UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Task implementation heterogeneity and wage dispersion

Visintin, S.; Tijdens, K.; Steinmetz, S.; de Pedraza, P.

DOI
10.1186/s40172-015-0036-2
Publication date
2015
Document Version
Final published version
Published in
IZA Journal of Labor Economics

Link to publication

Citation for published version (APA):
Visintin, S., Tijdens, K., Steinmetz, S., & de Pedraza, P. (2015). Task implementation heterogeneity and wage dispersion. *IZA Journal of Labor Economics*, *4*(20). https://doi.org/10.1186/s40172-015-0036-2

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:06 Sept 2026

Visintin et al. IZA Journal of Labor Economics (2015) 4:20
DOI 10.1186/s40172-015-0036-2

IZA Journal of Labor Economics
a SpringerOpen Journal

ORIGINAL ARTICLE Open Access

Task implementation heterogeneity and wage dispersion

Stefano Visintin[1*], Kea Tijdens[1], Stephanie Steinmetz[2] and Pablo de Pedraza[1]

* Correspondence:
stefano.visintin@gmail.com
[1]AIAS Amsterdam Institute for Advanced Labour Studies, University of Amsterdam, Amsterdam, Netherlands
Full list of author information is available at the end of the article

Abstract

Wage dispersion among observationally similar workers is still only partially unexplained by economists from both a theoretical and an empirical point of view. Given that jobs can be broken down into tasks, namely units of work activities producing output, we empirically test whether part of the observed variation in wages across similar individuals is related to differences in the intensity with which tasks are implemented. We then investigate whether the variety in task implementation shown across occupations is related to cross-occupation wage levels. We found that the variation in task implementation in different occupations is related both to within-occupation wage dispersion and to cross-occupation wage levels: workers in high-wage occupations are less defined around a typical worker than those in other occupations.

JEL codes: J22, J24, J31

Keywords: Tasks, Wages, Wage dispersion, Workers, Skills

1 Introduction

Despite long-standing efforts to understand labour income differences, theory and empirical studies can explain only a limited amount of wage dispersion. In particular, theoretical and empirical economic studies take into account only to a restricted extent the possibility that workers with similar productive characteristics receive different salaries. They do not consider a fundamental difference between similar workers, that is, the intensity with which they implement different work activities. Consequently, empirical approaches (based on human capital and search and matching theories) explain only a limited amount of the observed salary variation (Mortensen, 2005). As cross-worker differences in productivity are generally considered the main source of wage heterogeneity in economic literature (Becker, 2009), salary differences across workers with similar productive capacities remain mainly unexplained and cannot be fully accounted for by empirical models (Hornstein et al. 2011).

We observe that the dispersion of individual wages, which persists after controlling for workers' idiosyncrasies (e.g. educational level or professional experience) and job-specific characteristics (e.g. occupation and industry), accounts for more than two thirds of the total.[1] This measures the extent to which *a priori* similar workers receive different wages. This article extends the existing knowledge on this dispersion, which is limited by omitted variables of unobservable features. Our research approach brings the wage analysis to the occupational level of more than 430 occupations[2] and

Springer

© 2015 Visintin et al. **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

introduces measures of units of work activities that produce output, namely tasks. We explore whether task implementation heterogeneity is one of these unobservable features. This job-specific approach is justified because occupational boundaries can be used to group workers with a high level of homogeneity in terms of their productivity and their capacity to implement a job (skills). The idea that wages are related to job tasks is supported by the job analysis literature, where tasks are the basis of salary determination (Chatterji and Devlin, 2011).

Against this background, the main contribution of this paper lies in finding out whether and to what extent wage dispersion across similar workers can be explained by taking into account what had been an unobserved characteristic of the worker–job match: the tasks implemented in each occupation. In particular, the paper addresses two issues: (a) whether within-occupation wage dispersion is related to differences in task implementation and (b) whether cross-occupation wage dispersion is related to the level of task heterogeneity of occupations. The originality of the research is supported by the use of the WageIndicator data, which is a novel and unique web-based dataset (for more details, see Tijdens et al., 2010).[3] A specific strength of the WageIndicator is providing data on the intensity with which workers implement tasks. In order to assess the task intensity heterogeneity in each occupation, we set up measures of the role played by the 'typical worker' in each occupation, where 'typical worker' is defined as a worker who employs the most common way of performing an occupation. We set up these measures using the variation ratio, the entropy and the Gibbs index on individuals' information recorded in the dataset. The role of the typical worker in each occupation was used to explain part of the wage variation measured by the well-known Gini index and a wage dispersion index, once the effects of canonical human capital and firm-related factors were ruled out. In addition, we also observe how task implementation heterogeneity plays a role in explaining wage dispersion across occupations.[4]

Our exploration shows how heterogeneity in task implementation is related to wage dispersion among *a priori* similar workers when the latter are measured by the difference between the median and the reservation wage. A further exploration of our dataset and measures shows that high-wage (high-skilled) occupations have a high level of task implementation heterogeneity, whereas low-wage (low-skilled) occupations do not show any particular pattern. Our results suggests that the task approach to labour markets is a valuable tool that can extend and deepen existing knowledge on the sources of wage dispersion and that web survey data, given the high level of detail they provide for analysis, are a valuable data source to be increasingly considered in labour market research.

This article is organised as follows. Section 2 presents a review of the literature in the field of wage dispersion and tasks. Section 3 describes the data source and the methodological strategy implemented. Section 4 briefly depicts the extent of the wage dispersion in our dataset and presents the analysis results. Section 5 concludes this paper and offers some considerations and implications.

2 Literature review

2.1 Theories to explain wage dispersion

The determinants of wage dispersion among workers have been questioned from the outset of economic science. Among the numerous theories that deal with wage

differentials, human capital (HC) theory is possibly the most developed in terms of the quantity and quality of contributions.[5] Its essential message regarding wage dispersion is that individuals derive economic benefits as a return to a set of idiosyncratic characteristics, such as education, experience or ability: the higher the productive capacity, the higher the benefit. Accordingly, differences in productive capacities result in different wages across individuals. As a consequence, wage dispersion is entirely determined by differences in productivity across workers' human capital. Some extensions of the theory also consider other individual features that may modify the expected return given a specific productivity level. Models of discrimination suggest inequity for certain demographic groups on the basis of, for example, race, gender or religion (Becker, 2010). In addition, the theory of compensating differentials takes into account job- or occupation-related characteristics, especially undesirable attributes (e.g. high risk) but also geographical-related features (e.g. the attractiveness of the area where the job is located), and explains wage differences across jobs as the amount of income workers have to be offered to accept relatively unpleasant positions (for a complete discussion on this theory, see Rosen, 1986). According to all these theories, workers with similar individual and job-related features will achieve similar returns. Therefore, in HC, discrimination and compensating differential theories, wage disparities are limited to the heterogeneity across individuals and firm-specific features; in other words, similar workers receive similar wages.

An offshoot that is closely related to discrimination and compensating differential theories is the labour market segmentation theory. This theory, which is often said to compete with HC theory, considers the labour market as comprising distinct segments with different rules for wage setting and employment policies. Segments of the labour market where HC returns are lower than the rest of the market lead to wage disparities among similar workers (Lang and Dickens, 1992).[6]

Theories of search friction in labour market equilibrium go beyond the limitation presented above. By allowing for wage dispersion among workers with similar productive attributes and similar job characteristics, they add a further source of wage dispersion, namely labour market frictions (Mortensen 2005). The underlying idea of this research strand is that wage determination (among other labour market aspects, such as unemployment or job tenure) is influenced to a certain extent by the mechanism of the search and matching process between employer and employee. In particular, wage offers have some randomness from the individual jobseeker's point of view. Contributions build on the formal mathematical models of individual behaviour and labour market equilibrium developed by Stigler (1962), McCall (1970) and Mortensen (1970). Within the imposing literature that follows this line of reasoning, it is worth mentioning the scholars who formalised the main ideas on wage dispersion: Albrecht and Axell (1984), Eckstein and Wolpin (1990), Mortensen (1990, 2005) and Burdett and Mortensen (1998). According to this literature, wage dispersion across similar workers essentially arises from (a) search friction (workers looking for jobs can observe only a limited proportion of all wage offers in the market that match their skills) and (b) cross-firm differences in employee productivity (if the same worker is more productive in one firm than in another, the more productive firm will offer a higher salary).

HC and search theories assume that the same skill is remunerated uniformly across jobs and that wage differences arise from differences in workers' skills and market

frictions. Nevertheless, literature has also paid particular attention to the endogeneity between skills and wages, showing that workers' heterogeneity in skills leads to different rewards for the same skill across occupations. This introduced a new source of wage variation. The basic idea is that wages are not simply the result of operating skills in each occupation, but actually influence the skills of those who perform a specific job by determining the allocation of workers to jobs. Roy (1951) formalised this idea by developing a family of labour market models that treat the job–worker match as an endogenous process whereby workers decide upon their occupations on the basis of the return that each occupation offers to their set of skills. A job–worker match is seen as an outcome of a person's decision to choose the job for which he/she will achieve the highest earnings.[7]

2.2 Empirical approaches to analyse wage dispersion and common findings

Wage dispersion for homogeneous workers is one of the numerous issues (others being, for example, the return to skills or the unemployment duration) studied by an impressive amount of empirical research that operationalises the theories mentioned above. Empirical works mostly rely on two approaches: the estimation of (i) Mincerian wage equations and (ii) equilibrium model parameters. The former tool has often been implemented to test HC, discrimination and compensating differential theories, while the latter is the prevalent approach to make use of search theories and Roy-type models to empirically analyse labour markets. On the one hand, research relying on the Mincerian functional form is capable of testing how earning differences are determined by (a) individual heterogeneity in productivity (often using educational level or work experience as proxies for personal productive capacity) or (b) company productivity heterogeneity, measured by, for example, firm size or industry (see Heckman et al. 2003, as the most important contribution in the field). On the other hand, once properly calibrated, equilibrium models generate, among other things, a measure of wage dispersion that can be compared to actual data and serves as reference to understand wage dispersion sources.

Empirical findings show that, especially when using search models, most variation in wages can be explained by productivity heterogeneity, as opposed to worker's personal heterogeneity, firms' heterogeneity and market frictions. In other words, when comparing the contribution of different sources of wage variation, wages diverge mainly as a consequence of differentials in the productive capacity of the job–worker match as a result of individual productive characteristics. However, it should be pointed out that estimates of Mincerian wage regressions and equilibrium search models account for only a limited amount of the total observed wage variation. Mincerian models typically explain no more than 30% of salary variation across workers (Heckman et al., 2003, Mortensen, 2005). In addition, as earnings and the abovementioned variables can be determined simultaneously, endogeneity problems are a major drawback of this approach. Equilibrium models, supported by search friction and Roy-type theories, rely on a framework capable of handling this endogeneity. They allow for a comprehensive labour market analysis that takes into account such issues as unemployment, job duration and wage dispersion (see Eckstein and Van den Berg, 2007, for a detailed review). However, as Hornstein et al. (2011) point out, unless some parameters assume

implausible values, calibrated standard search models explain only a very small portion of observed wage dispersion. There are two main reasons for this: first, unobserved heterogeneity across individuals regarding individual, job and, particularly, job–worker match characteristics[8] that are not fully controlled by available data; and second, the theoretical search models fail to cover all aspects related to the wage determination mechanism. As a consequence, unsatisfactory results call for further investigation from both the theoretical and the empirical standpoint. Particularly, contributions to the latter aspect should be aimed at identifying measures of individual, job or job–worker match characteristics that are as yet unobservable.

2.3 Wage dispersion and the task approach

An approach that is capable of inspecting some aspects of the job–worker match is the task approach. It considers the tasks implemented in each occupation and relates them with workers' skills. The skill-task match is a characteristic of the job-worker match that has so far been only partially considered by the literature on wages determination. The importance of skills and tasks is however reflected in an emerging literature that is attracting increasing attention because of its capacity to explain changes in the labour market structure (Autor, 2013) by relating them to changes in the levels of routine activities within occupations over time. Acemoglu and Autor (2011) have developed a formal model in which wages emerge as the return to the labour factor in a production function and that is capable of dealing with the interaction among skill supplies, technological capabilities, and potential trade and offshoring opportunities. The assignment of workers to tasks in this model is determined by comparative advantages according to skills, as is the return to task. Although there is not yet a formal connection between this model and the search friction models, the task approach could contribute to the study of the equilibrium in search friction models and help explain wage dispersion among similar workers on the basis of the task assigned. Firpo et al. (2011) demonstrate how the causes of changes in the wage distribution over time can be captured by task measures, both theoretically by developing a model based on Roy's theories, and empirically.

According to Acemoglu and Autor (2011), a task is a unit of work activity that produces output. A task is also a job-specific characteristic since jobs can be decomposed into their tasks in a given time. On the other hand, a skill is a worker's endowment of capabilities for performing various tasks, which is closely related to his/her productive capacity. Consequently, the task intensity, measured for example as the frequency with which each worker implements a specific task, can be considered a characteristic of the skill–task, and consequently job–worker, match. It seems therefore reasonable to consider the intensities with which workers perform particular tasks as a proxy for the specific job–worker match productive capacity, and differences among workers as a proxy for some of the as yet unobserved heterogeneity. This means that task intensity measures could therefore describe some of the as yet unobservable heterogeneity across workers mentioned above. The main aim of this paper is to demonstrate the validity of this approach. In addition, an occupation-specific approach represents here a step forward in the task empirical literature, which mainly relies on tasks defined at the cross-occupational level. The latter approach allows tasks to be related to skills but

prevents tasks from being considered as units of input in the production process. In our occupation-specific framework, however, tasks can be directly considered as units of the production process.

The task approach is closely related to a psychological job analysis approach, which is aimed at determining which workers are a proper fit for specific jobs. In industrialised countries that have valid task lists for each job, job analysis is widely used as a systematic way to describe jobs and to classify them according to characteristics such as knowledge, responsibility and skills. Job evaluation is used to assess the value of the jobs that have been classified within an organisation. Both public and private organisations routinely rely on job evaluation as a means of constructing an appropriate pay scale (Chatterji and Devlin, 2011). While human capital approaches rely predominantly on individual characteristics to explain wage differentials, the job analysis approach bypasses the individual characteristics of the job holder in determining wages and, consequently, wage heterogeneity and mainly focuses on tasks implementation.

In order to complement existing empirical literature on wage variation with the job analysis perspective, we addressed jobs and tasks at an occupation level, which allowed for the observation of the intensity with which each worker performs the job-related tasks. This meant going a step further with respect to the task approach mentioned above, which relies on cross-occupation tasks. In fact, the units of production taken into consideration in the above mentioned task research are not job specific but general (e.g. tasks are labelled as routine or non-routine activities). Studying the tasks implemented within occupations allowed us to observe a worker–job characteristic and relate it to the principles of job analysis for which wages are determined according to the tasks performed. In this context, two approaches for measuring work activities in occupations can be distinguished. First is the generic approach, which measures the so-called generalised work activities (GWAs). However, GWAs are unavoidably vague, implying that they should be considered not as tasks, but as the underlying behavioural dimensions of work activities (Tippins and Hilton, 2010). Second is the job-specific approach, which measures highly specific tasks representing the most detailed descriptors commonly used in work analysis. In this approach, the work activities in a selected set of predefined occupations are typically investigated (Morgeson and Dierdorff, 2011; Dierdorff and Wilson, 2003). The generic work descriptions allow for non-occupation-specific sampling and can therefore be asked for in any survey targeting the workforce. In contrast, job-specific descriptions require either occupation-specific sampling or general sampling with an occupation-based screening question with follow-up questions. In this respect, the web-based data gathering methods, given their positive features such as timeliness and the high level of detail they allow for (Askitas and Zimmermann, 2015), have a clear advantage over other traditional techniques.

Given that tasks describe what had been an unobservable feature of the job–worker match and that they play a relevant role in salary determination, we hypothesise that their measure could contribute to a better explanation of wage variation across similar workers. In the first place, we believe that they might be able to explain within-occupational wage variation, a phenomenon that was largely unexplained across workers who share similar productive characteristics but perform tasks differently. Secondly, we believe that they might represent some occupation-specific feature, such as

the extent to which all workers perform their job similarly and that this feature could explain some of the cross-occupation wage variation.

3 Data and methodology

3.1 Source description

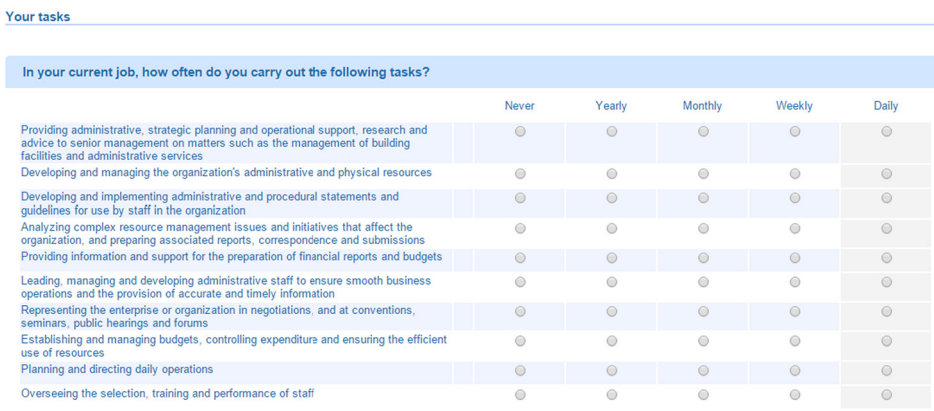
The data used in this research were derived from a multi-country volunteer web survey: the WageIndicator dataset (Tijdens et al., 2010). This survey is posted continuously on national WageIndicator websites (www.wageindicator.org). It is operational in 85 countries and attracts millions of visitors (25 million in 2014). The websites provide job-related content, labour law and minimum wage information, VIP wages, and a free 'salary check' presenting average occupational wages based on the web survey data. In return for free information, web visitors are asked to complete a questionnaire with a lottery prize incentive. Approximately 70,000 questionnaires are completed each year.[9]

The web survey contains detailed questions about demographic variables together with work-related ones such as occupation, industry, wage and firm size. Occupation is asked for not through an open text field with post-survey coding, but by means of self-identification using a database of approximately 1,700 occupational titles. Respondents can search the database by text string matching or with the help of a 3-step search tree. Respondents who are unable to identify their occupation are instructed to choose an occupational title that comes close to theirs. The 1,700 titles are classified according to the International Standard Classification of Occupations (ISCO–08), which is maintained by the International Labour Organisation (ILO) and is increasingly considered the global standard. The 1,700 titles are clustered into 433 occupational units at the ISCO's 4-digit level.

When respondents self-identify their occupation title, they are presented with a list of between 4 and 14 occupation-specific tasks related to their 4-digit occupational unit (designed by the ILO when defining the present ISCO–08; see Hunter 2009). They can then state how often they perform each of the tasks on the basis of five answer categories: *never*, *yearly*, *monthly*, *weekly* and *daily*. To provide an example of the measurement of task intensity, Fig. 1 shows a screenshot for the occupation 'Business services and administration managers', which distinguishes between 10 occupational tasks. As the tasks are occupation specific, they vary across occupations. A full list of occupations and corresponding tasks is available on the ILO website.[10]

Given our focus on occupations and occupation-specific tasks, an online survey seemed to be the best way to collect data on task implementation. In all other modes of data collection, the handling of a list of tasks for 433 occupations would have been too complicated and too time-consuming. Moreover, a tight fieldwork budget and other priorities very often hinder the inclusion of rather difficult and complex items in a questionnaire. Against this background, it is not surprising that task data are rarely available.

To our knowledge, only a few other datasets provide information about the task implementation of specific jobs; the best known are the North American O*NET and the German Qualification and Career Survey. The O*NET dataset provides, among other information, a measure of the relevance of (thousands of) tasks distributed across (hundreds of) occupations. However, this assessment is based on mixed data collection


Your tasks

In your current job, how often do you carry out the following tasks?

	Never	Yearly	Monthly	Weekly	Daily
Providing administrative, strategic planning and operational support, research and advice to senior management on matters such as the management of building facilities and administrative services					
Developing and managing the organization's administrative and physical resources					
Developing and implementing administrative and procedural statements and guidelines for use by staff in the organization					
Analyzing complex resource management issues and initiatives that affect the organization, and preparing associated reports, correspondence and submissions					
Providing information and support for the preparation of financial reports and budgets					
Leading, managing and developing administrative staff to ensure smooth business operations and the provision of accurate and timely information					
Representing the enterprise or organization in negotiations, and at conventions, seminars, public hearings and forums					
Establishing and managing budgets, controlling expenditure and ensuring the efficient use of resources					
Planning and directing daily operations					
Overseeing the selection, training and performance of staff					

Fig. 1 WageIndicator task question for the specific occupation 'Business services and administration managers' (ISCO–08 code 1219)

methods involving job incumbents and occupation experts, while the online survey approach, such as that of the WageIndicator is based on a bottom-up approach. The latter produces an intensity measure for each task listed within each occupation and captures individual variation within an occupation. This advantage is exploited in the present article by relating task intensities variation with wage variation. We do not exclude that similar exercises can be performed on the O*NET task rating data; however, the possibility of triangulation/cross-checking between the WageIndicator task data and the O*NET task ratings seems, at best, unlikely since they make use of different classifications for both occupations and tasks within occupations. The Qualification and Career Survey carried out by the German Federal Institute for Vocational Training and the Research Institute of the Federal Employment Service, provides information directly provided by employees over time. Nevertheless, tasks are not occupation specific but generic, and no information about the time spent on each activity is included.

Given the abovementioned characteristics, data collection through a web survey is the most suitable method when aiming to collect job-specific task information in the case of a large number of occupational titles (Tijdens et al., 2012).[11] The advantages are obvious: respondents can read the tasks themselves, thus reducing their time burden, and web surveys are often cheaper, as researchers do not have to consider fieldwork budgets. The authors regard the respondents' reported task frequencies as trustworthy because the tasks refer to activities that are considered to be the core of their occupation. Hence, respondents are familiar with a task because they perform it regularly, or they are not familiar with it because they never or hardly ever perform the task. It is our belief that respondents do not generalise on frequency of tasks from memory. In the case of the present analysis, the main advantage of the dataset was clearly the availability of comparable task intensities data at the 4-digit occupation level. The combination of such a specificity, the individuals' information provided and the large sample size allowed us to identify similar workers and, at the same time, to observe heterogeneities in salaries and task implementation. Moreover, it contributed to the improvement of empirical wage models by including what had been a rather unobserved variable that contributes to a better understanding of sources of wage heterogeneity.

Given the structural differences in national labour markets, we confined our analysis to the Dutch labour market, which is generally considered to be well-performing, as

evidenced by a prolonged period of low unemployment, high participation and stable wages (Gerritsen and Høj, 2013). These conditions reduced the possibility of any observed effect being due to some nation-specific shock in the market. In order to avoid a possible impact on the wage structure of changes in job tasks over time (see e.g. Acemoglu and Autor, 2011), we used pooled data for a limited time span (2013–14). Moreover, to guarantee the representativeness of the occupations considered, we selected the 100 most common 4-digit occupations, leading to a total number of 745 tasks for 5,230 observations (see Table 1 for a list and statistical descriptives of the 100 occupations). Finally, as self-reported wage data, which we used to estimate salary heterogeneity, are often noisy,[12] we relied on the following two-step filtering to improve the data: first, we excluded implausible data (values below €0.5 or above €2,000 per hour), and second, we concentrated on the 5th through 95th percentiles of the resulting distribution. Despite these restrictions, our range of values is in line with other works using survey wage data (e.g. Acemoglu and Autor, 2011, working on Current Population Survey data).

As the survey is voluntary, however, the question of representativeness needs to be addressed. Volunteer web surveys are an extreme case of a convenience sample which suffer from a combination of coverage, self-selection and non-response biases. Moreover, one person might be able to fill in the questionnaire several times. Previous research has shown that generally people who self-select into a (web) survey differ from those who do not with respect to socio-demographics but also in terms of time availability, web skills, or altruism to contribute to the project (Bethlehem 2010; Couper et al. 2007; Malhotra & Krosnick 2007). In case of the WageIndicator survey, which is used in the present study, previous findings show that high educated, younger and male respondents as well as people working predominantly in non-manual occupations are overrepresented in various country samples such as the Netherlands, Germany, Spain and Brazil (Bispinck et al. 2010; De Pedraza et al. 2010; Steinmetz et al. 2012). Moreover, it seems that the survey topic is an important criterion for the self-selection process. This leads to an increase of people with a particular interest in the topic, and those people might differ from those who have no strong interest in the topic. For instance, the study by Steinmetz et al. (2014) showed that respondents of the Dutch WageIndicator are more interested in wages and career advances. The authors, however, could also show that even if the web sample deviates to some extent from representative reference sample, the obtained conclusions are very often comparable. Against the background that literacy skills and internet penetration in the 100 most common Dutch occupations can be assumed to be at a medium or high educational level, the above described problems are less severe. The results for low-educated workers should be considered cautiously and be taken into account in approaches that are intended to go further than this exploratory study.

3.2 Analytical strategy

As our aim was to find out whether and, if so, to what extent a task approach could explain part of the unrevealed wage variation (a) across *a priori* similar individuals and (b) across occupations. Our empirical strategy is as follows. In the first step, we adjust the wage heterogeneity due to individuals' productive attributes, discrimination and

Table 1 Occupations and wage heterogeneity

ISCO code	Occ. name	n	Median wage	GINI index	ISCO code	Occ. name	n	Median wage	GINI index
4110	General office clerks	499	13.61	0.2093	9412	Kitchen helpers	36	4.77	0.2955
5223	Shop sales assistants	229	9.84	0.2357	2320	Vocational education teachers	35	16.96	0.1514
3313	Accounting associate professionals	170	16.04	0.2087	5322	Home-based personal care workers	35	15.56	0.1951
4312	Statistical, finance and insurance clerks	136	14.40	0.1749	9334	Shelf fillers	35	3.70	0.2852
3322	Commercial sales representatives	121	16.40	0.2071	3221	Nursing associate professionals	34	19.31	0.1166
4120	Secretaries (general)	116	14.43	0.1682	3259	Health associate professionals not elsewhere classified	34	14.10	0.1488
3343	Administrative and executive secretaries	113	16.29	0.1980	3423	Fitness and recreation instructors and program leaders	34	8.34	0.3273
5321	Health care assistants	109	14.66	0.2169	2330	Secondary education teachers	33	20.79	0.1361
3123	Construction supervisors	104	17.13	0.2078	2341	Primary school teachers	33	16.44	0.1403
2431	Advertising and marketing professionals	97	15.30	0.2024	9112	Cleaners and helpers in offices, hotels and other establishments	32	9.37	0.2024
4321	Stock clerks	93	12.70	0.2004	1330	Information and communications technology service managers	31	25.67	0.1949
2511	Systems analysts	92	19.49	0.4000	2523	Computer network professionals	31	20.21	0.1364
4226	Receptionists (general)	88	12.21	0.2124	2642	Journalists	31	12.12	0.2565
2221	Nursing professionals	85	17.40	0.1958	4311	Accounting and bookkeeping clerks	30	14.52	0.1747
2423	Personnel and careers professionals	80	19.98	0.2836	5244	Contact centre salespersons	29	10.18	0.2582
3512	Information and communications technology user support technicians	75	14.03	0.1745	2422	Policy administration professionals	27	19.43	0.1880
4222	Contact centre information clerks	75	11.26	0.1886	4323	Transport clerks	27	14.94	0.1677
5230	Cashiers and ticket clerks	75	9.67	0.3197	1211	Finance managers	26	31.14	0.3066
1120	Managing directors and chief executives	73	25.15	0.2569	1212	Human resource managers	25	23.85	0.1413

Table 1 Occupations and wage heterogeneity *(Continued)*

7231	Motor vehicle mechanics and repairers	71	11.65	0.2267	5151	Cleaning and housekeeping supervisors in offices, hotels and other establishments	25	11.51	0.1466
3118	Draughtspersons	67	15.47	0.1584	1412	Restaurant managers	24	16.39	0.1677
9111	Domestic cleaners and helpers	63	11.55	0.1646	3514	Web technicians	24	14.71	0.1115
2635	Social work and counselling professionals	62	17.53	0.2015	7212	Welders and flamecutters	24	14.19	0.1157
5120	Cooks	62	10.30	0.1930	3341	Office supervisors	23	18.04	0.1590
7412	Electrical mechanics and fitters	62	13.98	0.1503	5132	Bartenders	23	9.65	0.1899
1221	Sales and marketing managers	60	23.09	0.2614	5312	Teachers' aides	23	11.90	0.1262
2166	Graphic and multimedia designers	60	11.93	0.1953	2264	Physiotherapists	22	17.53	0.1713
3323	Buyers	58	15.73	0.1761	3342	Legal secretaries	22	15.59	0.1655
7233	Agricultural and industrial machinery mechanics and repairers	58	15.05	0.2741	4412	Mail carriers and sorting clerks	22	9.06	0.0781
1219	Business services and administration managers not elsewhere classified	56	22.47	0.1851	1420	Retail and wholesale trade managers	21	17.89	0.2110
5131	Waiters	55	8.08	0.2762	2522	Systems administrators	21	19.55	0.1752
9333	Freight handlers	55	10.84	0.1684	3251	Dental assistants and therapists	21	13.05	0.2523
4416	Personnel clerks	53	15.16	0.2523	3434	Chefs	21	14.57	0.0930
4419	Clerical support workers not elsewhere classified	53	12.70	0.1903	7512	Bakers, pastry-cooks and confectionery makers	21	13.68	0.1988
6113	Gardeners, horticultural and nursery growers	53	11.83	0.2482	1324	Supply, distribution and related managers	20	23.09	0.1782
2519	Software and applications developers and analysts not elsewhere classified	50	18.98	0.2342	5141	Hairdressers	20	9.36	0.2138
5222	Shop supervisors	47	11.55	0.1428	2141	Industrial and production engineers	19	17.32	0.1757
8332	Heavy truck and lorry drivers	47	12.12	0.1204	2359	Teaching professionals not elsewhere classified	19	17.85	0.1120

Table 1 Occupations and wage heterogeneity *(Continued)*

2411	Accountants	45	22.71	0.1959	7131	Painters and related workers	19	15.46	0.1009
5414	Security guards	45	13.65	0.1289	7543	Product graders and testers (excluding foods and beverages)	19	14.43	0.1439
2619	Legal professionals not elsewhere classified	43	16.18	0.3278	2513	Web and multimedia developers	18	13.67	0.2436
2310	University and higher education teachers	41	19.63	0.2250	2529	Database and network professionals not elsewhere classified	18	22.37	0.1672
5221	Shop keepers	41	13.17	0.1772	3113	Electrical engineering technicians	18	16.55	0.1097
3256	Medical assistants	40	16.29	0.1753	3240	Veterinary technicians and assistants	18	10.63	0.1290
3333	Employment agents and contractors	39	12.71	0.1191	4211	Bank tellers and related clerks	18	13.47	0.1131
7115	Carpenters and joiners	39	14.68	0.1411	7213	Sheet-metal workers	18	11.16	0.2467
8322	Car, taxi and van drivers	39	11.19	0.1583	3513	Computer network and systems technicians	17	16.94	0.1583
5311	Child care workers	37	13.64	0.2034	4214	Debt-collectors and related workers	17	14.60	0.1246
2514	Applications programmers	36	17.31	0.1196	4322	Production clerks	17	12.80	0.1565
3115	Mechanical engineering technicians	36	18.53	0.1421	7411	Building and related electricians	17	12.91	0.1261
	All occupations	5230	15.11	0.1888	(averages)				

Source: WageIndicator

Occupations listed in the table are those considered in our analysis, the n column specifies the number of individuals in each occupation in our sample. All Occupations values are average median wage and average Gini index. Wages are expressed as gross hourly values in Euro

compensating differentials. We therefore regress a common set of wage predictors on the log hourly wages and considered the exponential of the residuals for the rest of our analysis. The common wage predictors used here are a cubic of work experience,[13] years of education, and binary and categorical dummies for gender, migrant status, firm size and industry (according to the 1-digit NACE classification).[14] The exponential of the residuals of this first regression are individual wage measures that are orthogonal to the individual and firm productive characteristics. The dispersion of these residual wages is therefore independent with respect to the main individual and firm observable characteristics.

In the second step, we estimate within-occupational wage and task-implementation heterogeneities and test whether these measures are significantly related. We therefore correlate three task-heterogeneity indexes of each of the 100 selected occupations with two corresponding within-occupation wage dispersion indexes (see Section 3.3 for the full description of the measures implemented). Observing the relation between these measures, and in particular the share of explained wage heterogeneity variation, provides insights into whether a task approach can explain part of the unrevealed wage variation across *a priori* similar individuals. Using wage measures orthogonal to the individual and firm productive characteristics and considering workers of the same ISCO-08 4-digit disclosure allows us to address similar workers.

In the final step, we shift the analysis to a cross-occupation level and examine whether the observed heterogeneity in task implementation of each occupational group is related to its salary level. This is done by examining the relation between the three task heterogeneity indexes and the corresponding median wages received (i) by all workers in each occupation and (ii) by subgroups of workers differentiated by skill levels. This helps us better understand whether a task approach is capable of explaining some of the differences across occupations that determine wage diversity. In particular, it shows whether the role of standardisation in jobs is related to the salary structure.

3.3 Measures

The main measures we use in the framework presented in this paper are related to wage heterogeneity and task implementation heterogeneity. With respect to wage heterogeneity, we define two measures: the Gini index and the frictional wage dispersion (FWD) index. The Gini, which is the most widely accepted single-measure of wage inequality in economic literature, measures the inequality of a distribution. It ranges from 0 (= a continuous uniform distribution where all individuals receive the same income) to 1 (= the maximum inequality where a single individual obtains all available income and the rest of the population's income is 0).[15] The FWD is based on the difference between the average wage and the lowest (the reservation) wage paid in each group. This measure is similar to the mean–min ratio of Hornstein et al. (2011), which measures wage differentials induced by labour market search frictions.[16] Although slightly different, our measure shares the same elements, and we therefore refer to it as the FWD. The choice of the two measures is justified by the fact that they are differently respondent to the salary levels, and we know that this aspect could be of interest in our analysis.

We also include measures that would indicate the extent to which tasks are performed differently in each occupation. To proxy this within-occupational task heterogeneity, we rely on the relevance of the typical worker in each occupation, where the 'typical worker' is defined as a worker who employs the most common way of performing an occupation. In order to measure this dispersion consistently, we make use of three indexes: the variation ratio, the entropy and the Gibbs index, all computed on the self-assessed task intensities as explained below. The choice of the three measures is justified by the need to avoid possible index-specific bias. All three are measures of qualitative variation suited to assess dispersions in nominal distributions, which is the case of the self-assessed task intensities where respondents are asked how often they perform specific tasks and are given five frequency category options (see the previous section for details). Moreover, they are based on the weight of the most common class, which makes them suited to assess the idea of typical workers in an occupation. More concretely, the variation ratio is one of the most widely used measures of statistical dispersion in nominal distributions. It is defined as the proportion of cases that are not the mode (Wilcox, 1973). Entropy is a measure of the unpredictability of information content (based on the seminal work of Shannon, 1948) and is an established measure of dispersion for ordinal and nominal data (Vanhoy, 2008). It is a measure of uncertainty; in discrete distribution this is minimised when all cases belong to a single category and maximised in a uniform distribution (for further details, see Harris, 1982). Entropy makes use of the information provided by the probability distribution of the discrete variable and is elaborated as the negative of the sum of the category probabilities times the (base two) logarithms of the category probabilities. The Gibbs index was developed by Gibbs and Poston (1975)[17] and can be interpreted as one minus the likelihood that a random pair of samples will belong to the same category. It is therefore higher when the distribution across categories is uniform.

With regard to the interpretation of the indexes, for each task, all three measures produce low values when the proportion in the largest class of response is high; in other words, when the proportion of respondents performing a specific task with the same intensity is high. We first compute a task-specific index for each of the 745 tasks in our dataset. We then average the task heterogeneities corresponding to each occupation and obtain an occupation-specific value of tasks heterogeneity. By doing so, we create an occupation-specific index measuring the proportion of workers who perform their occupation in the same manner; that is, the index reflects how much workers are like the typical worker in each occupation. Low values of the indexes correspond to occupations with a large proportion of workers acting like the typical worker, whereas high values identify occupations where the weight of the typical worker is low, and therefore the task implementation is more heterogeneous.

4 Results

4.1 Observing wage heterogeneity

According to HC theories, workers with similar productive characteristics should earn similar salaries. Consequently, the dispersion of salaries in groups of similar workers should be lower than the total dispersion of salaries. Figure 2 provides a first descriptive elaboration of wage heterogeneity across workers with similar productive

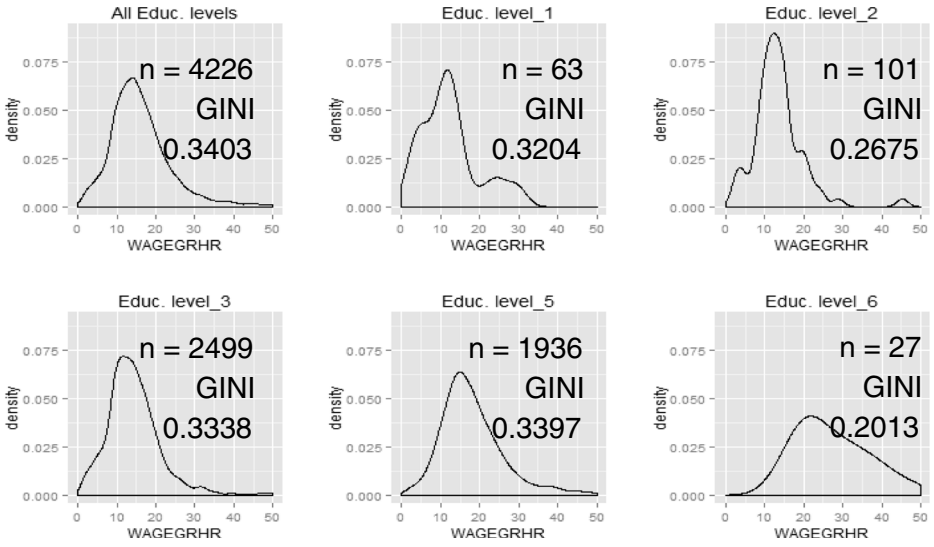


Fig. 2 Wage distribution per educational level (ISCED), the Netherlands (2013–14)

characteristics. It clearly shows that the shape of the hourly wage distribution is similar for the whole sample and groups of workers with similar characteristics, in this case similar educational levels. In this context it is worth noting that the Gini index varies by less than 2% between the dispersion across the whole population (All_Educ_levels, Gini index 0.34) and the dispersion across the most common group of workers (Educ_level_3, *upper secondary education*, Gini index 0.33). It shows that, even when workers are grouped according to their productive characteristics, in the case of education, wage dispersion persists.

When looking at the wage dispersion on the basis of a much more detailed classification criterion for workers, such as occupational groups, we can observe the Gini index varying widely across different occupations, ranging from 0.08 (Mail carriers and sorting clerks) to 0.40 (System analysts). Moreover, it becomes obvious (see Table 1) that the wage dispersion is much lower when based on 4-digit occupational groups (on average, within-occupational Gini = 0.19) rather than 5 educational groups (on average, within-educational Gini = 0.29). But more importantly, it is worth noting that within-occupational wage variation still accounts for, on average, approximately two thirds of the level measured for the total population (the average within-occupation Gini index is 0.19, whereas the measured Gini index of the whole population is 0.34). Thus, even after controlling for occupation, wage variation is still observable.

Against this backdrop, it is obvious that grouping workers according to occupation explains a consistent (although limited) amount of wage variation across individuals in the labour market. This seems logical, as the occupation variable includes information about both the workers' productivity and their capacity to implement specific job activities. It is worth noting that this information is often not explicitly formalised elsewhere, which makes the occupation information highly valuable when explaining wages. As a consequence, the remaining analyses presented in this paper were carried out at the 4-digit ISCO occupational level.

4.2 Wage and task heterogeneities

After establishing the persisting wage heterogeneity across similar workers and the usefulness of the occupational approach, we turned to the question whether within-occupational wage and task-implementation heterogeneities are significantly related. Therefore, as indicated in Section 3.2, we applied in a first step a typical Mincerian wage regression to adjust wage heterogeneity for individual human capital and job-specific characteristics. The results confirm common findings (see Table 2): experience and years of education are significant predictors of wage differences and therefore good proxies for human capital. Moreover, other individual and job-specific characteristics are also significantly related to salaries. When looking at the total wage variation explained, however, it is interesting to note that the consideration of all these characteristics accounts for around 33% of the total wage variation in the sample. This finding, in line with previous research, justifies the need to include measures that are capable of explaining wage variation across individuals to a larger extent.

In the second stage of the analysis, we use the (exponential of the) residuals of the regression as a proxy for the individual wages adjusted by individual productive attributes, discrimination and compensating differentials, to relate the task heterogeneity of each of the 100 selected occupations to wage dispersion indexes computed on the above-described proxy of workers' adjusted wages.

Figure 3 visualises the relation between the three heterogeneity indexes (variation ratio, entropy and Gibbs index) and the two wage dispersion indexes (Gini and frictional age dispersion). They represent the relation between job implementation heterogeneity

Table 2 Wage regression on individual productive attributes, discrimination and compensating differentials characteristics

Dependent variable: log of wage		Variable source/description
Experience	0.04491*****	Estimated years of work experience. From year of survey and year of first job.
	(0.00428)	
Experience^2	−0.001188*****	Square of the experience.
	(0.00023)	
Experience^3	0.00001****	Cubic of the experience.
	(0.00000)	
Years of Education	0.02523*****	Estimated years of education. From year of finishing education and year of birth (minus five).
	(0.00157)	
Female	−0.09495*****	Binary variable, 1 Female, 0 Male.
	(0.01226)	
Migrant	0.02228	Binary variable, 1 if not born in the country of residence, 0 otherwise.
	(0.02318)	
Industry dummies	*Yes*	Industries classified under NACE rev2.0, 1 digit disclosure (21 industries).
Firm size dummies	*Yes*	Ten firm-size groups
Year dummies	*Yes*	2013; 2014
(Intercept)	1.637*****	
	(0.1201)	
no of observations	2938	
R2	*0.326*	

Signif. codes: 0 '*****' 0.001 '****' 0.01 '***' 0.05 '**' 0.1 '*'

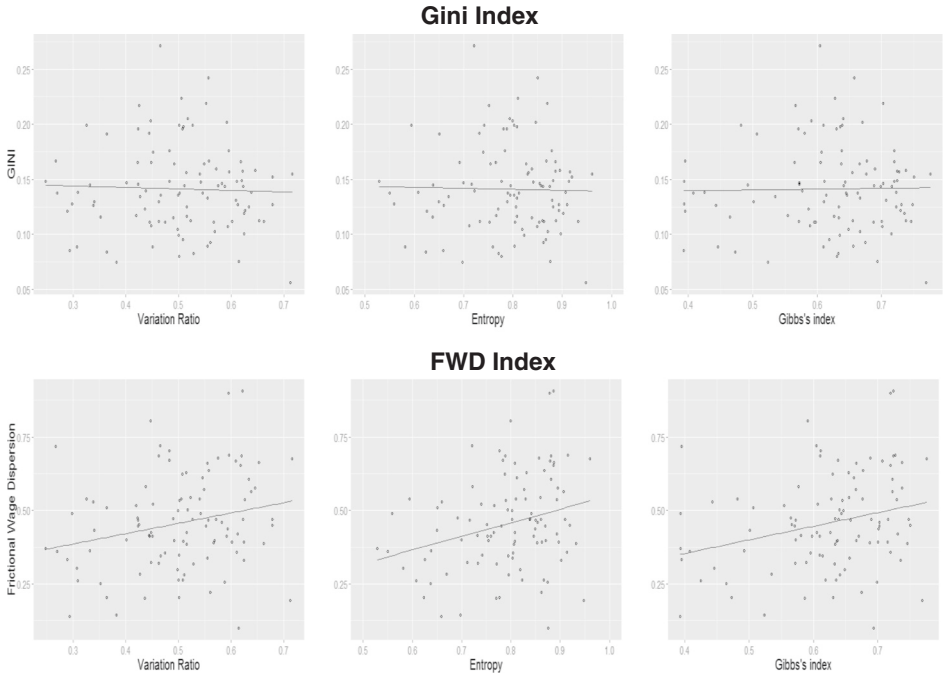


Fig. 3 Correlations between task heterogeneity indexes and the Gini Index and the frictional wage dispersion (FWD) index

and wage heterogeneity among *a priori* similar workers. The figures show that none of the three task heterogeneity indexes are related to the Gini index, but that they are positively related to the FWD index. This indicates that heterogeneity in each occupation implementation cannot explain within-occupation wage heterogeneity when the latter is measured with a more general wage dispersion measure like the Gini Index, but it is related to the wage variation measured by the FWD index.

The regression outcomes shown in Table 3 confirm the visual results. As expected, no significant effects can be found for the Gini index, whereas the FWD is positively and significantly related to all three task heterogeneity indexes. For a better understanding of this finding, it is worth recalling that all three task heterogeneity indexes measure the extent to which workers in the same occupations belong to a typical worker group. This implies that in occupations in which a large proportion of workers perform the same tasks with the same intensity, low levels of heterogeneity can be

Table 3 OLS results for cross-occupation task heterogeneity indexes on wage dispersion measures

	GINI			FWD		
	Variation	Entropy	Gibbs	Variation	Entropy	Gibbs
Heterogeneity index	−0.013	−0.016	0.007	0.351***	0.344***	0.464****
	(0.037)	(0.040)	(0.042)	(0.147)	(0.159)	(0.165)
(Intercept)	0.148*****	0.154*****	0.137*****	0.281*****	0.185	0.168
	(0.019)	(0.032)	(0.026)	(0.076)	(0.127)	(0.104)
N	99	99	99	99	99	99
R^2	0.001	0.002	0.003	0.055	0.046	0.075

Source: WageIndicator 2013–14, own calculations; signif. codes: 0 '*****' 0.001 '****' 0.01 '***' 0.05 '**' 0.1 '*' (standard errors are in parentheses)

measured, whereas in occupations in which the presence of typical workers is less relevant, the level of heterogeneity is higher. Looking at the results shown in Table 3 for the FWD, we can conclude that occupations in which typical workers are of less importance are those in which the internal wage disparities among *a priori* similar workers vary the most. In occupations in which workers perform their duties in a more standardised manner, those with similar productive characteristics receive similar wages.

In addition, it is interesting that around 5–7% of the wage dispersion, after controlling for individual and firm-related attributes within each occupation, can be explained by individual differences in task implementation. In sum, wage heterogeneity across similar workers, measured by the difference between the median and the minimum wage, can be partially explained by the heterogeneity in task implementation. Therefore, a task approach can be considered a useful tool to contribute to a further explanation of wage heterogeneity.

At this point, however, the question is why task heterogeneity seems to be related to the FWD but not to the Gini index even though the indexes are highly correlated (see Fig. 4). We believe that the answer might be related to the fact that the indexes differ with respect to their sensitivity to wage levels: whereas the FWD is more apt to account for wage dispersion at higher wage levels, the Gini index is more linear across the salary scale. This suggests that task heterogeneity can be related not only to the within-occupation wage heterogeneity but also to the salary scale.

In order to explore the possible relationship between task implementation heterogeneity within occupations and wage scales in more detail, we plot the task heterogeneity indexes and the log median wage in each occupation for high-skilled (orange) and medium-/low-skilled occupations (blue) (see Fig. 5).

For all task heterogeneity indices, a positive relation can be observed with the within-occupational median wage. This indicates that at the top of the salary scale, the mosaic of job task implementation is much more diverse and workers are less similar to the typical worker in such occupations. In contrast, the figure of the typical worker is clearly evident in low-paid jobs, where the heterogeneity indexes are lower. This possibly implies that high-wage occupations display more task heterogeneity, allowing for originality and inventiveness in performing high-wage jobs, whereas for medium- and low-wage salaries tasks are more standardised around the typical worker. Additionally, if we consider high-skilled and medium-/low-skilled occupations separately, the

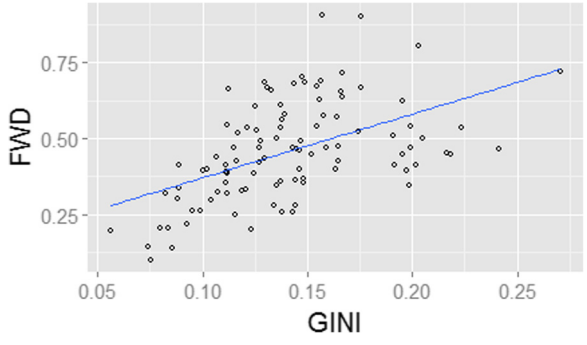


Fig. 4 Gini and frictional wage dispersion indexes correlation

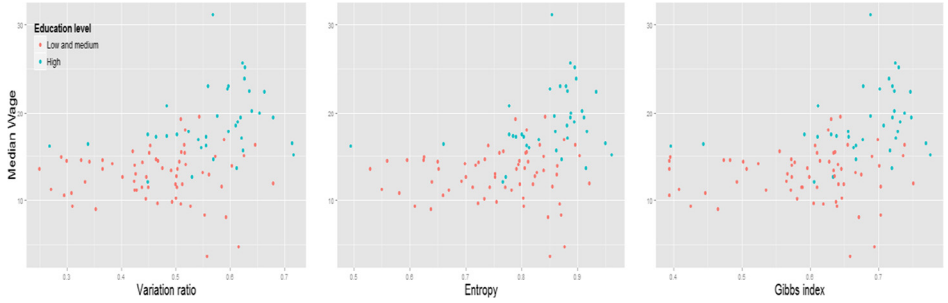


Fig. 5 Task heterogeneity indexes and median wages within occupations, across two skill levels

latter category (orange) shows no or only a weak relation between wages and any of the heterogeneity indexes. However, they do display a spectrum of task implementation heterogeneity (low as well as high), but for those occupations it is not related to the wage level. The opposite effect can be observed for high-skilled occupations: they are clustered around high values of heterogeneity that are also positively related to higher wage levels. This effect can be observed for all three heterogeneity indexes, which further strengthens and validates our results.

These exploratory findings are again confirmed by the regression results presented in Table 4. When the log median wages for all occupations are regressed on the task heterogeneity indexes, a significant positive relation can be observed for all three indexes. When specifying the regression for high-skilled and medium-/low-skilled occupations, the results confirm that median wages are positively related to task heterogeneity in high-skilled occupations (slightly significant), whereas no significant relation is found for the medium-/low-skilled ones. These outcomes suggest that the observed positive relation between task heterogeneity and wages for all occupations is driven by two forces, that is, by a composition effect (high-skilled occupations representing higher wage levels and task heterogeneity versus medium-/low-skilled occupations representing lower wage levels and task heterogeneity) and by a positive relation between wage levels and task heterogeneity within the high-skilled occupational group.

As emphasised at the beginning of this section (4.1), occupations define wage-dispersion groups more precisely. This analysis suggests that differences in the level of task heterogeneity across occupations could be one of the factors behind the observed classification effectiveness. In this context, the R-squared presented in Table 4 indicates that task heterogeneity accounts for approximately 9% of the cross-occupation wage variance, which again indicates the usefulness of a task approach in explaining wage differences.

5 Discussion and conclusion

This research empirically explores the explanatory contribution of tasks to wage dispersion across similar workers. This was done using a unique web-based dataset that allowed the measurement of the task implementation intensity at the individual level (10 occupation-specific tasks across hundreds of occupations and thousands of individuals) for the Dutch labour market. Previous research that used various theoretical and analytical approaches explained a consistent amount of wage differences for similar

Table 4 OLS results of cross-occupation task heterogeneity indexes on log median wage

	Variation			Entropy			Gibbs		
All occupations	0.867****			0.881****			0.993****		
	(0.282)			(0.305)			(0.318)		
High-skilled		0.623**			0.689**			0.779**	
		(0.345)			(0.378)			(0.356)	
Medium-/low-skilled			−0.223			−0.171			−0.1
			(0.361)			(0.373)			(0.391)
Intercept	2.234*****	2.566*****	2.627*****	1.974*****	2.337*****	2.653*****	2.051*****	2.392*****	2.589*****
	(0.145)	(0.199)	(0.172)	(0.243)	(0.321)	(0.285)	(0.201)	(0.279)	(0.234)
N	99	36	62	99	36	62	99	36	62
R^2	0.088	0.085	0.006	0.079	0.087	0.003	0.090	0.094	0.001

Source: WageIndicator 2013–14, own calculations signif. codes: 0 '*****' 0.001 '****' 0.01 '***' 0.05 '**' 0.1 '*' (standard errors are in parentheses)

workers, but left part of it unexplained. We have shown that a better understanding of wage heterogeneity among similar workers can be achieved by implementing a task approach.

Taking into account the limitations and complementarities of existing literature and tapping into new web-sourced data possibilities, we addressed two objectives: to explore (a) whether within-occupation wage dispersion is related to differences in task implementation intensity and (b) whether cross-occupation wage dispersion is related to the level of task heterogeneity of occupations. For that purpose, we used three indexes of task implementation heterogeneity: variation ratio, entropy and the Gibbs index. Wage disparities were measured by two indexes: the Gini index and the frictional wage dispersion (FWD) index, which considers the median and the reservation wages in each occupation. In order to compare *a priori* similar workers, the wage heterogeneity indexes were computed on the individual's wage portion orthogonal to personal and firm productive characteristics. In addition, task and wage heterogeneity indexes were calculated and compared at occupational level.

Our findings show that when using the FWD as a measure of wage distribution disparity within each occupation, a positive significant relation between wage and task implementation heterogeneities can be observed. This confirms that task implementation intensities can explain some of the wage differences across workers who should earn similar salaries, given their productive and job-specific characteristics. According to our estimates, task intensity heterogeneity explains 5–7% of the residual wage disparities after controlling for individual and firm-related productive characteristics. These results cannot be confirmed using the Gini index. This implies that, considering the difference between the Gini coefficient and the FWD index, which mainly concerns the weight of high salaries in computing the latter (differences in high-wage occupations produce larger FWD values than do differences in low-wage jobs), task implementation heterogeneity might be related not only to wage heterogeneity but also to wage levels. Therefore, we compared the 100 occupations' median wage with their task implementation heterogeneity measures and found a positive association: high median wages correspond to the high levels of task heterogeneity indexes. Moreover, this relation is more consistent when examining high-skilled occupations and when comparing high-skilled with medium-/low-skilled jobs.

In sum, our findings suggest that measuring task intensities helps to (marginally but significantly) explain wage differentials. This supports the idea already emphasised by the job analysis approach that labour market theories could be extended towards an integration of job tasks. Therefore, task-related theories, which have demonstrated their ability to explain employment structure and evolution over time, can also help in explaining wage variation.

All three heterogeneity measures that we implemented (variation ration, entropy and the Gibbs index) quantify how much workers are like the typical worker in each occupation. Therefore, the positive relation between within-occupation wage disparities and the relevance of the typical worker in each occupation observed indicates that the higher the proportion of workers who are different from the typical worker, the higher the wage dispersion.

Furthermore, we observed how the typical worker has a higher relevance in some low- and medium-skilled occupations, whereas high-skilled occupations are more diversified in

their implementation. Medium- and low-skilled occupations do not show any particular pattern of task heterogeneity when this dimension is related to wage levels.

The fact that well-paid and highly skilled occupations are less defined around a typical worker suggests that non-standardised roles require higher skills and are more productive. This is in line with the approach that explains changes in the labour market structure according to a routine/non-routine classification of the tasks performed in jobs. In this literature, the growth of the labour market share of some jobs is positively related to their low level of 'routinisation'. According to this approach, a consistent proportion of jobs that undergo growth for this reason are highly skilled occupations. Our results suggest that, in addition to explaining changes in the employment structure and wage levels according to the role of routine and non-routine tasks, this approach could also explain wage level differences across occupations. Further investigation should formulate and test these hypotheses.

6 Conclusions

In conclusion, the intensity of task implementation influences an individual's wage. Indeed, *a priori* similar persons (as measured by occupation, education and experience, and firm sector and size) earn different wages according, to a certain extent, to the difference in the intensity of tasks implemented. In addition, the level of heterogeneity in within-occupation task implementation is higher in high-salary occupations. Furthermore, cross-occupational wage differences can be explained by the different degrees of relevance of the typical worker in each occupation.

Endnotes

[1]See Section 2 for further details

[2]This disclosure corresponds to the most detailed occupational classification available, the ISCO–08 classification at 4 digits. See Section 3 for further details.

[3]We acknowledge the limitation of the dataset (see section 3.1) and therefore consider our findings to be exploratory rather than conclusive.

[4]Given the structural differences in national labour markets and the evolving division of labour, we confined the implementation of the approach to a single national market (the Netherlands) for a limited period (2013–14).

[5]Sweetland (1996) dates the roots of this theoretical branch back to Adam Smith, John Stuart Mill and Alfred Marshal. The record of studies and the attention received by HC theory is outstanding and includes Nobel prizes awarded to scholars connected to the field (Becker, 2009).

[6]Consequently, empirical studies that support this theory (e.g. Battisti, 2013) explain differences in returns to similar human capital as being exclusively the consequence of job characteristic disparities reflecting job segments.

[7]Empirical evidence shows that given each person's comparative advantages (skills heterogeneity), in the resulting allocation some skills are rewarded more than others across jobs, and consequently workers can earn more than others with similar skills. This process can be better understood when observing some empirical evidence based on Roy's models (e.g. Heckman and Sheinkman, 1987; Gibbons et al., 2005), which introduced a new source of wage heterogeneity variation that, as in the case of search models, originates from the worker–job match.

[8]Hornstein et al. (2007) stress how unobservable heterogeneity concerns the search friction aspects of the match.

[9]'Completed questionnaires' are those filled in by respondents who have provided information about at least their wage, gender, time of first job, occupation title and level of education. In 2014, there were approximately 240,000 completed and uncompleted questionnaires.

[10]http://www.ilo.org/public/english/bureau/stat/isco/isco08/

[11]For a general discussion on the advantages and disadvantages of using online data in the field of labour economics, see Kurekova et al. (2014).

[12]This is often due to non-response and under-reporting bias in high incomes (see Piketty and Saez, 2003) and by income-consumption disparities in low incomes (Meyer and Sullivan, 2008). As noted by Acemoglu and Autor (2011), these discrepancies can be due to measurement errors, under-reporting and transfer of income among individuals.

[13]The third degree of the polynomial was chosen following the main literature references in wage regression.

[14]The persistence of inter-industry wage differences for similar workers is one of the most pervading phenomena in labour economics (Abowd et al., 2012).

[15]The most recent World Bank (http://data.worldbank.org/indicator/SI.POV.GINI) estimates of the Gini index for the whole population in the Netherlands are 30.3 (2007) and 28.9 (2010).

[16]Hornstein et al. (2011) analytically derive the mean-min ratio from three search friction models.

[17]Here, we computed the first of the differentiation measures proposed in their article 'The unstandardised index M1' (p. 471).

Competing interests
All authors declare to observe the IZA Guiding Principles of Research Integrity. None of the authors have have any competing interests in the manuscript.

Acknowledgements
The present work received the financial support of the Eduworks Marie Curie Initial Training Network Project (PITN-GA-2013-608311) of the European Commission's 7th Framework Program. All work reported in this article has benefited from discussions and collaborations developed within the project. The authors also acknowledge the WageIndicator Foundation to provide the data used in this paper. The Foundation bears no responsibility for the analyses or interpretation of the data reported here. The authors would also like to show their gratitude to Brian Fabo and Tomáš Kabina for their help in preparing and to the anonymous reviewers for their insights.
Responsible editor: Klaus F Zimmermann

Author details
[1]AIAS Amsterdam Institute for Advanced Labour Studies, University of Amsterdam, Amsterdam, Netherlands. [2]Faculty of Social Science, University of Amsterdam, Amsterdam, Netherlands.

Received: 30 August 2015 Accepted: 20 October 2015
Published online: 05 November 2015

References

Abowd JM, Kramarz F, Lengermann P, McKinney KL, Roux S (2012) Persistent inter-industry wage differences: rent sharing and opportunity costs. IZA Journal of Labor Economics 1(1):1–25
Acemoglu, D. and Autor, D. (2011) Skills, tasks and technologies: Implications for employment and earnings, in: David Card and Orley Ashenfelter (eds), Handbook of Labor Economics, Elsevier, 2011, Volume 4, Part B, 1043–171
Albrecht JW, Axell B (1984) An equilibrium model of search unemployment. J Polit Econ 92:824–840
Askitas N, Zimmermann KF (2015) The internet as a data source for advancement in social sciences. Int J Manpow 36(1):2–12
Autor D (2013) The 'task approach' to labour markets: an overview. Journal of Labour Market Research 46:185–199
Battisti M (2013) Reassessing segmentation in the labour market: an application for Italy 1995–2004. Bull Econ Res 65(s1):s38–s55

Becker GS (2009) Human capital: a theoretical and empirical analysis, with special reference to education. University of Chicago Press, Chicago
Becker GS (2010) The economics of discrimination. University of Chicago Press, Chicago
Bethlehem J (2010) Selection bias in web surveys. Int Stat Rev 78:161–188
Bispinck R, Dribbusch H, Öz F (2010) Das Projekt LohnSpiegel: Tatsächlich gezahlte Löhne und Gehälter. WSI-Mitteilungen 63(1):42–49
Burdett K, Mortensen DT (1998) Wage differentials, employer size, and unemployment. Int Econ Rev 39:257–273
Chatterji M, Devlin S (2011) How heavy is a job?: A critical survey of job evaluation as a payment device. University of Dundee, Economic Studies, vol 259, Discussion Papers
Couper M, Kapteyn A, Schonlau M, Winter J (2007) Noncoverage and nonresponse in an Internet survey. Soc Sci Res 36:131–148
De Pedraza, P., Tijdens K., Muñoz de Bustillo R., and Steinmetz S (2010) A Spanish continuous voluntary web survey: Sample bias, weights and efficiency of weights. Revista Española de Investigaciones Sociológicas N° 131 (Julio–Septiembre 2010), pp. 109–130. ISSN 0210–5233 (print), 1988–5903.
Dierdorff E, Wilson W (2003) A meta-analysis of job analysis reliability. J Appl Psychol 88:635–646
Eckstein Z, Van den Berg GJ (2007) Empirical labor search: a survey. J Econ 136(2):531–564
Eckstein Z, Wolpin KI (1990) Estimating a market equilibrium search model from panel data on individuals. Econometrica: Journal of the Econometric Society 58:783–808
Firpo S, Fortin NM, Lemieux T (2011) Occupational tasks and changes in the wage structure, vol 5542, Discussion paper series//Forschungsinstitut zur Zukunft der Arbeit
Gerritsen M, Høj J (2013) The Dutch labour market: preparing for the future, vol 1012, OECD Economics Department Working Papers. OECD Publishing, Paris
Gibbons R, Katz LF, Lemieux T, Parent D (2005) Comparative advantage, learning, and sectoral wage determination. J Labor Econ 23(4):681–724
Gibbs JP, Poston DL (1975) The division of labor: conceptualization and related measures. Social Forces 53(3):468–476
Harris B (1982) Entropy, Encyclopedia of statistical sciences
Heckman JJ, Scheinkman J (1987) The importance of bundling in a Gorman–Lancaster model of earnings. Rev Econ Stud 54(2):243–255
Heckman L, Todd, L (2003) Fifty years of Mincer earnings regressions. NBER working papers series No 9732
Hornstein A, Krusell P, Violante GL (2011) Frictional wage dispersion in search models: a quantitative assessment. Am Econ Rev 101(7):2873–2898
Hunter D (2009) ISCO–08 Draft definitions. ILO Bureau of Statistics, Geneva
Kurekova, L., Beblavy, M., and Thum, A. (2014) Using internet data to analyse the labour market: A methodological enquiry. IZA Discussion Paper 8555, IZA, Bonn. Available at SSRN: http://ssrn.com/abstract=2514748
Lang K, Dickens WT (1992) Labor market segmentation, wage dispersion and unemployment. No. w4073. National Bureau of Economic Research
Malhotra N, Krosnick J (2007) The effect of survey mode and sampling on inferences about political attitudes and behaviour: comparing the 2000 and 2004 ANES to Internet surveys with nonprobability samples. Polit Anal 15:286–324
McCall JJ (1970) Economics of information and job search. Q J Econ 84:113–126
Meyer BD, Sullivan JX (2008) Changes in the consumption, income, and well-being of single mother headed families. Am Econ Rev 98(5):2221–2241
Morgeson FP, Dierdorff EC (2011) Work analysis: from technique to theory. In: Zedeck S (ed) APA Handbook of industrial and organizational psychology, vol 2. American Psychological Association, Washington DC, pp 3–41
Mortensen DT (1970) A theory of wage and employment dynamics. In: Phelps ES et al (eds) The microeconomic foundations of employment and inflation theory. Norton, New York
Mortensen DT (1990) Equilibrium wage distributions: a synthesis. In: Hartog J et al (eds) Panel data and labour market studies. North-Holland, Amsterdam
Mortensen DT (2005) Wage dispersion: Why are similar workers paid differently? MIT Press, Cambrige
Piketty T, Saez E (2003) Income inequality in the United States, 1913–1998. Q J Econ 118:1–39
Rosen S (1986) The theory of equalizing differences, vol 1, Handbook of labor economics. Elsevier, New York, pp 641–692
Roy AD (1951) Some thoughts on the distribution of earnings. Oxf Econ Pap 3(2):135–146
Shannon CE (1948) A note on the concept of entropy. Bell System Tech J 27:379–423
Steinmetz S, Raess D, Tijdens K, de Pedraza P (2012) Measuring wages worldwide - exploring the potentials and constraints of volunteer web surveys. In: Sappleton N (ed) Advancing social and business research methods with new media technologies. IGI Global, Hershey, PA
Steinmetz S, Bianchi A, Tijdens K, Biffignandi S (2014) Improving web survey quality: potentials and constraints of propensity score adjustments. In: Callegaro M, Baker R, Bethlehem J, Göritz A, Krosnick JA, Lavrakas PJ (eds) Online panel research: A data quality perspective (Wiley series in survey methodology). Wiley, Chichester, pp 273–298
Stigler GJ (1962) Information in the labor market. J Polit Econ 70(Part 2):94–105
Sweetland SR (1996) Human capital theory: foundations of a field of inquiry. Rev Educ Res 66(3):341–359
Tijdens KG, van Zijl S, Hughie-Williams M, van Klaveren M, Steinmetz S (2010) Codebook and explanatory note on the WageIndicator dataset: a worldwide, continuous, multilingual web-survey on work and wages with paper supplements, AIAS Working Paper. University of Amsterdam, Amsterdam, pp 10–102
Tijdens KG, De Ruijter E, De Ruijter J (2012) Measuring work activities and skill requirements of occupations: experiences from a European pilot study with a web-survey. European Journal of Training and Development 36(7):751–763
Tippins N, Hilton M (2010) A database for a changing economy, Review of the Occupational Information Network (OA database for a changing economyNET®). The National Academies Press, Washington
Vanhoy, M. (2008) An entropy estimator of population variability in nominal data. Journal of Scientific Psychology, June, 25–30
Wilcox AR (1973) Indices of qualitative variation and political measurement. The Western Political Quarterly 26(2):325–343