if manual jobs were worse paid, but in these regressions we also include controls for hourly wages. Second, routine cognitive content appears to matter, but the sign of the effect runs against our expectations. Routine cognitive jobs are related to larger wage inequality. Table 3 indicates that the effects are driven mainly by wage inequality between the 50^{th} and 10^{th} percentile. Among top earners, routine cognitive task content appears as a poor predictor of dispersion.

Table 3: Specific tasks and wage inequality

	log(p90/p10)			log(p90/p50)			log(p50/p10)		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
NR Cog.	0.12***	0.00	0.07***	0.06***	-0.01	0.04***	0.07***	0.01	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
NR. Per.	0.07***	0.07***	0.04***	0.04***	0.04***	0.02***	0.03***	0.03***	0.02***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
NR. Man.	-0.09***	-0.02	-0.01	-0.05***	-0.01	-0.01	-0.03***	-0.01	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
R. Cog.	0.04***	0.05***	0.01**	0.00	0.01***	0.00	0.03***	0.03***	0.01***
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R. Man.	-0.04***	-0.01	0.00	-0.04***	-0.01	0.01	-0.01	0.00	-0.01
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
N	1862	1862	1862	1862	1862	1862	1862	1862	1862
R^2		0.65	0.59		0.55	0.51		0.55	0.52

Notes Table shows regressions of within occupation inequality measures on task content of jobs. In (1), each cell represents a separate regression. Columns (2) and (3) repeat the specifications from column (5) and column (7) of Table 2. All regressions include year and country fixed effects plus a control for average hourly wages. Standard errors in parentheses. *,**, *** represent significance at the 10%, 5% and 1% level.

Since routine jobs are not related to more equal wage distribution, the negative relation reported between RTI and wage inequality presented in Table 2 arises solely from non-routine jobs. As Table 3 presents beta coefficients, their values indicate the relative strength of the associations. The combination of non-routine personal and non-routine cognitive outweighs the importance of routine cognitive tasks in every specification. In short, routine cognitive jobs might present wage dispersion, but wage dispersion in non-routine jobs is larger, particularly at the top of the income distribution. This conclusion is aligned with our initial hypothesis: workers in non-routine intensive jobs face larger wage inequality.

4.1 Changes in occupational structure

Results from the previous section provide support to the claim that high values of wage inequality within occupations were associated to the task content of these occupations. The more non-routine intensive an occupation is, the larger the wage dispersion. A question that remains is whether this difference comes from some particular characteristic of occupations, for example if non-routine tasks are more sensible to differences in productivity, or if larger wage inequality arose as a result of reallocation of workers. As the number of non-routine workers increased, the number of bad matches in those occupations might have risen as well, even if the proportion remained constant.

In order to estimate the change in occupational structure, we complement EU-SES data with data from EU-LFS. In spite of the many advantages of employing EU-SES noted above, this database is collected less frequently and is only available since 2002. If we were to work with EU-SES data, the first year would be loss, as we would lack information on occupational changes. Moreover, as the 2010 edition of EU-SES employs a different classification of occupations it is excluded from the main analysis. EU-LFS, in contrast, provides information on the occupational structure for some EU member states since 1992 and for the majority of the EU since 1997. Data from EU-LFS has then two advantages to measure occupational changes. First, given its availability one can compute change before 2002; second, it allows to test whether different time horizons affect the result. In order to keep estimates comparable to previous tables, countries that included only one digit ISCO codes in the EU-LFS were excluded from further analysis, as this occupation classification is too broad.

Our preferred measure of change in employment at the occupational level is defined as the change in the employment share of an occupation between two periods. In short, our measure of changes in occupation j is given by:

$$\Delta_{j,t} = N_{j,t}/N_t - N_{j,t-5}/N_{t-5} \, ,$$

where $N_{j,t}$ represents employment in occupation j in period t and N, total employment. The selection of a five year period reflects a tradeoff between two characteristics of the data. First, changes in the occupational structure might be better reflected if a longer time span is the considered. Second, the longer the period under analysis, the shorter the list of countries covered by the sample.

Sample includes only 2002 and 2006 editions of the EU-SES as the 2010 edition includes a different classification of occupations. Given data availability on EU-SES, it is impossible to provide an error-free crosswalk of occupations, even if one were to convert three digits ISCO 08 into 1 digit ISCO 88. Moreover, given that new classification was developed to recognize new occupations, one cannot expect that the error will not be correlated with employment growth. Given these considerations, we opted to leave the 2010 edition of the EU-SES out of this analysis.[10] While the exclusion of EU-SES 2010 could affect outcome, the results presented in Table A5 suggest that differences would be small.

Table 4 displays the coefficients from the regressions of within occupation wage inequality on task content (β) and changes in occupation growth (δ).

Table 4 repeats specifications from Table 2. In columns 1 to 3, dependent variables are measures of wage dispersion; whereas in the remaining columns dispersion were obtained from residuals of a wage regression. All regressions include controls for country and year fixed effects.

Following earlier discussions, growing occupations tend to present greater wage dispersion, as shown in first and fourth columns. Change in dispersion appears to be driven by changes at the bottom of the wage distribution, while dispersion at the top appears to be less correlated to changes in employment. This pattern suggests that changes in occupation size were related to an increase in the number of low quality matches. Once task content is included in the regression, coefficient on changes in employment becomes not statistically significant. A possible explanation is that 5 years might be a too short period. Table A6 explores this possibility measuring changes over a 10 year period, though this specification limits the number of countries to those for which information on the occupational structure is available in 1992. If this long-term approach is considered, then we observe that employment changes remain significant, though their size decreases in absolute terms after the inclusion of RTI, and average wages in the occupation. In principle, these results could emerge from growing occupations having more wage dispersion, as much as from decaying occupations having less.

Both patterns are consistent with models such as Jung and Mercenier (2014). Yet, it becomes interesting to analyze if the sign of change affected the relation for policy considerations. Table A7 tests which direction matters most by including a dummy variable that takes the value of 1 if employment in the occupation grew in the previous five years and zero otherwise plus an interaction of this dummy with the employment change. The sign and size of coefficients suggest that growing occupations presented greater wage dispersion than shrinking occupations.

Coefficient on RTI remains significant in Table 4 and with the same sign as in Table 2. More non-routine occupations present a more disperse wage structure, though the effects are much

[10] In the appendix, we elaborate further on the difficulties involved with matching occupations and provide a measure of misclassification errors.

Table 4: Within occupation wage inequality and occupational change

	(1)	(2)	(3)	(4)	(5)	(6)
			log(p90/p10)			
β		-0.10***	-0.00		-0.08***	-0.02**
		(0.01)	(0.01)		(0.01)	(0.01)
Δ	3.23***	1.55	0.76	2.35***	1.01	0.49
	(1.19)	(1.09)	(0.95)	(0.90)	(0.82)	(0.74)
N	778	778	778	778	778	778
R^2	0.42	0.52	0.64	0.51	0.60	0.67
			log(p90/p50)			
β		-0.04***	-0.01*		-0.05***	-0.01***
		(0.00)	(0.01)		(0.00)	(0.00)
Δ	0.86	0.15	-0.10	0.90*	0.13	-0.12
	(0.64)	(0.62)	(0.59)	(0.54)	(0.50)	(0.47)
N	778	778	778	778	778	778
R^2	0.40	0.46	0.50	0.43	0.53	0.58
			log(p50/p10)			
β		-0.06***	0.01*		-0.03***	-0.00
		(0.00)	(0.01)		(0.00)	(0.00)
Δ	2.38***	1.41**	0.87	1.45***	0.87*	0.60
	(0.72)	(0.67)	(0.56)	(0.48)	(0.45)	(0.41)
N	778	778	778	778	778	778
R^2	0.33	0.44	0.61	0.48	0.55	0.62
Country FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Log(hwage)			Y			Y

Notes: The table presents the relation between changes in employment, task content and wage inequality. Each cell represents a different regression. In (1), (2) and (3) the dependent variable is within occupation wage inequality. In (4),(5) and (6), the dependent variable measures inequality in residuals. Standard errors presented in parentheses. *,**,*** indicate significance at the 10%, 5% and 1% level.

smaller once one considers that the demand for workers in those occupations grew. In Table A6, coefficients on RTI are of a similar magnitude, but not statistically different from zero. The fall in the significance level might be a side product of reducing the sample size.

5 Discussion

Literature on the task content of jobs has focused on the analysis of polarized growth of employment and wages. Occupations at the bottom and top of the income distribution have shown relative increase in their employment shares, whereas those at the top also presented an important wage growth. Without rejecting this hypothesis, we argue that such analysis of redistributive consequences of technological change is incomplete, as it lacks a within occupation dimension. From a theoretical perspective, such analysis is the logical conclusion of the premise of opening up occupations and exploring what people do in their jobs. From an empirical perspective, inequality within occupation has gained importance recently.

Greater wage dispersion in non-routine intensive occupations can result from several mechanisms. First, the type of tasks that workers undertake in such occupations might have changed over time as a consequence of the introduction of new technologies. If workers differed on their ability to adapt to these new tasks, then wage inequality in those occupations is bond to increase. Second, it follows from the definition of routine jobs that workers are better substitutes for each other in those jobs. In other words, distribution of productivity is more compressed, and so it should be the distribution of wages. Moreover, following the theory of offshoring, these occupations are also subject to more competition with foreign workers, which might further reduce level and dispersion of wages. Third, the reallocation process itself could create a more spread distribution of productivity in non-routine jobs. As the demand for workers in these occupations increases, new positions might be filled with workers that were better matched to other occupations, which in turn would increase inequality at the bottom of the income distribution.

Empirical evidence presented in this chapter suggests that within occupation wage inequality is related to the task content of jobs. Non-routine intensive jobs presented greater dispersion of wages in a broad set of regressions. The relation exists both at the top and at the bottom of the within occupation wage distributions. An increase in the non-routine content by one standard deviation increases the breach between the 90 to 10 decile ratio by a value between 2% and 10%, depending on how one chooses to control for differences in individual characteristics within occupations. In our preferred specification, where we measure dispersion in residuals, the coefficient indicates an increase of 8%. These results seem to confirm our first hypothesis.

We explored two possible reasons for this increment: a) workers in routine occupations having more similar productivities and consequently routine intensive jobs presenting a more compressed wage structure, and b) increase in the number of low quality matches following the higher demand for non-routine tasks. To test the first hypothesis, we regress our measures of wage dispersion on the task content of occupations. As expected, non-routine intensive occupations presented greater wage dispersion; however, so did routine cognitive intensive occupations. Workers in routine jobs might be better substitutes for each other, but only in the case of routine manual tasks, and even then, only in some specifications. Two possible reasons might stand behind the positive relation between routine cognitive content and wage dispersion. First, it might be a pure measurement issue. An indicator used to estimate routine cognitive task content, i.e. importance of being accurate, might also take high values in non-routine jobs. Analysts or engineers are examples of non-routine occupations where accuracy matters. Such an explanation was also presented in Autor and Dorn (2013) to indicate the increase in routine content at the top of the income distribution. Second, if technological progress has increased the productivity of high skilled workers, and they are now required to perform a broader variety of tasks, as in the model of Acemoglu and Autor (2011), then the increase in inequality would reflect a higher participation of high skilled workers in routine tasks. If this was the case, wage dispersion in these occupations would be expected to fall as more high skilled workers fill new positions.

Second, we explored whether wage dispersion in non-routine occupation resulted from the increase in demand and, particularly, of the increase in lower quality matches. Results on this issue are sensitive to the time horizon over which changes are computed. Within shorter time horizons, the effect is positive; occupations that grew faster show higher wage dispersion; though the effect becomes not significant in most specification after the inclusion of our measure of task content. Following our expectations, change in employment appears to be connected with greater dispersion at the bottom of the income distribution, but the effect is small after the controlling for task content. Over longer time horizons, the relation between changes in employment and within occupation wage dispersion is statistically significant, even after the inclusion of RTI among covariates. The fact that the coefficient on RTI becomes smaller after the inclusion of changes in size suggests that part of differences in wage dispersion that we attributed to RTI resulted not from the task content itself, but from the reallocation process. Notwithstanding, the fact that the coefficient on RTI remains significant indicates that tasks still have some explanatory power, especially in the case of wage dispersion at the top of income distribution. These results on the change in employment confirm the second hypothesis.

One limitation on the analysis of the relation between growth in occupation and within occupation wage inequality comes from classifications. In fact, given data limitations, we employed relatively broad groups, which might be growing either because more workers join a particular occupation, or because new occupations are classified within old codes. Before the release of ISCO-08, there was no formal category for web developers, and workers in those occupations were classified along system analysts and software developers under the code "computer specialists". The increase in size of computer specialists over time might reflect then an increasing variety of jobs performed in that occupation, which might lead to further differences in wages, for reasons not related to workers' productivity.

We postulate here that if task content remains significant this might be due to workers' differences in abilities to perform non-routine tasks; but, this is not the only explanation. Segregation of tasks across workers of the same occupation could lead to similar predictions (Deming 2015, Garicano and Hubbard 2016). An explanation based on task segregation relies on tasks being more segregated in non-routine jobs than in routine jobs, otherwise efficiency gains from segregation might not differ across job types (or might even run in the opposite direction). One caveat on this literature is that if tasks are divided on a qualitative basis, as in the case of senior lawyers taking more difficult cases, then differences reflect workers' abilities to perform tasks, and not necessarily task segregation. Empirical evidence on task dispersion is not conclusive either. In a sample of British workers, Akçomak et al. (2015) find more specialization in jobs at the middle of the income distribution; whereas in the sample of participants in the WageIndicator project, Visintin et al. (2015) high-skilled occupations present greater dispersion of tasks.

Unfortunately, EU-SES data do not allow to test these hypotheses directly, as it does not include information on the task content at the individual level. A first possibility to sidestep this shortcoming is to measure dispersion of task based on O*NET measures obtained for more disaggregated occupational codes. One could obtain a proxy for dispersion at the two digit level by recovering task content at the four digit level and calculating some measure of distance within the two digit level. However, without proper occupational weights, estimates from such approach give too much importance to smaller occupations, which could result in upward biased estimates of dispersion. A better alternative might be merging EU-SES information on earnings with other databases that present information on tasks, such as PIAAC. Yet, we do not expect results to change significantly. The analysis of task dispersion carried by Visintin et al. (2015) suggests that dispersion in tasks explains 5% to 7% of wage inequality, and only in some measures of wage inequality.

Yet, the arguments presented here present only the introduction to a larger discussion on the role of technological change and within occupation wage inequality. In fact, there is still much to learn on the division of tasks at the workplace, its causes and what implies in terms of wage inequality, within the occupation and in the economy. First, one might test directly on whether differences in productivity stand behind the differences in wage dispersion between non-routine intensive occupation, for example via simulating counterfactual wage distributions for each occupation. Second, as emphasized by the literature some occupations change their task content over time: secretaries today and in the 1990s might have only the name in common. Then, it would be interesting to analyze how wages reflected this transformation of the task content, both in terms of level and dispersion.

Whether wage inequality within occupations could be dealt with the help of the right policy package is still an open question. On the one hand, it seems necessary to address problems connected with the reallocation of workers. These problems refer to both correcting the number of bad matches, which might be reduced with longer unemployment benefits, and skilling workers. Linking unemployment benefits and re-skilling programs could help to reduce wage inequality within occupation. On the other hand, one cannot expect such policies to be a silver bullet. While these variables were connected to wage inequality, coefficients were relatively small; hence, the situation might not improve dramatically.

References

Acemoglu, D. and Autor, D.: 2011, Skills, tasks and technologies: Implications for employment and earnings, *Handbook of Labor Economics* **4**, 1043–1171.

Akçomak, S., Kok, S. and Rojas-Romagosa, H.: 2015, Technology, offshoring and the task-content of occupations: Evidence from the United Kingdom, *International Labour Review* **115**(2).

Akerman, A., Helpman, E., Itskhoki, O., Muendler, M.-A. and Redding, S.: 2013, Sources of wage inequality, *American Economic Review* **103**(3).

Altonji, J. G., Kahn, L. B. and Speer, J. D.: 2014, Trends in earnings differentials across college majors and the changing task composition of jobs, *American Economic Review* **104**(5), 387–93.

Autor, D. and Dorn, D.: 2009, This job is "getting old": Measuring changes in job opportunities using occupational age structure, *American Economic Review* **99**(2), 45–51.

Autor, D. and Dorn, D.: 2013, The growth of low-skill service jobs and the polarization of the US labor market, *American Economic Review* **103**(5), 1553–97.

Autor, D. and Handel, M.: 2013, Putting tasks to the test: Human capital, job tasks, and wages, *Journal of Labor Economics* **31**(2), S59–S96.

Autor, D., Katz, L. F. and Kearney, M. S.: 2006, The polarization of the US labor market, *American Economic Review* **96**(2), 189–194.

Autor, D., Levy, F. and Murnane, R. J.: 2003, The skill content of recent technological change: An empirical exploration, *Quarterly Journal of Economics* **118**(4), 1279–1333.

Carrillo-Tudela, C. and Visschers, L.: 2013, Unemployment and endogenous reallocation over the business cycle, *Discussion Papers 7124*, Institute for Study of Labor (IZA).

Cortes, G. M.: 2016, Where have the middle-wage workers gone? A study of polarization using panel data, *Journal of Labor Economics* **34**(1), 63–105.

Cowell, F.: 2011, *Measuring inequality*, 3 edn, Oxford University press, Oxford, United Kingdom.

Deming, D. J.: 2015, The growing importance of social skills in the labor market, *Working paper 21473*, National Bureau of Economic Research.

Firpo, S., Fortin, N. M. and Lemieux, T.: 2011, Occupational tasks and changes in the wage structure, *Discussion Papers 5542*, Institute for Study of Labor (IZA).

Frey, C. B. and Osborne, M.: 2013, *The future of employment: How susceptible are jobs to computerisation?*, Oxford Martin School, University of Oxford, Oxford, United Kingdom.

Garicano, L. and Hubbard, T. N.: 2016, The returns to knowledge hierarchies, *Journal of Law, Economics, and Organization* . Forthcoming.

Gathmann, C. and Schonberg, U.: 2010, How general is human capital? A task based approach, *Journal of Labor Economics* **28**(1), 1–49.

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

Goos, M., Manning, A. and Salomons, A.: 2014, Explaining job polarization: Routine-biased technological change and offshoring, *American Economic Review* **104**(8), 2509–2526.

Jung, J. and Mercenier, J.: 2014, Routinization-biased technical change and globalization: Understanding labor market polarization, *Economic Inquiry* **52**(4), 1446–1465.

Kim, C. and Sakamoto, A.: 2008, The rise of intra-occupational wage inequality in the United States, 1983 to 2002, *American Sociological Review* **73**(1), 129–157.

Marcolin, L., Miroudot, S. and Squicciarini, M.: 2016, GVCs, jobs and routine content of occupations, *OECD Trade Policy Papers* **187**.

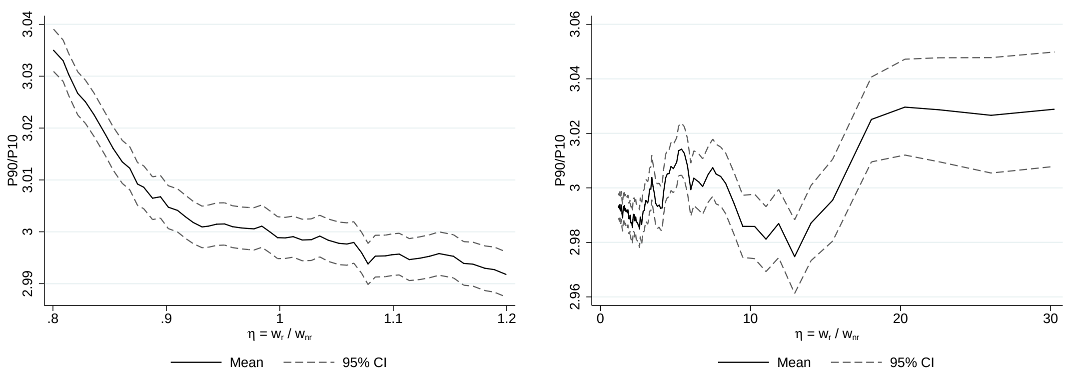
Oldenski, L.: 2012, The task composition of offshoring by US multinationals, *International Economics* **131**(1), 5 – 21.

Roy, A. D.: 1951, Some thoughts on the distribution of earnings, *Oxford Economic Papers* **3**(2), 135–146.

Visintin, S., Tijdens, K., Steinmetz, S. and de Pedraza, P.: 2015, Task implementation heterogeneity and wage dispersion, *IZA Journal of Labor Economics* **4**(1), 1–24.

Appendix

Figure A1: Relation between η and wage inequality in non-routine occupations



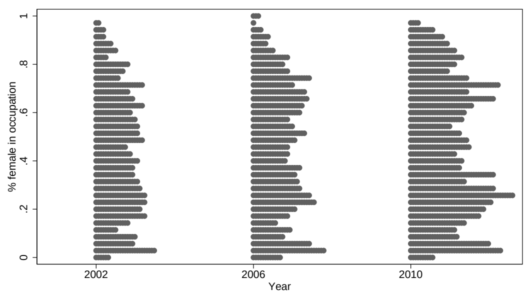
Notes: Panels show the relation between wage inequality in non-routine occupations and η^*, the cutoff productivity value. Each measure was obtained from a sample with 10000 observations and simulated a thousand times. According to Autor et al. (2003), the effects of technological change correspond to a fall in η, thus one should read the Figure from the right. Given that η also represent the ratio of wages, it is possible to argue that the left Figure represents a more interesting case. In both For values of η below 0.8, the increase in inequality measured by the ratio grows exponentially, these values were excluded to keep picture clean.

Table A1: Occupation classification in EU-SES

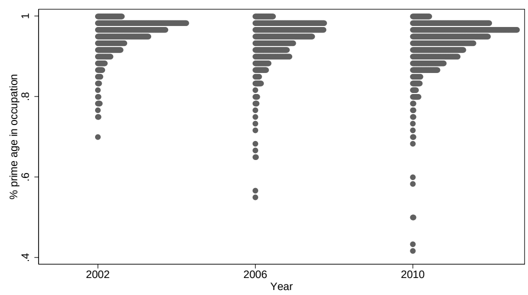
Occupation code	Country (Year)
a. 2 digits	Bulgaria,Belgium, Finland, France (2002), Germany (2006-2010), Greece (2006-2010), Hungary (2010), Italy, Netherlands, Norway (2002-2010), Poland, Portugal, Romania, Spain, Sweden (2010)
b. 3 digits	Cyprus, Czech Republic, Estonia, France (2006-2010), Greece (2002), Hungary (2002-2006), Latvia, Lithuania, Luxembourg, Norway (2006), Slovakia, Slovenia, Sweden (2002-2006), Great Britain

Notes: Years in parentheses indicate the information available for the country in those years. Countries with no parentheses use similar level of aggregation for occupations in every year.

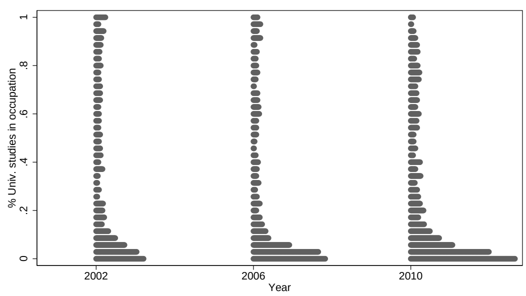
Figure A2: Dispersion of characteristics across occupations



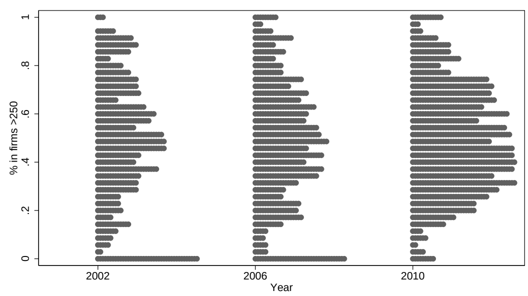
% Female



% Age 20-60



% Tert. studies



% in Firms >250 empl.

Notes: Graphs present the percentage of several variables within different two digit occupations for EU-SES divided by sample year. The larger number of observations in 2010 reflects the update in the classification system.

Figure A3: Non-linear effects of task content

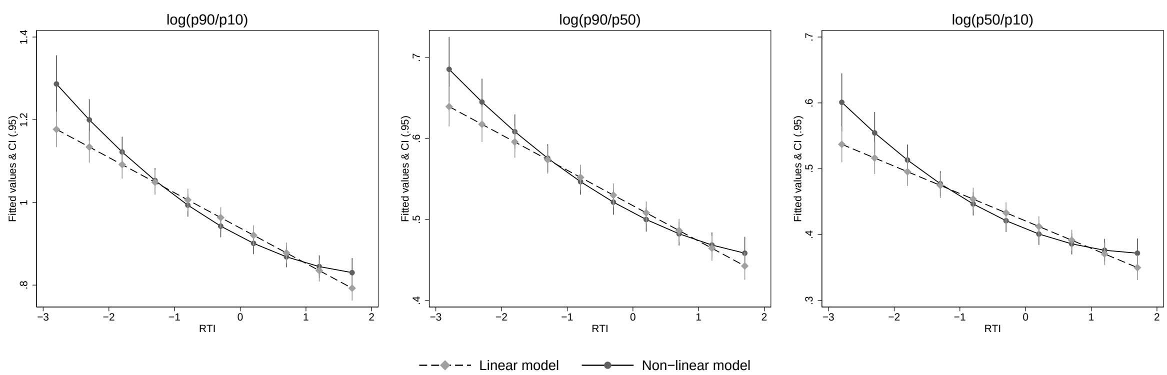


Table A2: Wage inequality in European regions and task content

	(1)	(2)	(3)	(4)
	Non Routine Cognitive			
β	0.03***	0.11***	-0.00	0.10*
SE	(0.01)	(0.03)	(0.03)	(0.06)
N	202	202	202	202
R^2	0.02	0.65	0.00	0.74
	Non-routine Personal			
β	0.02**	0.10***	-0.02	0.09
SE	(0.01)	(0.04)	(0.02)	(0.08)
N	202	202	202	202
R^2	0.01	0.64	0.00	0.74
	Non-routine Manual			
β	-0.02**	-0.07***	0.05***	-0.03
SE	(0.01)	(0.03)	(0.03)	(0.05)
N	202	202	202	202
R^2	0.01	0.63	0.02	0.73
	Routine cognitive			
β	0.02	-0.08**	0.04	-0.05
SE	(0.01)	(0.05)	(0.03)	(0.11)
N	202	202	202	202
R^2	0.01	0.61	0.01	0.73
	Routine Manual			
β	-0.03***	-0.07***	0.04**	-0.04
SE	(0.01)	(0.03)	(0.02)	(0.05)
N	202	202	202	202
R^2	0.02	0.65	0.01	0.73
F.E.	No	Yes	No	Yes

Notes: Each cell represents a different regression. The dependent variable in columns 1 and 2 is the within occupation wage inequality in European regions, whereas in columns 3 and 4 it is the percentage contribution to overall wage inequality to that region. Fixed effects included are interactions between country and year. Robust standard errors presented in parentheses. *,**,*** indicate significance at the 15%, 10% and 5% level.

Table A3: Non-linear effects tasks content on wage inequality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				log(p90/p10)			
RTI	-0.10***	-0.08***	-0.02**	-0.05***	-0.01	-0.08***	-0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
RTI*RTI	0.02***	0.02***	0.01**	0.01***	0.01***	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
N	1,842	1,842	1,842	1,842	1,842	1,842	1,842
R^2	0.46	0.49	0.55	0.57	0.63	0.53	0.54
				log(p90/p50)			
RTI	-0.04***	-0.04***	-0.02***	-0.02***	-0.02***	-0.04***	-0.04***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
RTI*RTI	0.01***	0.01***	0.01**	0.00	0.00*	0.01***	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
N	1,842	1,842	1,842	1,842	1,842	1,842	1,842
R^2	0.42	0.43	0.45	0.50	0.55	0.47	0.48
				log(p50/p10)			
RTI	-0.05***	-0.04***	0.01	-0.02***	0.01*	-0.03***	-0.03***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
RTI*RTI	0.01***	0.01***	0.01*	0.01***	0.01**	0.01***	0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
N	1,842	1,842	1,842	1,842	1,842	1,842	1,842
R^2	0.36	0.40	0.48	0.47	0.53	0.46	0.47
Year FE	Y	Y	Y	Y	Y	Y	Y
Country FE		Y	Y	Y	Y	Y	Y
Log(hwage)		Y	Y	Y	Y		Y
Ind. Charact.			Y		Y		
Var. Charact.				Y	Y		

Notes: Table repeats the specifications from Table 2 but it allows for non-linear effects of routine content on task inequality. Standard errors presented in parentheses. *,**,*** indicate significance at the 10%, 5% and 1% levels.

Table A4: Within occupation wage dispersion and task content

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				log(p90/p10)			
β	-0.06***	-0.03***	0.02***	-0.01*	0.03***	-0.05***	-0.03***
SE	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
N	2769	2769	2769	2769	2769	2769	2769
R^2	0.34	0.40	0.48	0.51	0.60	0.42	0.46
				log(p90/p50)			
β	-0.03***	-0.02***	0.00	-0.01	0.00	-0.03***	-0.02***
SE	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
N	2769	2769	2769	2769	2769	2769	2769
R^2	0.29	0.31	0.35	0.41	0.46	0.35	0.38
				log(p50/p10)			
β	-0.02***	-0.03***	-0.01***	0.02***	-0.00	0.02***	-0.02***
SE	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
N	2769	2769	2769	2769	2769	2769	2769
R^2	0.25	0.33	0.42	0.42	0.50	0.38	0.42
Year FE	Y	Y	Y	Y	Y	Y	Y
Country FE		Y	Y	Y	Y	Y	Y
Log(hwage)		Y	Y	Y	Y		Y
Ind. Charact.			Y		Y		
Var. Charact.				Y	Y		

Notes Table repeats the specifications from Table 2 for three digit ISCO codes. The list of country/years included corresponds to those presented in the bottom panel of Table A1. *,**,*** indicate significance at the 10%, 5% and 1% levels.

Table A5: Sensitivity to country inclusion

Excluded country / year	log(p90/p10)		log(p90/p50)		log(p50/p10)		N
BEL	-0.09***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,790
BGR	-0.09***	(0.01)	-0.05***	(0.00)	-0.04***	(0.00)	1,771
CZE	-0.09***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,771
DEU	-0.09***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,802
ESP	-0.09***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,772
EST	-0.09***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,772
FIN	-0.09***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,773
FRA	-0.09***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,773
GBR	-0.08***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,771
GRC	-0.08***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,757
HUN	-0.09***	(0.01)	-0.05***	(0.00)	-0.04***	(0.00)	1,772
ITA	-0.08***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,773
LTU	-0.09***	(0.01)	-0.05***	(0.00)	-0.04***	(0.00)	1,770
LVA	-0.08***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,771
NLD	-0.09***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,772
NOR	-0.09***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,770
POL	-0.09***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,771
PRT	-0.08***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,773
ROM	-0.06***	(0.01)	-0.03***	(0.00)	-0.03***	(0.00)	1,772
SVK	-0.08***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,772
SWE	-0.09***	(0.01)	-0.04***	(0.00)	-0.04***	(0.00)	1,772
2002	-0.05***	(0.01)	-0.03***	(0.00)	-0.02***	(0.00)	1,363
2006	-0.09***	(0.01)	-0.05***	(0.00)	-0.04***	(0.00)	1,305
2010	-0.08***	(0.01)	-0.03***	(0.00)	-0.04***	(0.01)	1,056

Notes: Each cell represents the coefficient on RTI from a different regression. Columns indicate the dependent variable and rows show the excluded country/year. Standard errors in parentheses. The last column presents the number of remaining observations. *,**,*** indicate significance at the 10%, 5% and 1%.

Table A6: Change in employment and wage dispersion

	(1)	(2)	(3)	(4)	(5)	(6)
			log(p90/p10)			
RTI		-0.12***	0.00		-0.09***	-0.01
		(0.01)	(0.01)		(0.01)	(0.01)
Δ	3.09***	1.62*	1.58**	2.54***	1.39**	1.36**
	(0.98)	(0.88)	(0.75)	(0.75)	(0.66)	(0.59)
N	430	430	430	430	430	430
R^2	0.20	0.37	0.55	0.23	0.41	0.54
			log(p90/p50)			
RTI		-0.04***	0.00		-0.05***	-0.01**
		(0.01)	(0.01)		(0.00)	(0.01)
Δ	1.06**	0.54	0.52	1.28***	0.62	0.61*
	(0.50)	(0.48)	(0.45)	(0.43)	(0.38)	(0.35)
N	430	430	430	430	430	430
R^2	0.23	0.31	0.40	0.20	0.38	0.49
			log(p50/p10)			
RTI		-0.07***	-0.00		-0.04***	-0.00
		(0.01)	(0.01)		(0.00)	(0.01)
Δ	2.03***	1.09*	1.06**	1.26***	0.77**	0.75**
	(0.62)	(0.55)	(0.47)	(0.41)	(0.38)	(0.35)
N	430	430	430	430	430	430
R^2	0.15	0.33	0.52	0.20	0.31	0.42

Notes The table repeats specifications from Table 4, but changes in employment are measured over a ten-year period. Sample restricted to countries that reported occupations in 1992: Belgium, Denmark, France, Great Britain, Greece, Netherlands, Portugal. Each cell represents a different regression. *,**,*** indicate significance at the 10%,5% and 1% level.

Table A7: Assymetric effects of employment changes and wage inequality

	log(90/10)			log(90/50)			log(50/10)		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
RTI		-0.08***	-0.01		-0.04***	-0.01**		-0.03***	0.00
		(0.01)	(0.01)		(0.00)	(0.01)		(0.00)	(0.00)
Δ	-2.09	-0.92	-3.05**	-1.81**	-1.14	-2.16***	-0.28	0.22	-0.88
	(1.39)	(1.33)	(1.18)	(0.85)	(0.81)	(0.74)	(0.89)	(0.89)	(0.84)
$(\Delta > 0)$	0.09***	0.03	0.02	0.05***	0.02	0.02	0.04***	0.01	0.01
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$\Delta * (\Delta > 0)$	1.92	1.76	5.59***	1.37	1.27	3.12***	0.55	0.49	2.47**
	(2.35)	(2.13)	(1.85)	(1.36)	(1.26)	(1.17)	(1.39)	(1.32)	(1.19)
N	778	778	778	778	778	778	778	778	778
R^2	0.52	0.60	0.68	0.44	0.53	0.59	778	778	778

Notes: Table presents additional specifications on the role of employment change. $(\Delta > 0)$ represents a dummy variable that takes the value of 1 if the change in employment is positive. All specification include year and country fixed effects. In column 3, a variable representing the average (log) hourly wage in the occupation is also included. Robust standard errors in parentheses. *,**,*** indicate significance at the 10%, 5% and 1% level.