CENTRE *for* ECONOMIC
PERFORMANCE

CEP Discussion Paper No 1545

May 2018

Specialization Matters in the Firm Size-Wage Gap

Maria Molina-Domene

THE LONDON SCHOOL
OF ECONOMICS AND
POLITICAL SCIENCE

Abstract
This study applies the O-ring theory to explain the firm-size wage premium. It focuses on the joint role of the division of labor and employee characteristics. Including the firm heterogeneity of occupations in a standard wage regression with individual fixed effect shrinks the size coefficient by a third. Labor productivity follows a similar pattern as wages. The intuition is that individuals who work for large firms focus on a limited number of tasks become more efficient and productive, and earn higher wages. Additional predictions originating from the labor specialization hypothesis receive support from the data.

1 Introduction

Why do small establishments pay employees less than large establishments? There is a large literature both documenting and attempting an explanation of the firm size-wage premium. Nevertheless, the existing explanations are incomplete and an unexplained firm size-wage gap remains (Lester, 1967; Brown and Medoff, 1989; Burdett and Mortensen, 1989; Green et al., 1996; Troske, 1999 and Barth et al., 2016). I investigate the idea that workers in large firms focus on fewer tasks, exhibit higher skills (quality) and hence are more productive. I find that greater labor specialization explains about a third of the firm size-wage gap.

Numerous studies conclude that larger employers select more skilled workers and therefore pay higher wages. Other explanations revolve around employer characteristics like market power, capital intensity, and unionization. Also there is at least one organizational explanation, which is that the firm size-wage premium relates to efficiency wages. The idea is that larger firms pay efficiency wages because they find it more difficult to monitor workers. Apart from this, the literature has not focused much on the different organizational structures of small and large firms. In this paper I suggest and test such an explanation.

Large employers make better use of the division of labor by assigning their employees more specialized roles. That is, employees tend to work on a fairly limited number of more specific and closely related tasks. Conversely, small employers need to perform a comparable set of tasks with fewer workers. Their employees need to be able to perform a greater variety of tasks, which sometimes go beyond their specific roles.

For example, if biscuits are manufactured by a large firm, some employees will be specialized in utilizing the baking oven or different parts of the production line, while others will have expertise in quality control and good manufacturing practices. There could also be experts in the areas of marketing, product development and customer satisfaction. In contrast — if the same process is executed in a small firm an employee working at the baking oven will likely also be working on the production line and carrying out quality controls, while whoever is working in marketing might also be in charge of customer satisfaction, and so on.

I use linked employer-employee data covering all private sector workers for the period 1993-2010 from the German Social Security system[1]. This data allows me to construct the entire occupation distribution within each firm surveyed. I include measures capturing the heterogeneity of the occupations in a firm in wage regressions. Firms where workers perform dissimilar occupations or where they are focused on fewer tasks, pay higher wages, and including this effect attenuates the firm size-wage premium. For a subsample of firms, business characteristics like sales are available from the IAB Establishment Panel survey. For these firms, I obtain results for labor productivity, which have the same pattern as those for wages — suggesting a productivity based explanation. Given this setting, the productivity advantage may translate into higher wages through rent sharing.

Following Autor and Handel's (2013) approach, I delve into the 'task framework'. The underlying intuition is that high levels of division of labor entail performing distinct tasks. To link the occupations to their related tasks, I merge the German Social Security data with the U.S. Occupational Information Network (O*NET) data and build a measure utilising the simple idea of assigning employees to occupations where they focus on distinct tasks. Including this measure in wage and productivity regressions, I find that firms assigning fewer and more distinct tasks pay higher wages.

---

[1] This study uses the Linked-Employer-Employee Data (LIAB) [version] from the IAB. Data access was provided via on-site use at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and subsequently remote data access (fdz1059-1060-1061). LIAB version: longitudinal model 1993-2010 (LIAB LM 9310). Data access was provided via on-site use at the Research Data Centre (FDZ) of the German Federal Employment Agency (BA) at the Institute for Employment Research (IAB) and subsequently remote data access. Data documentation longitudinal model 1993-2010 Heining, Jörg; Klosterhuber, Wolfram; Seth, Stefan (2013): Linked-Employer-Employee-Daten des IAB: LIAB-Lngsschnittmodell 1993-2010 (LIAB LM 9310). FDZ-Datenreport 08/2013 Klosterhuber, Wolfram; Heining, Jörg; Seth, Stefan (2013): Linked-employer-employee-data from the IAB: LIAB longitudinal model 1993-2010 (LIAB LM 9310). (FDZ-Datenreport, 08/2013 (en)), Nuremberg, 62 S. Heining, Jörg; Klosterhuber, Wolfram; Seth, Stefan (2014): An overview on the Linked Employer Employee Data of the Institute for Employment Research (IAB). In: Schmollers Jahrbuch. Zeitschrift fr Wirtschafts- und Sozialwissenschaften, Jg. 134, H. 1, S. 141-148. Fischer, Gabriele; Janik, Florian; Müller, Dana; Schmucker, Alexandra (2009): The IAB Establishment Panel things users should know. In: Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften, Jg. 129, H. 1, S. 133-148.

There may be different organizational choices underpinning my results[2] but I focus on labor specialization. I argue the O-ring theory (Kremer, 1993) helps to explain two mechanisms through which specialization may foster labor productivity: the specialization due to job assignment by skills and the sorting of workers into firms due to the role of skills held by the workers.

Intuitively, under the first mechanism, employees working for large employers make less mistakes in the Smithian and Beckerian way, because they specialize and concentrate on a narrow set of tasks. Under the second mechanism, given that large firms perform a higher number of tasks, large employers avoid mistakes by employing highly skilled individuals. I suggest that workers skill represent their expertise at a task or the probability of successfully completing that task (i.e. even beyond the educational attainment).

To illustrate the first mechanism of job assignment by skill, assume that both large and small firms hire the same type of workers. While large firms assign employees to the tasks they are best at, small firms assign them to multiple tasks with the drawback that workers might not be equally good at all tasks. In a nutshell, employers who make better use of the division of labor assign their employees fewer tasks and therefore achieve efficiency and productivity gains because the possibility of mistakes lessens when you have focused employees.

I evaluate this pattern using survey-based data from the German Federal Institute for Vocational Education and Training. I regress firm size and wage on measures that account for the number of tasks performed by individuals and the time spent on them, and the results confirm my intuition.

The second mechanism relates to the sorting of workers into firms. Because of job assignment, large firms look for individuals who have expertise in certain skills and consequently are more productive under higher division of labor (specialists) while small firms seek out individuals who can switch between different tasks (generalists). I am interested in whether employees who are highly skilled at task 1 work with employees highly skilled at task 2 and I expect to find this matching of workers predominantly in large firms.

---

[2]For instance, Caliendo et al. (2015a,b) find that firms with more hierarchical layers are larger and pay higher wages.

My prediction is that individuals with specific talents (e.g.very knowledgeable on particular subjects) are more likely to work in large firms where they can make best use of their expertise. In contrast, generalists will prefer to work in small firms because they are well versed in a variety of fields but do not excel in any of them. In this sense the individual who works for a small employer will be an extension of Lazear's (2005) 'jack of all trades'[3].

I test this implication with two datasets: the U.S. National Longitudinal Survey (NLSY79) and the Survey of Adult Skills (PIAAC). I regress firm size and wage on measures reflecting focus (specialization) in skills. The results for both datasets display the same feature: specialists tend to be concentrated within large firms and earn a firm size-wage premium, while generalists more likely work for small firms. The NLSY79 results are particularly compelling because they suggest that information about individuals at a young age could predict the size of the employer these individuals will work for and the wage premiums they could potentially earn. Another aspect of the sorting of workers into firms derived from specialization, is that that expertise results in individuals who are more likely to work within a similar occupation when they change job. Experts (specialists) typically possess less transferable talents and are less able to move to a very different occupation if they change job[4].

To investigate this implication I study individuals who change job. Again, using data from the German Social Security system, I characterize somebody as a specialist if the individual works in a similar occupation after changing jobs. In a cross section analysis, I regress firm size and wage on this additional measure of specialization and confirm the robustness of my previous results. Specialists tend to work for large firms and earn higher wages than generalists do.

The remainder of the paper is organized as follows. Section 2 suggests a concep-

---

[3]Lazear (2005) and Moog et al. (2015) suggest that individuals will be inclined to specialize in one skill if they have a strong absolute advantage in doing so. These authors argue that the ability to combine talents can trigger entrepreneurship. On the other hand, narrow expertise or specialization (e.g. in fields where knowledge is deeper) predicts working in larger teams (Jones, 2008).

[4]A learning-by-doing mechanism could explain how working in large firms allows employees to become better specialists (i.e. in the extreme, they replicate the same production process indefinitely), while working in small firms allows employees to become better generalists.

tual framework and Section 3 lays out my empirical strategy and describes the data used to illustrate specialization in terms of productivity and wages. Section 4 presents and discuss the results and in Section 5, I conclude.

2   Conceptual Framework

Consider a firm performing a production process, which consists of n tasks. For simplicity, let's assume n is technologically fixed and each task requires a single worker [5].

Let's call q a worker's quality or skill at a task and define the expected percentage of maximum value the product retains if a certain worker performs the task. Each worker has a chance of performing the task perfectly and the probability of mistakes by different workers is independent.

Define B as the output per worker with a single unit of capital if all tasks are performed perfectly. There is a fixed supply of capital k.

The production function is:

$$y = k^{\alpha}(\prod q_i)nB \tag{1}$$

The O-ring production function differs from the standard efficiency unit formulation of labor skill, in that it does not allow quantity to be substituted by quality. Specifically, mistakes in a task or production process reduces the product's value. Combining the O-ring theory and Becker (1981), firms maximize profit if they match together workers of high quality q in the n tasks (i.e. those who make few mistakes). Taken to the extreme[6], a specialized firm operates the ray where the ratio between the number of workers to production activity (input) is constant and cannot reassign production activities across employees[7].

Firms pay wages $w(q)$, based on worker i quality q. The first-order condition of

---

[5]Additionally, workers supply labor inelastically and do not face a labor-leisure choice.

[6]The production function should be a Leontief, where $y = \min y(n)$

[7]I assume specialized firms operate under optimal level of division of labor. That is, firms specialize as much as the market allows (Kim, 1989), conditioned on other forces, such as coordination and communication costs, supervision complexity, etc.

each q_i is:

$$\frac{dw(q_i)}{dq_i} = \frac{y}{dq_i} = (\prod_{j \neq i} q_j)nBk^\alpha \tag{2}$$

The marginal product of quality, $d(y)/dq_i$ must equal the marginal cost of quality $dw(q_i)/dq_i$.

Specialized firms operate under high levels of division of labor and hire high quality workers (those who have less chances to make mistakes when they perform tasks). I argue this matching process takes place predominantly within large firms.

2.1 The firm size-wage gap

2.1.1 First mechanism: the division of labor

Large firms make more use of the division of labor. Indeed, there is a positive correlation between the number of tasks and the number of workers. As each employee focuses —taken to the extreme on one task, the probability of making mistakes decreases. The employee working at the baking oven is less prone to make mistakes, if he just performs that task. The same for the engineer who supervises the production line, and so on.

Workers are more productive because they concentrate on a single production activity, combining their outputs with employees who work on other production activities. This job design produces superadditivity (Rosen, 1978) and leverages labor productivity, due to the multiplicative effect arising from the complementarity between employees.

Large firms can hire workers of similar or dissimilar educational attainment or skills (Kremer and Maskin, 1996). For instance, there are production line workers and engineers, but what counts is that each worker performs a task at the maximum quality q.

2.1.2 Second mechanism: the sorting of workers into firms

I argue specialists (proficient workers) sort into large firms and earn higher wages. Consistent with Abowd et al., (1999) person effects, especially those not related to

observable characteristics, are an important source of the positive firm size-wage rate relation.

My intuition is that the specialist's proficiency goes beyond traditional observable patterns of skills such as education achievement, experience and earnings (Iranzo et al., 2008). Building on Lazear (2009), I suggest the dispersion of talents or of education interests can complement traditional proxies for skills.

Going back to the previous example, the production line worker does a better job within a large firm if he has the ability to perform optimally a specific task. Instead, if he is a multitasker, he will benefit from working for a small employer.

As in the O-ring theory, the specialist has high probability of successfully completing a specific task. Given his ability to concentrate, the specialist make the most of his capacities, working under high levels of division of labor.

3 Empirical Strategy and Data Description

3.1 Firm size-wage gap assessment using German linked employer-employee data

A number of explanations for the firm size-wage gap account for observed worker cross-sectional differences (Brown and Medoff, 1989; Oi and Idson, 1999 among others). I start testing these differences and then I evaluate the specialization explanation I suggest for the firm size-wage gap.

To assess the firm size-wage gap I work with the linked employer-employee data (LIAB LM9310) provided by the German Federal Employment Agency, which matches establishment data (BHP Establishment History Panel) to administrative biographies of individuals (IEB Integrated Employment Biographies).

From all the biographies I focus on the Employee History (BeH) data which are the annual and end-of-employment notifications submitted by employers to the social security agencies for individuals subject to social security (i.e. excluding civil servants, self-employed, students and individuals receiving earnings replacement benefits) and part-time employees (since 1999).

I build a sample which includes around 35 million employee notifications (observations) and corresponds to 1,686,783 employees. It excludes employees with reported zero daily wages (i.e. due to employment interruption notifications such as maternity leave), establishments with less than two employees and those who work for public administration and defense; political parties; educational, scientific and cultural organizations; Christian churches and representations of foreign countries. After deleting observations that do not have information on the variables I study, the LIAB LM9310 sample includes around 12 million observations.

It is worth noting that the FDZ Datenreport provides a precise definition of establishment as a regionally and economically delimited unit in which employees work and to which, according to specific principles, an establishment number is to be allocated (e.g. if the company has several branch offices in several municipalities each of these branch offices is an establishment[8]). More leniently, establishment and firm are used as interchangeable in the present paper.

The empirical strategy focuses on evaluating the impact of the levels of occupation heterogeneity on the firm size-wage gap. I start running a standard wage regression, controlling for worker characteristics X_{it} (age, age squared, gender, education, occupational status[9] and occupation title[10]). I also control for year, industry and region of the employer F_{jt}:

$$W_{ijt} = \alpha_i + \beta Size_{jt} + \pi X_{it} + \gamma F_{jt} + \varepsilon_{ijt} \tag{3}$$

The dependent variable W_{ijt} is the natural logarithm of gross daily wages of individual i at time t, working in firm j. The reported gross daily wages in LIAB are top coded; that is they are reported up to the upper earnings limit for statutory pension insurance. These limits might change every year and differ between Eastern and Western Germany. I address this potential issue below.

The size variable is represented by the natural logarithm of the number of employees in the establishment. Including industry and region dummies as additional

---

[8]Therefore, I am unable to observe firm patterns such as decentralization, integration, etc.

[9]Distinguishes blue-collar, white collar, trainee, apprentice, etc.

[10]Occupation titles come from the 'Classification of Occupations. Systematic and Alphabetical Directory of Job Titles, KldB88.'

controls in my regressions allows me to capture the idiosyncrasies of different economic activities and geographic realities.

The industry dummies are 3-digit codes of the relevant industry in accordance with the Classification of Economic Activities for the Statistics of the Federal Employment Services (WZ73). The region dummies correspond to the place of residence of the employee (10 regional directorates: North , Berlin-Brandenburg, Saxony-Anhalt/Thuringia, Saxony, Lower Saxony-Bremen, North Rhine-Westphalia, Hesse, Rhineland-Palatinate-Saarland, Baden-Wuerttemberg and Bavaria).

My second specification scrutinizes the effect of occupation heterogeneity on wages. Therefore, I add measures of occupation heterogeneity to my basic specification as described in the following subsection:

$$
\begin{aligned}
W_{ijt} = \alpha_i + \beta Size_{jt} + \delta OccupationHeterog_{jt} + \pi X_{it} \\
+ \gamma F_{jt} + \varepsilon_{ijt}
\end{aligned}
\tag{4}
$$

Firms' occupation heterogeneity can be interpreted in different ways and in this study I suggest that it can be an indicator of specialization[11]. I expand on this in Section 3.3. In addition to the organizational choice, labor quality may be playing a role in the firm size-wage gap. I use individual fixed effect ψ_i to control for the unobserved dimensions of the employee productivity as in Brown and Medoff (1989):

$$
\begin{aligned}
W_{ijt} = \alpha_i + \beta Size_{jt} + \delta OccupationHeterog_{jt} + \psi_i + \pi Controls_{it} \\
+ \gamma Controls_{jt} + \varepsilon_{ijt}
\end{aligned}
\tag{5}
$$

---

[11]Departing from Garicano (2000) in which only a layer of the organization is formed by specialist problem solvers (knowledge-based hierarchy), I do not impose restrictions on the structure of the firm's hierarchy. For instance, my hypothesis leaves space for explanations like Fox's (2009); that increasing job responsibility raises the firm-size wage premium. This mechanism could be supporting the remaining unexplained firm-size wage gap.

3.1.1 Measures of occupation and tasks heterogeneity

I work with two measures of occupation heterogeneity. Both measures rely on the fact that each employee is assigned a unique occupational title for the job he performs within a firm. This information is included in the Employee History administrative data. Employers encode an employee's occupation with the title that best defines the main activity performed (i.e. even if more than one title could apply to one employee), in accordance with any of the German systematic classification of occupations. My data contains information of around 330 titles provided in the 3-digit coded 'Classification of Occupations. Systematic and Alphabetical Directory of Job Titles. KldB88'.

The first measure is simple and straightforward but could leave uncovered industry idiosyncrasies in terms of occupation heterogeneity. Therefore, I complement it with a second measure which helps to address this potential weakness.

Firstly, I compute the Simpson's Interaction Index (hereafter S) which allows measuring the proportion of different occupations within each firm every year (see Appendix). This index is appealing[12] because it measures occupation diversity within firms and it is highly comparable across different firms (i.e. S equals zero when there is complete segregation or division of labor: employees are evenly distributed among different occupations).

Next, I work with the dynamic version of the Ellison and Glaeser Index[13] (Dumais et al., 2002) (hereafter EG) applied in this case to compute the distribution of occupations in a firm. Given its construction, one advantage of working with EG is that it tells us to what extent the firms' heterogeneity of occupations departs from the occupation heterogeneity in a typical firm, within a specific industry every time there is a change in the firm employment composition (i.e. due to the start and the end of an employment relationship) (see Appendix).

---

[12]Reardon and Firebaugh (2002) evaluate desirable properties of S and other multigroup concentration/segregation indexes. One of these properties is size invariance. My measure S satisfies this property because if the number of employees in each occupation is multiplied by a constant factor, heterogeneity remains unchanged.

[13]Dumais et al.,(2002) apply the Ellison and Glaeser Index to measure the geographic concentration of industries. The original version was proposed by Ellison and Glaeser in 1997. I multiply this indicator by minus one to focus on dispersion instead of concentration (see Appendix)

Finally, I merge the LIAB-LM9310 with the O*NET 21.2 database (U.S. Department of Labor), to measure whether different individuals focus on distinct tasks. I use a crosswalk to link occupations classified according to the Standard Occupational Classification 2010 (O*NET-SOC 2010) and the 3-digit coded Classification of Occupations (Systematic and Alphabetical Directory of Job Titles–KldB88), present in the German data. Consistent with Becker's (1994) view, the intuition is that within large firms, each employee concentrates on a set of definable and diverse tasks.

The O*NET data contains information on standardized and occupation-specific descriptors (levels and importance) organized to reflect the characteristics of workers (abilities, values, knowledge, and licensing) and jobs (tasks, work context, and labor market information). Within job patterns, I focus on tasks, in particular in a set of variables called Generalized Work Activities (GWAs), which are 41 task statements specific to occupations and generalized work activities that apply to multiple occupations (see Appendix).

The GWAs tasks range from thinking creatively; estimating the quantifiable characteristics of products, events, or information; to handling and moving objects, etc. They are based on the primary activity among multiple activities found in a task statement and are also based on the meaning of a given task statement (one including work activity, purposes, context and technology).

The importance of the GWAs tasks ranges from 0 to 100. Taken to the extreme, within large firms, each employee concentrates on a different task, giving that task very high importance. In parallel, co-workers in a different occupation assign that task less importance and focus on a different task and so on. As a natural outcome, individuals working for large firms are more likely to develop firm-specific human capital than those working for small firms.

To compute my specialization proxy, I firstly standardize the task variables and then compute the standard deviation of the importance of the 41 tasks present in a firm occupation by occupation. Next, I build my measure of the heterogeneity of tasks across occupations, which is the average (i.e.among all the individuals who work for a firm) of the standard deviations of the importance of tasks (see

Appendix) and I call it SD. The intuition is that at a maximum level of SD, each employee focuses on fewer distinct tasks.

3.1.2 Productivity implication using German linked employer-employee data

Following the Becker and Murphy (1994) model an intuitive implication is that specialization impacts labor productivity, leading to increasing returns[14]. The model explains that labor productivity could increase due to higher division of labor and knowledge. This is because workers do better if they specialize in a subset of tasks, and they can be even more productive at particular tasks depending in part on how much knowledge they have[15]. The increasing returns of concentrating on a narrower set of tasks should raise the productivity of the specialist over the generalist.

After testing the impact of occupation heterogeneity on wages I evaluate its potential impact on productivity. My strategy is to study whether occupational heterogeneity works through a productivity channel. Again I use the German administrative data and build a subsample, merging LIAB LM9310 with the waves of surveys from the unbalanced[16] IAB Establishment Panel (1994-2010).

This additional data is useful for my analysis because it contains information on firm characteristics such as sales, investments, etc. Given that the survey's sampling frame is all establishments covered by the social security system, stratified according to industry, firm size and federal state, this data should be representative of the German firm population.

My subsample consists of around 3.5 million observations and also excludes civil servants, self-employed, students and individuals receiving earnings replacement benefits. These observations correspond only to employees who worked in the firms surveyed in specific years (i.e. they exclude employees in firms that while part of

---

[14]Greater specialization can result in increasing returns to scale depending, for instance, on the level of indivisibility of labor (Edwards and Star, 1987).

[15]Some of the human capital individuals acquire can be firm-specific or even task-specific human capital (Gibbons and Waldman, 2004).

[16]This panel is unbalanced because only the majority of the same establishments are interviewed every year.

the panel did not hold interviews in those corresponding years and they exclude employees in firms not selected for the IAB Panel survey).

The empirical strategy here is to mimic the wage regressions at the worker level using labor productivity as the dependent variable for a subsample of firms (i.e. for which sales are available). I run productivity regressions at the firm level and not at the worker level because otherwise, having more employees, large firms could have more weight on the results. This allows me to discern whether specialization affects productivity in the same way as wages.

The first productivity specification controls for age, age squared, education, year, industry and region:

$$P_{jt} = \alpha_j + \beta Size_{j(t-1)} + \pi Controls_{it} + \gamma Controls_{jt} + \varepsilon_{jt} \qquad (6)$$

The labor productivity variable P_{jt} is defined as the value added divided by the number of employees of the firm j at time t. Value added is defined as the natural logarithm of the business volume (sales in euro) minus intermediate inputs (e.g. all raw materials and supplies purchased, external services, rents, etc.). As the dependent variable is computed using the current number of employees, I use the lagged size to avoid working with a simultaneously determined regressor. The rest of the independent variable definitions are the same as in the previous subsection. The second productivity specification adds to the basic specification measures of occupation heterogeneity:

$$\begin{aligned} P_{jt} = \alpha_j + \beta Size_{j(t-1)} + \delta OccupationHeterog_{jt} \\ + \pi Controls_{it} + \gamma Controls_{jt} + \varepsilon_{ijt} \end{aligned} \qquad (7)$$

where the occupation heterogeneity measures are again the S and EG Indexes and the SD of the importance of tasks, computed as described in subsection 3.1.1. Table 2 shows a comparison of the impact of these heterogeneity measures on the LM9310 and LM9310+IAB samples.

14

3.2 The division of labor/specialization hypothesis. Evaluation of mechanisms using different datasets

The size-wage strategy I present could admit different explanations. S, EG and SD measure firm occupational distribution and a concern is that occupational distribution could be consistent with various interpretations. Different occupational distribution could reflect different organizational choices such as the degree of integration, more or less extensive hierarchical controls, etc.

An alternative interpretation is that the occupational distribution reflects the fact that bigger firms carry out more tasks but the level of specialization is the same. To rule out this alternative and provide support for my specialization hypothesis, I investigate independent evidence of the two key mechanisms: very specific and clear task assignments (i.e. high levels of division of labor) and sorting of workers into firms. These two mechanisms can obviously coexist.

For the first mechanism, assume that small and large employers hire the same type of worker. My intuition is that lower specialization is frequently small firms' organizational choice which might result in lower efficiency and productivity (e.g. an employee of a small manufacturer of biscuits sometimes undertakes quality controls, documents anomalies and suggests corrective actions in the production process; other times he focuses on the production process and when necessary he helps his colleague working on the production line). On the other hand, within large firms there should be higher specialization because every worker carries out very few tasks (taken to the extreme, only one task) and this translates into a considerable number of very specific tasks. I expect tasks within large employers to be properly assigned to employees on a regular basis, which might represent gains in efficiency and productivity[17].

For the second mechanism, which is the sorting of workers into firms, I drop the assumption that small and large firms hire the same type of workers and study the sorting of generalists into small firms and specialists into large firms.

---

[17]Complementary to my argument, Bolton and Dewatripont (1994) predict returns to specialization, when the firm is able to minimize costs of processing and communicating information. Otherwise, benefits from specialization can be offset by the increased costs of communication.

My basic empirical analysis does not provide information on these mechanisms (e.g. whether specific abilities matter for working in small or large employers). To provide more insights on this I employ two additional strategies: I run firm size and wage regressions on variables capturing the specialization of individual workers.

The specification for the firm size regressions is:

$$S_{jt} = \alpha_i + \lambda Specialization_{it} + \pi Controls_{it} + \gamma Controls_{jt} + \varepsilon_{ijt} \qquad (8)$$

where S_{jt} is the firm size and $Specialization_{it}$ are the different measures of specialization adopted in different contexts and datasets.

The specification for the wage regressions is:

$$W_{ijt} = \alpha_i + \lambda Specialization_{it} + \pi Controls_{it} + \gamma_C ontrols_{jt} + \varepsilon_{ijt} \qquad (9)$$

and W_{ijt} is the wage of individual i at time t, working in firm j and $Specialization_{it}$ are the different measures of specialization.

In the following paragraphs I present alternatives for testing these strategies using different datasets.

3.2.1 Specialization mechanism. Evaluation using BIBB data

I study the first mechanism of my hypothesis with the Employment Survey data[18]. These surveys are considered representative among full-time employed persons in Germany. They collect information on vocational qualifications and the working conditions of individuals aged 15 or older who work in paid employment for at least 10 hours. I work with two cross-section employment surveys conducted in

_____

[18]BIBB/BAuA Employment Survey of the Working Population on Qualification and Working Conditions in Germany 2005/2006, 2012 and Supplemental Task Survey 2012 to Employment Survey 2012.

2005/2006 and 2011/2012 (approximately 20,000 individuals), and the second part of the Supplemental Task Survey to employment (approximately 2,000 individuals).

The surveys provide data on the professional requirements for performing tasks as well as the different tasks performed by employees, which is fundamental for my analysis. This data set is compelling because workers are asked about the number of the different tasks performed and their frequency. These two dimensions enable us to measure whether an employee is carrying out either more or fewer tasks and parallel to that — if an employee undertakes the same tasks frequently.

My empirical strategy is to run size and wage regressions on proxies of specialization. The firm size variable is the firm size class where the individual works (i.e. including owner and trainees). It is an ordered categorical variable which ranges from 1 to 10.[19]

To build the proxies of specialization I gather information about tasks from two sources. One source is the information provided by the main surveys. They contain categorical ordinal variables which record the choice of types of work performed in the job and their frequency (frequently, sometimes, never). These are standardized tasks such as manufacturing, quality control, working with computers, etc.[20] and considered exhaustive by around 88 percent of the respondents (i.e. respondents replied they do not need to mention any other activity performed in the job). The second source is the supplemental survey, which provides additional information to the main surveys, measuring how much time respondents spend on these standardized tasks on any working day of their choice.

The time spent on tasks should be more informative than frequencies, but it is only available for a small subsample of 2,000 respondents. I take advantage of the fact that time is continuous and relate the first moment of time variables distribution

---

[19]This variable takes the following values: 1 (2 people), 2 (3-4 people), 3 (5-9 people), 4 (10-19 people), 5 (20-49 people), 6 (50-99 people), 7 (100-249 people), 8 (250-499 people), 9 (500-999 people) and 10 (1000 and more).

[20]The survey gives a choice of 19 tasks in 2005/2006 and 20 tasks in 2011/2012. There is a new question about using the Internet or editing e-mails. This is a categorical ordinal variable which gives employees 3 options: each task could have been performed sometimes, frequently or never.

to the frequency variables. This is valuable because being frequent or sporadic means different spans of time spent on the different tasks. For instance quality controls are considered frequent if the time spent on them in mean is 49 minutes while working with computers is considered frequent if the time spent on it is in mean 115 minutes.

Working with the BIBB data I include two regressors of interest: the Simpson's Interaction Index and the count of tasks, which is the number of different types of task performed frequently or sometimes in the job by each employee. Within this context, the Simpson's Interaction Index reflects individual's focus on specific tasks and it is based on the fraction of total work time spent on each task on the reporting day (see Appendix). My intuition is that performing fewer tasks regularly could more likely make an employee a specialist, rather than performing many tasks seldom[21]. Hence, maximum level of specialization entails performing only one task frequently (i.e. all the time spent in the working time is dedicated to one task).

I drop observations containing missing answers (non response, refusal, etc.) on the variables I use for my analysis and also delete observations with imputed wages. My samples include around 14,000 observations.

The regressions with the BIBB data control for school education (highest general school leaving certificate), occupation (3-digit coded occupation based on KldB1992) and industry (2-digit coded branch of industry according to WZ2003).

3.2.2 Sorting of workers into firms mechanism. Evaluation using NLSY79 and PIACC data

For the second mechanism I investigate the prediction that employees who work for more or less specialized establishments are intrinsically different in terms of skills balance. That is, individuals who have more specialized knowledge might end up working for large employers while those who have more balanced skills

---

[21]Within this scenario, there is higher efficiency in avoiding multitasking and allocating effort to specific tasks. This contrasts with Holmstrom and Milgrom (1991) model where there are efficiency gains in assigning effort across a given set of tasks.

might work for small employers. Regarding wages, more specialized individuals could earn higher wages than less specialized ones.

Firstly, I work with NLSY79 data. The NLSY79 data is provided by the U.S. Department of Labor and originally consisted of information on the lives of a cohort of $12,686$ American respondents (which now become $9,964$), first surveyed in 1979. The data refers to 25 rounds of surveys from 1990 to 2012 because respondents were interviewed annually (until 1994) and biennially (since then).

My sample consists of around $41,000$ observations. I exclude missing values (i.e. refusal, don't know, valid skip, non-interview) and self-employed. I only use observations of individuals who work 6 or more hours a day (i.e. individuals who have finished their education and are working full time) and are not self-employed.

Secondly, I study this implication using the PIAAC surveys report. Working with this data is appealing because the measured abilities should be related to employee productivity and in turn to firm size. PIAAC Surveys of Adult Skills (adults aged 16-65) were conducted by the OECD in 2008-2013 in selected countries. They refer to around $5,000$ individuals in each participating country, representative of the population of adults living in that country.

The survey questions are designed to be valid cross-nationally, however participating countries adapt questions to reflect national characteristics (e.g. educational systems). For this reason I include in my sample only the fraction of individuals currently working in comparable countries such as Germany, Austria, Finland and the UK, where the assessment questions are the same. My sample then consists of around $13,000$ observations, excluding missing values.

My empirical strategy in this case is to run regressions of employer size and wage using the standard deviation of abilities as a key independent variable for this analysis, controlling for mean ability. I analyze if there is an association between excelling on any aptitude tested and the size of the firm where the individual works and the wage this individual earns.

For the NLSY79 data firm size is given by the number of employees at the place where the respondent works. The wage variable is the natural logarithm of the gross pay rate received in the job. This pay rate accounts for tips, overtime,

bonuses and reflects how much the respondent usually earns at that job.

My regression includes the standard deviation of all the Armed Services Vocational Aptitude Battery (ASVAB) vocational test scores controlling for mean ASVAB. The ASVAB scores are based on exams (administered to the respondents in 1980) on 10 areas of knowledge.[22]. I compute the standard deviation and the mean of standardized variables using the 10 tests.[23]

My additional controls are employee age, age squared, education (completed education on the survey year: completed less than 12th grade, enrolled in high school or college, high school graduate), occupation (3-digit coded according to the 2000 census), the industry (4-digit coded according to the 2000 census) where the respondent works and the year of the interview.

Once again I run regressions of employer size and wage on the standard deviation of the tested abilities, controlling for the mean.

For the PIAAC data, firm size is given by an ordered categorical variable which ranges from 1 to 5[24] and I treat it as continuous. The wage variable for this dataset is the monthly income rank category.

I focus on worker abilities used in the workplace. These abilities are based on proficiency levels shown in tests on three areas of knowledge: literacy, numeracy, and problem solving. The PIAAC survey provides ten plausible values for each test and each individual. I average these ten values to obtain a single test score. I then compute the average and standard deviation of the standardized values of the three test scores.

In addition I control for employee age class (10 classes, from 16 to 65 years old), education (highest qualification level), occupation (according to the International

---

[22]This is a battery of multiple choice tests administered to high school students (when they are in 10th, 11th and 12th grade), which is used to determine qualification for enlisting in the U.S. Armed Forces.

[23]I use the scale scores related to the 10 tests that measure knowledge and skill in the following areas: general science, arithmetic reasoning, word knowledge, paragraph comprehension, numerical operations, coding speed, auto and shop information, mathematics knowledge, mechanical comprehension, electronics information.

[24]This variable takes the following values: 1 (1-10 people), 2 (11-50 people), 3 (51-250 people, 4 (251-1000 people), 5 (More than 1000 people) plus other categories for refused, don't know, valid skip, not stated or inferred answers.

Standard Classification of Occupations 2008) and industry (1-digit code from the International Standard Industrial Classification of All Economic Activities) where the respondent works.

3.2.3 Sorting of workers into firms mechanism. Evaluation using LIAB-MM9308 data

Another way of using the data to measure whether someone is a specialist or a generalist is to concentrate on movers. Knowing that an employee tends to work in a similar occupation when he changes job could suggest that this employee is a specialist. The motivation is that a specialist's very specific talents might bind him to work in very similar occupations.

To characterize specialists I work again with the LIAB data, but this time with the Movers Model (LIAB-MM9308). This dataset contains around 129 million employee notifications.

I investigate whether individuals who change job (movers) persistently behave as generalists or as specialists. The question here is how likely is it that a worker from a large firm remains in a similar position after changing jobs. Knowing that individuals are more versatile and able to change occupation should be a sign of being a generalist, who tends to work in small firms and might earn lower wages. Conversely, individuals who are likely to work within similar occupations suggest that they are specialists.

My empirical strategy is then to run firm size and wage regressions on a specialization variable. Once more the size variable is represented by the natural logarithm of the number of employees in the establishment and the wage variable is the natural logarithm of the gross daily wages of individuals.

My specialization variable within this context is called proximity. Higher proximity should be interpreted as working within similar occupations while changing employer.

I build the proximity variable based on the employee change of occupation. The first digit of the 3-digit KldB88 coding identifies the occupation class. I consider a substantial change of occupation when the worker changes occupation class (i.e.

the new occupation has a different first digit KldB88 code) and a small change when the worker keeps working within the same occupation class (i.e. the new occupation has the same first digit KldB88 code but the second or third digits are different).

I compute a categorical variable which equals 2 if the individual makes a substantial change of occupation, equals 1 if the individual moves to a similar occupation, and equals 0 if the individual stays in the same occupation. The average of this variable reflects how far the individual moves in terms of job class while changing employer. My proxy of specialization is the inverse additive of this variable, which I call proximity.

After deleting observations that do not have information on the variables I study, there remains around 38 million observations. I build a cross-section sample on the last observation of each employee, which includes around 3.8 million observations. In my basic specification I control for worker characteristics (age, age squared, gender, education and occupational status). The region dummies correspond to the place of residence of the employee (10 regional directorates as stated in 2.1) and the industry dummies correspond to the economic sector (WZ73).

An alternative point of view would be that job changes reflect a search process, where workers locate better firms and job matches, consistent with Topel and Ward (1992). Towards the end of this process they may end up at larger firms and make more marginal changes. In order to control for this potentially confounding situation my modified specification controls for the number of times the individual changes employer.

4 Results

4.1 The firm size-wage gap: results

I plot the wage against firm size for both LIAB samples in Figure 1. The size-wage relationship appears, in both samples, as strictly increasing. In practice these results suggest that an employee working for an establishment of the smallest size class earns in mean around 20 percent less than an employee working for an es-

tablishment of the largest size class. Furthermore the former establishments are on average around 20 percent less productive than the latter ones.

An important insight is that the S and EG index and SD (see Appendix) monotonically increase with establishment size. This implies that there is a higher division of labor in large establishments and it might mean that each employee works on a limited number of more specific tasks. Vice versa, smaller establishments display a higher S and EG index as well as SD, which I interpret as a lower division of labor.

Figure 2 presents the relationship between an employee's highest schooling qualification and firm size. The plot shows that there is only a weak pattern relating to these two dimensions. This is relevant for my empirical strategy for two reasons. Firstly, the fact that larger establishments do not necessarily employ individuals with the highest schooling qualifications but end up paying employees higher wages leaves space for the organizational explanation. Secondly, it points to additional employee characteristics, beyond those of educational qualifications, which give to the employee a comparative advantage when working in firms of a different size.

Descriptive statistics of the LIAB9310 dataset are shown in Table 1. Some results about the core question why small firms pay employees less than large firms are displayed in Tables 2 to 5. These are the outcomes of wage regressions using the two samples of German data (LIAB LM9310). The odd columns in Tables 3 to 5 present results in line with some of the traditional explanations for the firm size-wage premium. That is, I check if controlling for worker characteristics and including year, industry and region dummies as additional controls still leaves part of the firm size-wage premium unexplained.

For instance the statistically significant firm size-wage premium shown in Table 3, column 1 implies that a 10 percent increase in size augments wages by around half a percent, controlling for worker characteristics, industry and region.

The main feature of these results is that the positive and highly significant coefficient of size shrinks after controlling for specialization. Even columns in Tables 3 to 5 provide evidence that including my proxies for specialization (S, EG and SD) contributes to explaining the positive relationship between employer size and

23

wages. Indeed, including the S variable shrinks the size effect between 20 percent and 35 percent, including the EG variable reduces the size effect between 24 percent and 38 percent, and including the SD variable diminishes the size effect between 20 percent and 31 percent.

Columns 3 to 6 in Tables 3 to 5 show alternative specifications, controlling for additional employee characteristics such as occupation and occupational status. After including these additional controls the pattern of results changes little when using the three proxies of specialization.

The previous literature has documented that firm size is related to unobserved dimensions of employee productivity, therefore I complement my analysis using individual fixed effect. The results given in columns 7 and 8 of Tables 3 to 5, suggest that at least around 28 percent of the wage gap due to firm size vanishes when S, EG or SD are included within the fixed effects specifications.

Recalling that wages are censored, the estimates I obtain could be biased. I check the previous results computing Tobit estimates based on censored wages (Maddala, 1983). In addition I focus on the lower education subgroup since this group is less affected by censoring. In both cases the estimates and standard errors presented in Tables 3 to 5 remain unchanged. These findings suggest that even if wages are top coded, the OLS estimates are reliable and I adhere to this methodology for the rest of the paper[25].

My hypothesis is that specialization contributes to explaining the firm size-wage premium through a productivity channel. Consequently I check whether specialization matters for productivity using the LIAB LM9310 + IAB Panel subsample. Table 6 presents the results for the productivity regressions. Given that productivity is computed at firm level I do not mimic the individual fixed effect strategy in the labor productivity regressions. As anticipated above, columns 2 to 4 illustrate that the coefficients of the one-year lagged size decrease between 9 percent and 21 percent after controlling for occupation heterogeneity (S and EG) and by 16 percent after controlling for task concentration (SD) that I interpret as specialization.

---

[25]Dustmann et al.(2009) impute the censored part of the wage distribution under alternative distributional assumptions. They also find that OLS estimates and the standard errors based on imputed wages and Tobit estimates based on censored wages, are almost identical.

This reflects a similar pattern as shown in the wage regressions (Tables 3 to 5), suggesting that specialization might be playing a role through productivity.

4.2   Mechanisms of the specialization hypothesis: results

The assessment of the two mechanisms of my specialization hypothesis is performed with different datasets. Descriptive statistics of these datasets are shown in Table 7.

Firstly, I evaluate if employees working for large employers may be carrying out more distinct tasks. That is, leaving aside the type of employees hired, I test whether employees of firms of different sizes work under different levels of division of labor.

Large firms should make more use of the division of labor, assigning employees more distinct tasks. The results from a firm size regression on the specialist variable working with the BIBB survey-based data are presented in Table 8. Columns 1 and 2 exhibit the outcomes for the relationship between firm size and being a specialist and columns 3 and 4 show the results for the relationship between income and being a specialist. The odd and even columns present the outputs for 2005-2006 and 2011-2012 respectively.

Controlling for education, vocational degrees, skills, occupation, industry and year of the interview, the coefficient of the Simpson's Interaction Index variable is positive and highly significant and the coefficient of the count of tasks is negative and significant (except for the wage regression in 2012) within the firm size and wage regressions. This suggests that there is as strong association between systematically spending more time on fewer tasks and working at large firms as well as earning higher income. These findings seems to be in line with the first mechanism of my hypothesis. That is, a one standard deviation increase in the concentration of the time dedicated to tasks is associated with a rise in wages by at least 2 percent. In addition the negative coefficients of count of tasks (columns 1 to 3) reinforce my intuition that employees who work at large firms are assigned fewer tasks. Regarding the count of tasks, the interpretation is that performing sometimes or frequently 1 more task (out of 19 or 20 tasks, in 2005-2006 and 2011-2012

respectively), is related to a wage decrease of 1 percent (column 2). This effect becomes close to negligible in 2011-2012 (column 4).

Secondly, I test the mechanism of the sorting of workers into firms. Table 9 presents firm size and wage regressions on the dispersion of workers' skills. Even columns show the outcomes obtained working with NLSY79 data and odd columns exhibit the outcomes working with PIAAC data. In column 2, the coefficient of the average skill (AVG) shows the familiar pattern that more able individuals work for larger firms.

More important for my investigation is that the coefficients of the standard deviation(SD) are also positive and point to the fact that having a different bundle of skills (being a specialist) is linked to working for large employers while not excelling in any skill or talent but balancing them (being a generalist reflected in a low standard deviation) is associated with working for small employers. The same pattern holds for wages, in line with my hypothesis.

In addition, I test whether individuals who tend to switch around between similar occupations when they change job are likely to work for large firms and get higher wages while those who are more versatile and can switch to very different occupations typically work for small firms earning lower wages. To evaluate this implication I work with the LIAB MM9308 data.

Table 10 presents the results of two specifications. The results from the basic specification are displayed in columns 1 and 3. The highly significant and positive coefficient of proximity in column 1 confirms that individuals who move to similar occupations when they change job (specialists) tend to work in large firms. I obtain similar evidence for wages. The positive and significant coefficient of proximity displayed in column 3 makes apparent that more specialized individuals should get higher wages.

The second specification (columns 2 and 4) adds to the basic specification an additional control regarding the number of times an individual changes job. The results suggest the presence of a fairly important omitted variable bias in the basic specifications. Controlling for the number of times individuals change job shrinks the proximity effect in the size regression (column 2) and grows the proximity

effect in the wage regression (column 4).

This could mean that the basic specification is capturing other job mobility aspects such as different stages of the employment careers or simply a higher propensity to change job. Once these potential confounders are controlled for, it is confirmed that specialists tend to work for large firms at a lower extent than originally suspected. An interesting finding is that, given the positive relationship between the number of job changes and wages, the negative omitted variable bias in the wage regression suggests that higher wages are associated with less job changes and probably higher specialization.

5 Discussion and Conclusion

This paper has applied the O-ring theory to explain the firm size-wage premium. Large firms frequently make more use of the division of labor and make the most of specialists. This appears to increase employee's concentration on tasks and to decrease the probability of being inefficient.

I found that firm occupation heterogeneity explains around a third of the wage premium due to firm size. These results may be consistent with different explanations and I present supporting evidence for labor specialization.

Working with the linked employer-employee German administrative LIAB LM9310 and the IAB Panel I verified that observed worker cross-sectional differences, such as education attainment and vocational qualifications, did not completely account for the wage gap due to firm size. Next, my empirical strategy was to include proxies for occupational heterogeneity (S and EG) and task focus (SD) in wage regressions. I found that occupation heterogeneity helps to shrink the firm size-wage premium. Furthermore, these results were similar even after controlling for individual fixed effects.

I also investigated whether specialization plays a role in labor productivity. I replicated the wage strategy and have been able to confirm this intuition, working again with the German administrative data and O*NET data. An implication was that the productivity advantage could translate into wages through rent sharing.

I provided additional evidence for two mechanisms through which the O-ring theory could apply. Firstly, under the assumption that small and large firms hire the same type of workers, I explored whether specialization means specialized job assignment by skills. Secondly, dropping this assumption I investigated the sorting of workers into firms.

Given my basic empirical analysis did not provide information on the time spent on individual occupational tasks, abilities or talents, I evaluated these two mechanisms with different datasets. The intuition was that individuals who perform fewer tasks and possess very focused talents, have the chance to make less mistakes and therefore exhibit higher quality. These individuals tend to work for larger firms and to earn higher wages.

For the first mechanism, I worked with waves of the BIBB surveyed-based data which provide detailed information on job tasks of the German labor force. I found that being systematically focused on fewer tasks (as per specialists) is strongly associated with working for large firms and earning higher wages.

The second mechanism extended Lazear (2005), asking whether individuals who work for small firms possess balanced talents while individuals who work for large firms have more concentrated talents. Using the NLSY79 and PIAAC data the results showed that generalists (those who do not excel in any skills but balance them) are more likely to work for small employers and to earn lower wages.

Using the LIAB MM9308 movers, I studied the intuition that taking the same or similar occupations while changing jobs could confirm the pattern of having very specific talents. I also found that workers who move to more similar occupations when they switch jobs tend to work for larger firms and earn higher wages.

I discussed a number of ways in which specialists and generalists can be characterized. All of them helped to explain the sorting of individuals into firms of different sizes. I regarded this sorting process as evidence that specialists prefer to work within large firms, probably with higher levels of division of labor.

My findings lined up with the O-ring theory and provided empirical support for various implications of the division of labor/specialization hypothesis which relates to firm size. I found the joint role of the job assignment by skills and the sorting of workers by skills plausibly explain part of the firm size-wage gap.

References

Abowd, John M., Francis Kramarz and David N. Margolis. 1999. High Wage Workers and High Wage Firms. *Econometrica* 67, 251333.

Autor, David and Michael J. Handel. 2013. Putting Tasks to the Test: Human Capital, Job Tasks, and Wages. *Journal of Labor Economics* 31: S59-S96.

Barth, Erling, James David and Richard B. Freeman. 2016. Augmenting the Human Capital Earnings Equation with Measures of Where People Work. National Bureau of Economic Research, Working Paper 22512.

Becker, Gary S. 1981. Treatise on the Family (Cambridge:Harvard University Press). *Journal of Labor Economics*, 3: S33-S58.

Becker, Gary S. and Kevin M. Murphy 1992. The Division of Labor, Coordination Costs, and Knowledge. *The Quarterly Journal of Economics*, 107: 1137-1160.

Becker, Gary S. and Kevin M. Murphy. 1994. *The Division of Labor, Coordination Costs, and Knowledge in Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education.* Chicago: University of Chicago Press.

Bolton, Patrick and Matthias Dewatripont. 1994. The firm as a Communication Network. *Quarterly Journal of Economics* 109: 809-839.

Brown, Charles and James Medoff. 1989. The Employer Size-Wage Effect. *Journal of Political Economy* 97: 1027-1059.

Burdettt, Kenneth and Dale T. Mortensen 1989. Equilibrium Wage Differentials and Employer Size. 1989. Discussion Paper 860, Northwestern University.

Caliendo, Lorenzo, Ferdinando Monte and Esteban Rossi-Hansberg. 2015a. The Anatomy of French Production Hierarchies. *Journal of Political Economy* 123: 809-852.

Caliendo, Lorenzo, Giordano Mion, Luca David Opromolla and Esteban Rossi-Hansberg. 2015b. Productivity and Organization in Portuguese Firms. National Bureau of Economic Research, Working Paper 21811.

Dumais, Guy, Glenn Ellison and Edward L. Glaeser. 2002. Geographic Concentration as a Dynamic Process *The Review of Economics and Statistics* 84: 193-204.

Dustmann, Christian, Johannes Ludsteck, and Uta Schönberg. 2009. Revisiting the German Wage Stucture. *The Quarterly Journal of Economics* 124: 843-881.

Edwards, Brian K. and Ross M. Starr. 1987. A Note on Indivisibilities, Specialization, and Economies of Scale. *The American Economic Review* 77: 192-194.

Ellison, Glenn and Edward L. Glaeser. 1997. Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy* 105: 889-927.

Fox, Jeremy T. 2009. Firm-Size Wage Gaps, Job Responsibility, and Hierarchical Matching. *Journal of Labor Economics* 27: 83-126.

Garicano, Luis. 2000. Hierarchies and the Organization of Knowledge in Production. *Journal of Political Economy* 108: 874-904.

Garicano, Luis and Thomas N. Hubbard. 2008. Specialization, Firms, and Markets: The Division of Labor within and between Law Firms. *The Journal of Law, Economics and Organization* 25: 339-371.

Gibbons, Robert and Waldman Michael. 2004. Task-Specific Human Capital. *The American Economic Review* 94: 203-207.

Green, Francis, Stephen Machin and Alan Manning. 1996. The Employer Size-Wage Effect: Can Dynamic Monopsony Provide an Explanation? *Oxford Economic Papers* 48: 433-455.

Holmstrom Bengt and Paul Milgrom. 1991. Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics and Organization* 7: 24-52.

Iranzo, Susana, F. Schivardi, and E. Tosetti. 2008. Skill Dispersion and Firm Productivity: An Analysis with Employer-Employee Matched Data. *Journal of Labor Economics* 26: 247-285.

Jones, Benjamin F. 2008. The Burden of Knowledge and the Death of the Renaissance Man: Is Innovation Getting Harder? *Review of Economic Studies* 76: 283-317.

Kim, Sunwoong. 1989. Labor Specialization and the Extent of the Market. *Journal of Political Economy* 97: 692-705.

Kremer, Michael. 1993. The O-Ring Theory of Economic Development. *The Quarterly Journal of Economics* 108: 551-575.

Kremer, Michael and Eric Maskin. 1996. Wage Inequality and Segregation by Skill. *National Bureau of Economic Research*, Working Paper 5718.

Lazear, Edward P. 2005. Entrepreneurship. *Journal of Labor Economics* 23: 649-680.

Lazear, Edward P. 2009. Firm-Specific Human Capital: A Skill-Weights Approach. *Journal of Political Economy* 117: 914-940.

Lester, Richard. 1967. Pay Differentials by Size of Establishment. *Industrial Relations* 7: 57-67.

Maddala, Gangadharrao S. 1983. Limited-Dependent and Qualitative Variables in Econometrics. *Cambridge: University of Cambridge Press.*

Moog, Petra, Arndt Werner, Stefan Houweling and Uschi Backes-Gellner. 2015. The Impact of Skills, Working Time Allocation and Peer Effects on the Entrepreneurial Intentions of Scientists. *Journal of Technology Transfer* 40: 493-511.

Oi, Walter Y. and Todd L. Idson. 1999. Firm Size and Wages, in *Handbook of Labor Economics, Vol. 3*, by O. Ashenfeher and David Card Eds.: Elsevier, Chapter 33: 2165-2214.

Reardon, Sean F. and Glenn Firebaugh. 2002. Measures of Multigroup Segregation. *Social Methodology* 32: 33-67.

Rosen, Sherwin. 1978. Substitution and Division of Labour. *Economica* 45: 235-250.

Topel, Robert H. and Michael P. Ward. 1992. Job Mobility and the Careers of Young Men. *The Quarterly Journal of Economics* 107: 439-479.

Troske, Kenneth R. 1999. Evidence on the Employer Size-Wage Premium from Worker-Establishment Matched Data. *The Review of Economics and Statistics* 81: 15-26.

Tables and Figures

Table 1: LIAB 9310 Data - Descriptive Statistics

	Mean	25Percentile	75Percentile	St.Dev.	N (in millions)
Mean Ln Wages	4.18	3.92	4.73	0.87	12,300
Age	40.37	31	49	11.25	12,300
Female	0.39			0.49	12,300
EG	-1.25	-1.00	-1.21	0.50	12,300
S	0.66	0.88	0.52	0.26	12,300

Table 2: Sample comparison of the impact of heterogeneity measures: S and EG (LIAB data) and SD (LIAB+O*NET data)

	(1)	(2)	(3)
Without FE	Ln(Wage)	Ln(Productivity)	Number of Observations
Individual data + Establishment File	Yes	—-	12,264,916
Individual data + Establishment File + IAB Establishment Panel	Yes	Yes	3,480,924

Table 3: Wage regressions in the linked employer-employee data (LIAB-LM9310)

| | Without Fixed Effect | | | | With Fixed Effect | | |
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln Size	0.051***	0.033***	0.046***	0.037***	0.053***	0.037***	0.039***	0.023***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.004)	(0.004)	(0.003)	(0.003)
S Index		0.510***		0.224***		0.439***		0.385***
		(0.034)		(0.002)		(0.027)		(0.019)
Occup.Status Dummies	No	No	Yes	Yes	No	No	No	No
Occupation Dummies	No	No	No	No	Yes	Yes	No	No

Dependent variable: ln gross daily wage.
Independent variable: average of S Index across years.
Number of observations: 12,264,916.
All regressions control for education, age, age squared, year, industry and region.
Fixed-effect estimates at individual level.
Standard errors, clustered at the firm level, in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

35

Table 4: Wage regressions in the linked employer-employee data (LIAB-LM9310)

	Without Fixed Effect						With Fixed Effect	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln Size	0.051***	0.032***	0.046***	0.035***	0.053***	0.036***	0.039***	0.027***
	(0.004)	(0.004)	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)
EG Index		0.153***		0.087***		0.136***		0.075***
		(0.004)		(0.002)		(0.004)		(0.002)
Occup.Status Dummies	No	No	Yes	Yes	No	No	No	No
Occupation Dummies	No	No	No	No	Yes	Yes	No	No

Dependent variable: ln gross daily wage.
Independent variable: average of EG Index across years.
Number of observations:12,264,916.
All regressions control for education, age, age squared, year, industry and region.
Fixed-effect estimates at individual level
Standard errors, clustered at the firm level, in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Wage regressions in the linked employer-employee data (LIAB-LM9310 and O*NET)

	Without Fixed Effect				With Fixed Effect			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln Size	0.051***	0.035***	0.046***	0.038***	0.053***	0.039***	0.039***	0.028***
	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)
SD Imp Tasks		0.358***		0.169***		0.314***		0.278***
		(0.026)		(0.022)		(0.019)		(0.015)
Occup.Status Dummies	No	No	Yes	Yes	No	No	No	No
Occupation Dummies	No	No	No	No	Yes	Yes	No	No

Dependent variable: ln gross daily wage.
Independent variable: average of SD Imp Tasks across years.
Number of observations: 12,101,880.
All regressions control for the average importance of tasks, education, age, age squared, year, industry and region.
Fixed-effect estimates at individual level
Standard errors, clustered at the firm level, in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 6: Productivity regressions in the linked employer-employee data (LIAB-LM9310 + IAB Panel)

| | Without Fixed Effect | | | |
	(1)	(2)	(3)	(4)
Ln Lagged Size	0.116***	0.092***	0.106***	0.097***
	(0.017)	(0.018)	(0.020)	(0.017)
S Index		0.789***		
		(0.164)		
EG Index			0.248***	
			(0.103)	
SD Imp Tasks				0.763**
				(0.166)

Dependent variable: ln of (value added/number of employees).

Number of observations: 3,480,924.

All regressions control for age, age squared, education, occup. status, year, industry and region.

Column (4) controls for the average importance of tasks.

Standard errors, clustered at the firm level, in parentheses.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 7: Other Databases - Descriptive Statistics

BIBB/BAuA Employment Survey - Year 2006

	Mean	25Percentile	75Percentile	St.Dev.	N
Ln Wages	8.21	7.38	8.37	1.57	19,214
Count of Tasks	8.55	6	11	3.32	19,214
S	0.82	0.80	0.88	0.10	19,214

BIBB/BAuA Employment Survey - Year 2012

	Mean	25Percentile	75Percentile	St.Dev.	N
Ln Wages	7.81	7.50	8.70	0.67	17,460
Count of Tasks	9.46	7	12	3.37	17,460
S	0.82	0.80	0.88	0.10	17,460

Note: data from the main and supplemental surveys

NLSY79 Data

	Mean	25Percentile	75Percentile	St.Dev.	N
Ln Wages	9.89	9.47	10.60	1.19	39,007
SD Skills(std)	0.77	0.55	0.99	0.17	31,789
AVG Skills(std)	-0.16	-0.61	0.25	0.61	31,789

PIACC Data

	Mean	25Percentile	75Percentile	St.Dev.	N
Monthly Income	3.49	1	6	1.58	14,156
SD Skills(std)	0.33	0.19		0.43	20,029
AVG Skills(std)	0.00	-0.60	39 0.65	0.88	20,029

Table 8: Size and income regressions. Data from BIBB

	Ln Size			Ln Income	
	(1) 2005-2006	(2) 2011-2012		(3) 2005-2006	(4) 2011-2012
Specialist (Simpson's Index)	1.624***	2.167***		2.203***	1.336***
	(0.393)	(0.484)		(0.129)	(0.137)
Count of tasks	-0.100***	-0.126***		-0.010***	0.006**
	(0.010)	(0.012)		(0.003)	(0.003)
N.Observations	14,624	13,809		14,624	13,809

Dependent variable col (1) and (2): size classes; col (3) and (4): ln monthly income.
All regressions control for age, age squared, education, vocational degree, occupation,
year of the interview and industry.
Standard errors, clustered at the individual level, in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 9: Size and wage regressions. Data from NLSY79 and PIAAC

	Ln Size			Ln Wage	
	(1) NLSY79	(2) PIAAC		(3) NLSY	(4) PIAAC
SD Skills(std)	0.142*	0.112*		0.064*	0.108*
	(0.072)	(0.056)		(0.038)	(0.058)
AVG Skills(std)	-0.043	0.076***		0.192***	0.167***
	(0.045)	(0.014)		(0.011)	(0.015)
N.Observations	23,710	13,224		23,707	14,156

Dependent variable col (1): ln employees at location of current job.
Dependent variable col (2): size classes.
Dependent variable col (3) and (4): ln total income from wages and salary in the
past year.
All regressions control for education, occupation and industry.
NLSY79 regressions control for age, age squared and year. PIACC regressions
control for age (categorical variable).
Standard errors, clustered at the individual level, in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 10: Size and wage regressions. Data from LIAB Mover Model (MM9308)

	Ln Size			Ln Wage	
	(1)	(2)		(3)	(4)
Proximity	0.134***	0.082***		0.035***	0.044***
	(0.001)	(0.001)		(0.001)	(0.001)
Number of job changes	No	Yes		No	Yes

Dependent variable col (1) and (2): ln employees at location of last job.
Dependent variable col (3) and (4): ln (gross daily wage).
Number of observations: 3,817,416.
All regressions control for age, age squared, education, occup. status, industry and region.
Robust standard errors, in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

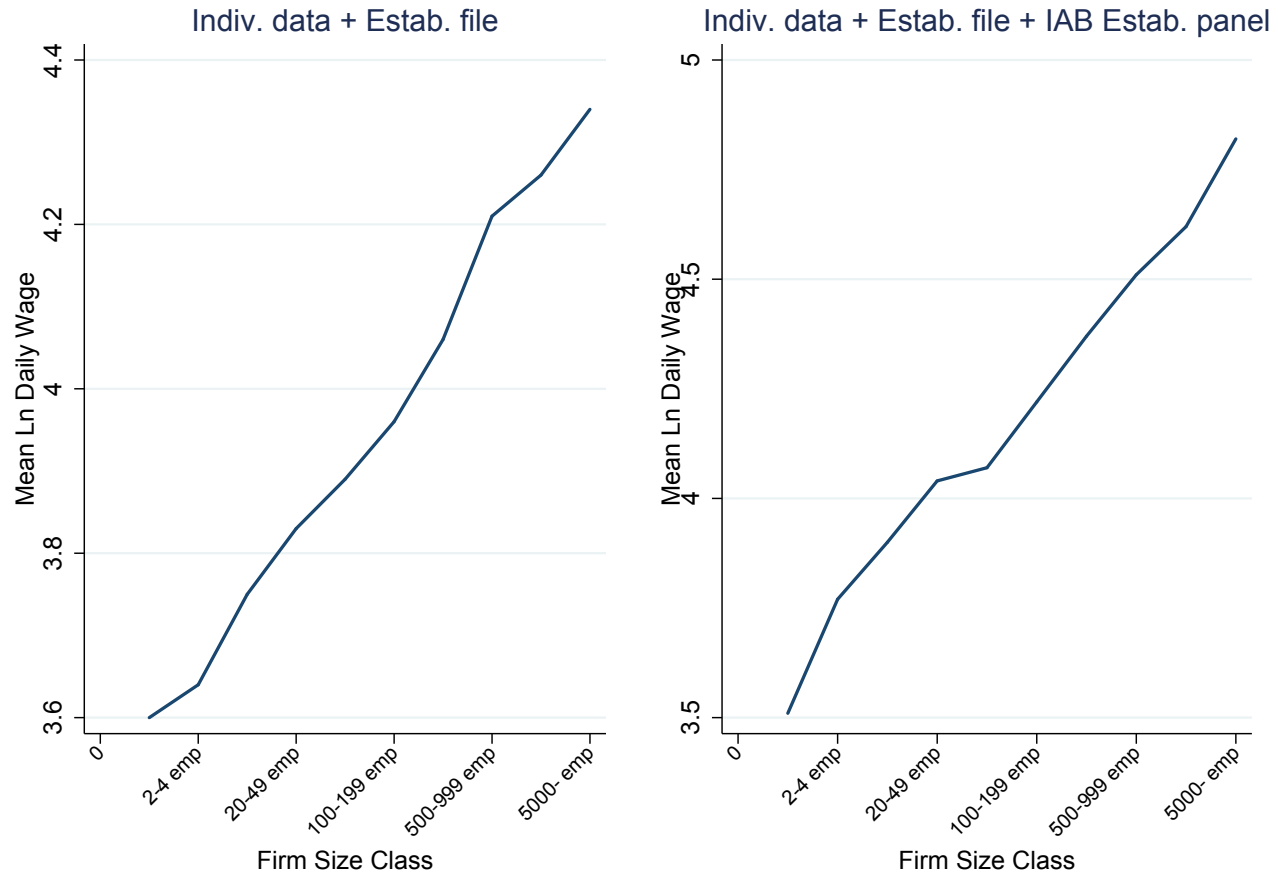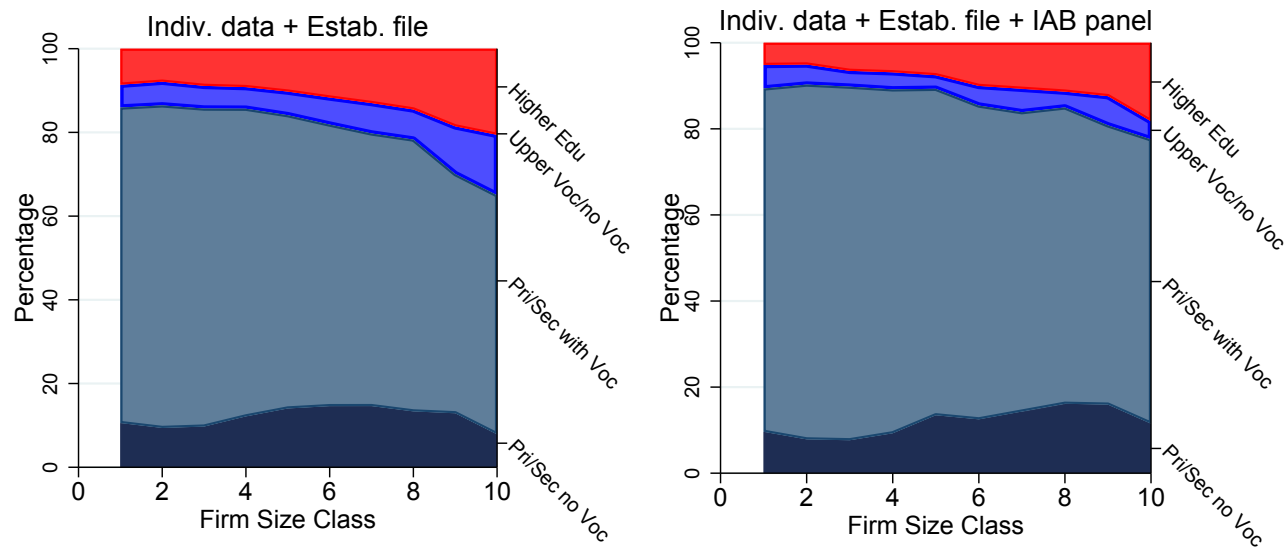## Figure 1: Mean Ln Wage by firm size

Figure 2: Percentage of employees in each education category (by size)

Appendix: Different proxies for labor specialization

MEASURE 1. The Simpson's Interaction Index (Reardon et al. 2002) with LIAB LM9310 and LIAB LM9310 + IAB Panel data

It is computed as a measure of concentration of occupations within firms. For simplicity, I work with its inverse additive, which represents heterogeneity of occupations (i.e. more specialized firms exhibit a higher S). S is given by:

$$S_{jt} = (-1) \sum_k \pi_{kt}(\pi_{kt} - 1) \tag{10}$$

where:

$k = 1, ., K$ are the occupations (job titles) assigned by employers to employees based on 3-digit codes from the Classification of Occupations. Systematic and Alphabetical Directory of Job Titles (KldB88).

$t = 1, ., T$ are the different years in which the firm is observed.

N_{kjt} is the number of workers in occupation k working in establishment j at time t.

N_{jt} is the total number of workers in establishment j, at time t.

$\pi_k = N_{kjt}/N_{jt}$ is the proportion of employees working in each occupation.

MEASURE 2. A dynamic version of the Ellison and Glaeser Index (1997) (Dumais et al. 2002) with LIAB LM9310 and LIAB LM9310 + IAB Panel data

This index is corrected by the Herfindahl-style measure to account for the fact that the concentration of occupations should be larger in small firms. It compares the degree of concentration of occupations within an establishment to the concentration of occupation of other establishments within the same economic activity. For simplicity, I work with its inverse additive, which represents heterogeneity of occupations (i.e. more specialized firms exhibit a higher EG). The EG index is given by:

$$EG_{jt} = (-1)\frac{G_{jt}/(1 - \sum_s \pi_{st}^2) - H_{it}}{1 - H_{it}} \tag{11}$$

where:

N_{kjst} is the number of workers in occupation k working in establishment j, sector s, at time t.

N_{jst} is the total number of workers in establishment j, sector s, at time t.

$k = 1, ., K$ are the occupations described in Measure 1.

i=1,.,n indicate the different establishments.

$s = 1, ., m$ represent 3-digit industry according to the WS73 or Classification of Economic Activities for the Statistics of the Federal Employment Services (1973). Before 2003 the variable contains the original values and from 2003 this information is continued or recoded (if necessary). It includes primary economic activities, manufacturing, construction and services.

$t = 1, ., T$ are the different split episodes, which are non-overlapping periods.

π_{jst} is the establishment occupation share computed as N_{kjst}/N_{jst}.

π_{st} is the average of π_{jst} within each industry.

$G_{jt} = \sum_s \pi_{st} - \pi_{jst})^2$ is the sum of squared deviations of establishment occupation share π_{jst} from a measure π_{st} of the share of occupations within a specific industry.

$H_{it} = \sum_k b_{jt}^2)/(\sum_k (b_{jt})^2$ is a Herfindahl-style measure where b_{jt} is the number of occupations within an establishment at different split episodes.

MEASURE 3. The dispersion of the importance of tasks with LIAB9310 and O*NET

Firstly, I standardize the s task variables and then I compute the standard deviation of the importance imp of the 41 tasks for each occupation.

$$sd_{imp_{(task(s))_b}} = sd(imp_(task1)_{(std)}, ..., imp(task41)_{(std)}) \qquad (12)$$

where subscripts s are the 41 tasks, b is the occupation present in a firm. Subsequently, I consider all the employees i who work for firm j and compute the average of the standard deviations of the importance of tasks $(sd_{imp_{(task(s))_b}})$ by firm j and by year t:

$$sd_{AVGjt} = mean(sd_{imp_{(task(s(std)))}})_{jt} \qquad (13)$$

Specialization is considered a firm pattern (i.e. time invariant), therefore after computing these measures of occupation heterogeneity, I work with averages across years.

MEASURE 4. The Simpson's Interaction Index (Reardon et al. 2002) with BIBB data

It is computed as a measure of individual's focus on certain tasks performed. S is given by:

$$S(tasks)_i = \sum_k \pi_k(\pi_k - 1) \qquad (14)$$

where:

$k = 1, ., K$ are the standardized tasks performed.

Tki is the time individual i spend in task k.

T_i is the individual i work time.

$\pi_k = T_{ki}/T_i$ is the share of the time spent in tasks performed frequently or seldom.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1544	Jonathan Colmer	Weather, Labor Reallocation and Industrial Production: Evidence from India
1543	Andrew B. Bernard Esther Ann Bøler Swati Dhingra	Firm-to-Firm Connections in Colombian Imports
1542	Gerard H. Dericks Hans R.A. Koster	The Billion Pound Drop: The Blitz and Agglomeration Economics in London
1541	Andrew B. Bernard Andreas Moxnes	Networks and Trade
1540	Does Gentrification Displace Poor Households? An 'Identification-Via-Interaction' Approach	Sevrin Waights
1539	Zhuan Pei Jörn-Steffen Pischke Hannes Schwandt	Poorly Measured Confounders Are More Useful On the Left than On the Right
1538	Konstantin Büchel Stephan Kyburz	Fast Track to Growth? Railway Access, Population Growth and Local Displacement in 19th Century Switzerland
1537	Stephen Gibbons Maria Sánchez-Vidal Olmo Silva	The Bedroom Tax
1536	Pawel Adrjan Brian Bell	Pension Shocks and Wages
1535	Tom Kirchmaier Stephen Machin Matteo Sandi Robert Witt	Prices, Policing and Policy: The Dynamics of Crime Booms and Busts

1534	Laurent Bouton Paola Conconi Francisco Pino Maurizio Zanardi	Guns, Environment and Abortion: How Single-Minded Voters Shape Politicians Decisions
1533	Giulia Giupponi Stephen Machin	Changing the Structure of Minimum Wages: Firm Adjustment and Wage Spillovers
1532	Swati Dhingra Rebecca Freeman Eleonora Mavroeidi	Beyond Tariff Reductions: What Extra Boost From Trade Agreement Provisions?
1531	Doruk Cengiz Arindrajit Dube Attila Lindner Ben Zipperer	The Effect of Minimum Wages on Low-Wage Jobs: Evidence from the United States Using a Bunching Estimator
1530	Stephen Gibbons Vincenzo Scrutinio Shqiponja Telhaj	Teacher Turnover: Does it Matter for Pupil Achievement?
1529	Ghazala Azmat Stefania Simion	Higher Education Funding Reforms: A Comprehensive Analysis of Educational and Labour Market Outcomes in England
1528	Renata Lemos Daniela Scur	All in the Family? CEO Choice and Firm Organization
1527	Stephen Machin Matteo Sandi	Autonomous Schools and Strategic Pupil Exclusion
1526	Stephan E. Maurer	Oil Discoveries and Education Spending in the Postbellum South