Wage and Productivity Dispersion: Labor Quality or Rent Sharing?*

Jesper Bagger †
Royal Holloway, University of London and LMDG

Bent Jesper Christensen ‡
Aarhus University, LMDG and CREATES

Dale T. Mortensen §
Northwestern University, Aarhus University, NBER and LMDG

February 15, 2010

Abstract

Wage and labor productivity differ across firms, and more productive firms tend to pay higher wages. We consider a model that allows for differences in capital, employment and labor quality as well as rent sharing, all of which should help explain these observations. We estimate the model using detailed matched employer-employee data from the manufacturing sector in Denmark. The production function estimation is embedded in a structural equation system involving worker, firm, time, and occupation effects from an individual wage decomposition and accounting for labor input components that are substitutes and complements, while accommodating stochastically varying factor productivity. We find that both input heterogeneity and intrinsic differences in total factor productivity across firms are important explanations. In the case of Manufacturing, about 41% of the dispersion in log value added per worker is attributable to cross-firm differences in the levels of capital per worker, while another 39% of the variation stems from intrinsic TFP differences. Only 5% is associated with quality differences in the labor input. In the case of individual log wages, 70% of the variation is due to individual characteristics, whereas only 13% is attributable to firm differences. Our results suggest that there are major gains to reallocation of labor from less to more productive firms. Rent sharing enhances the reallocation process by inducing wage dispersion that motivates worker search. The relatively small contribution of firm heterogeneity to individual wage dispersion is thus consistent with the inefficient allocation of labor across firms.

Keywords: Matched employer-employee panel data, Production function input differences, Two-step GMM, Wage bargaining, Worker complementarity

JEL classification: C33, E23, E24, J31

*We thank John Kennan, Jean-Marc Robin, participants at the 2009 Structural Labor Market Models and Policy Analysis conference at Sandbjerg Manor in Denmark, and seminar participants at Royal Holloway, Oxford University and University of Essex for helpful and constructive comments.

†Department of Economics, Royal Holloway College, University of London, Egham, Surrey TW20 0EX, United Kingdom; E-mail: jesper.bagger@rhul.ac.uk

‡School of Economics and Management, Aarhus University, 1322 University Park, DK-8000 Aarhus C, Denmark; E-mail: bjchristensen@econ.au.dk

§Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, Illinois 60208, U.S.A.; E-mail: d-mortensen@northwestern.edu

1 Introduction

Both labor productivity and the average wage paid vary substantially across firms and are positively correlated. In the case of Denmark, the 90/10 percentile ratio for the distribution of labor productivity is 2.3, and for the distribution of average wage bill paid per worker it is 1.8, while the cross firm correlation between the two variables is 0.61.[1] In the case of France, similar relationships by industry are documented by Postel-Vinay and Robin (2006). Comparable results by industry based on Danish matched employee-employer data are reported in this paper.

Recent empirical research on foreign trade activities of firms provides further evidence of substantial and important firm productive heterogeneity. Only a few firms supply products outside their home country and those that do tend to be larger and more productive than the rest. Exporting firms appear to pay significant wage premiums. Finally, exporting firms are found scattered across all industries in the manufacturing sector. Bernard and Jensen (1995) and Bernard, Eaton, Jensen, and Kortum (2003) document these facts using microeconomic data on sales, inputs, and foreign trade for the US. The same observations are replicated for France by Eaton, Kortum, and Kramartz (2004, 2008) and by Pedersen (2009) for Denmark.

Although researchers have been aware of these properties of firm micro data for some time (Bailey, Hulton, and Campbell, 1992 and Foster, Haltiwanger, and Krizan, 2001), there is no consensus regarding the explanation of the positive wage-productivity premium or the correlation between firm productivity and wages. Some argue that the observed productivity and wage dispersion simply reflects differences in the composition and quality of labor and other productive factors, as would be implied by the simplest competitive model (Murphy and Topel, 1990). Others contend that productivity varies systematically across firms for whatever reasons, that these differences and labor market friction induce differential quasi rents, and that these are shared in some way between employer and worker.[2] Undoubtedly, both reasons play a role, but so far there has been no quantitative assessment of the relative importance of the two.

[1] These numbers relate to the data used in our empirical analysis. See section 3 for a description of this data including sample selection issues.

[2] These include Burdett and Mortensen (1998), Postel-Vinay and Robin (2002a,b), Cahuc, Postel-Vinay and Robin (2006), Cahuc, Marque, and Wasmer (2007), and Mortensen (2003, 2009), among others.

In this paper we construct and estimate a model that incorporates both explanations of the dispersion in wage and productivity, and the relation between the two, in order to determine the extent to which each of the two stories can account for the observations in the case of Danish data. Our model of wage determination is composed of the firm's production function with inputs that include capital and quality adjusted labor and a wage equation that allows for rent sharing. The wage equation is based on the Stole and Zwiebel (1996) model of bilateral bargaining between an employer and each of its employees.

In the Stole and Zwiebel model, the number of potential employees is predetermined in the short run by those who stay and new applicants. The game is composed of a potentially infinite sequence of negotiation rounds, as originally proposed in Rubinstein (1982). Given that participation constraints are satisfied, the game continues until an agreement is reached. In the version of the game proposed by Mortensen (2009), the employer cannot recruit a substitute and the worker cannot search for an alternative job while negotiating. Hence, delay is the only default option while negotiations are taking place. The threat to stop negotiating is not credible, given that there is 'money on the table,' as in Hall and Milgrom (2008).

The literature on the impact of 'human capital' variables on wages and productivity is huge. Early attempts to take account of changes in human capital variables in the decomposition of aggregate productivity growth by Jorgenson et al. (1987) concluded that these were small contributors. Although the availability of matched worker-employer data with information on education and labor market histories of individual workers offers the prospect of directly testing the hypothesis that firm productivity differences primarily reflect differences in the quality of inputs, surprising little research has been done on the subject. However, we are aware of two recent projects that exploit these data bases: Fox and Smeets (2009) use Danish matched employer-employee data and Irarrazabal et al. (2009) use similar data for Norway.

Fox and Smeets estimate Cobb-Douglas and translog production functions with quality weighted employment and capital inputs. The worker characteristics available in the Danish data for each individual employee include gender, years of completed education, total labor market experience, industry experience, and firm tenure, as well as age. These variables are included as explanation of the total labor input measured in efficiency units with weights determined by the data. The estimated firm specific residual in this specification is interpreted as the firm's total factor productivity (TFP). The authors then compare the variance explained by these labor quality variables above and beyond that implied by simply including a measure of total firm employment as the labor input and capital. The estimation is done separately by industry, including some services as well as manufacturing.

Fox and Smeets find that the weights on the 'human capital' variables are significant, well determined, and of the expected signs. However, including them explains relatively little of the variance in firm productivity in any of the industries. Averaging over the six manufacturing industries, the ratio of the 90th to the 10th percentile of the Danish distribution of the standard TFP measure is large, at 3.74. The ratio is reduced, but only to 3.36, when the human capital variables are included. They obtain similar results using the wage bill, a wage weighted measure of employment, to correct for labor input quality. They conclude that the observable component of 'input quality' explains very little of the dispersion in firm productivity observed in Danish data.

Irarrazabal et al. approach the issue of labor input quality from the perspective of the empirical trade literature, where the fact that the firms engaged in foreign trade are more productive and pay more is documented. Hence, they estimate the extent to which observable worker characteristics explain the 'productivity premium.' Both their methods and data are similar to those of Fox and Smeets. They too include measures of labor force age, education, and tenure, as well as the capital stock of each firm. Using either actual worker characteristics or the wage bill in the analysis, they conclude that labor input quality explains about 25% of the average productivity differential between exporting an non-exporting firms in Norway. Although they conclude that the potential for 'gains from trade' are overstated by the measured TFP differences, it is clear that labor input quality differentials fail to explain the bulk of the differences in productivity.

The method introduced in this paper contributes to the estimation of production functions by embedding them in a structural equation system involving worker, firm, time, and occupation effects from an individual wage decomposition and accounting for labor input components that are substitutes and complements, while accommodating stochastically varying factor productivity. The structural model allows for differences in input quality as well as rent sharing, both of which may potentially help explain the observations that wage and labor productivity are dispersed across firms, and that more productive firms tend to pay higher wages. From our empirical results, which focus on the manufacturing sector only, both input heterogeneity and intrinsic differences in total factor productivity across firms are important for dispersion. We find that 41% of the dispersion in log value added per worker within the manufacturing sector is attributable to cross-firm differences in the levels of capital per worker, while another 39% of the variation stems from intrinsic TFP differences across firms. Only a smaller portion observed dispersion in log value added per worker (5%) is associated with quality differences in the labor input. These results suggest that that there are major gains to reallocation of labor from firms with low marginal labor productivity to firms with high marginal labor productivity. The same would not be the case if the dispersion in marginal labor productivity were due to differences in labor quality alone. Rent sharing provide a link between individual wages and firms' marginal labor productivity and thus ties the dispersion in labor productivity to the wage distribution. Hence, rent sharing is a potentially important vehicle for reallocation as wage dispersion motivates job search. We provide a decomposition of individual log wages in the manufacturing sector and find that 70% of the individual log wage variation is due to individual characteristics, whereas only 13% is attributable to firm differences (i.e. rent sharing). The relatively small contribution of firm heterogeneity to individual wage dispersion is thus consistent with the inefficient allocation of labor across firms.

The rest of the paper is laid out as follows. The model is introduced in Section 2. Section 3 describes the construction of our MEE panel data set. Section 4 presents the estimation method and the empirical results. Section 5 concludes. Some further details on the data are provided in an Appendix.

2 The Model

In the bargaining model, each 'period' may be thought of as divided into three subintervals. In the first, the employer bargains individually with all the workers available to enter an employment agreement. In the second subperiod, production takes place, using the subset of workers who agree to a bargaining outcome. Search and recruiting occur in the third subperiod, thus determining the allocation of job applicants among firms at the beginning of the next period. Information is complete and symmetric in the sense that both sides know the values of their match and all options available at each node of the bargaining game.

In the first subperiod, bargaining between each worker-employer pair takes place in a sequence of rounds as in Rubinstein (1982). The crucial assumption is that worker applicants cannot search for another employment opportunity and firms cannot recruit more applicants during the negotiations. Finally, employers commit not to respond to any alternative employment opportunity available to the worker. Hence, those workers who were employed must choose to bargain either with her current employer or the alternative. The default position for either party in any round of the negotiation is therefore only delay, as in Hall and Milgrom (2008). During a potential delay, taking place during the production phase of the period, the worker engages in home production yielding output at rate b, and production takes place with one worker less.

In the special case of identical workers, the gross profit flow of the firm can be represented as

$$\pi(n,p) = pf(n) - wn, \tag{1}$$

where n is employment, p is the firm's factor productivity, w is the wage, and $f(n)$ is the baseline production function, an increasing concave function of employment. We introduce bargaining power by supposing that nature selects the worker to propose the deal with probability β and the employer with complementary probability $1-\beta$ at the beginning of each negotiation round. As information is complete by assumption, the proposer offers a wage that makes acceptance at least as attractive to the other party as continued negotiation. Let the value of home production, represented as b, denote the value of delay to the worker, and assume there is no out of pocket cost or benefit of delay to the employer. Then Mortensen (2009) finds that bargaining ends after the first round of negotiation with an expected division of the flow surplus that can be represented as the solution to the differential equation

$$\beta\pi_n(n,p) = \beta\left(pf'(n) - w - \frac{\partial w}{\partial n}n\right) = (1-\beta)\,(w-b)\,. \tag{2}$$

As Stole and Zwiebel (1996) show, the solution to (2), the wage bargaining outcome function, is

$$w(n,p) = (1-\beta)b + p\int_0^1 z^{\frac{1-\beta}{\beta}} f'(zn)dz. \tag{3}$$

In the constant marginal product case, the wage reduces to the average of the outside option b and the value of the marginal product, as in the canonical search and matching model (Pissarides (2000)). The Cobb-Douglas case, $f(n) = n^{\alpha}$, produces a linear relation between the firm's labor productivity and wage,

$$\begin{aligned} w(n,p) &= (1-\beta)b + p\alpha\int_0^1 z^{\frac{1-\beta}{\beta}}(zn)^{\alpha-1}dz \\ &= (1-\beta)b + \alpha\left(\int_0^1 z^{\frac{1}{\beta}+\alpha-2}dz\right)\frac{pf(n)}{n} \\ &= (1-\beta)b + \frac{\beta\alpha}{1-\beta+\alpha\beta}\left(\frac{pf(n)}{n}\right). \end{aligned} \tag{4}$$

Thus, the wage is equal to the average product when workers have all the power ($\beta = 1$) and the default option when they have none ($\beta = 0$). If the elasticity of output with respect to employment is $\alpha = 2/3$, then the weight on the average product is 0.4 in the symmetric bargaining power case ($\beta = \frac{1}{2}$). In general, the wage is an increasing function of both workers' share of the rent β and the labor elasticity of output α.[3]

Cahuc, Marque, and Wasmer (2007) generalize the outcome of the Stole and Swiebel bargaining problem to allow for any number of different types of labor that are imperfect substitutes in general, and other quasi-fixed factors, such as capital. Given a general production function of the form

$$Y = pf(K, \mathbf{L}), \tag{5}$$

[3]Of course, the bilateral bargaining outcome (3) determines the wage if and only if the it is consistent with participation conditions. Mortensen (2009) shows that these do not bind in equilibrium.

where p represents firm TFP, K is the capital stock, and $\mathbf{L}=(L_1, L_2, ..., L_H)$ the vector of labor inputs, the generalization of (3) to the case of H different worker types is

$$w_h(K, \mathbf{L}, p) = (1-\beta_h)b_h + p\int_0^1 z^{\frac{1-\beta_h}{\beta_h}} f_h(K, \mathbf{L}A_h(z))dz, \quad h = 1, ..., H, \tag{6}$$

where $pf_h = \partial Y/\partial L_h$ is the value of marginal product of type h labor and $\mathbf{L}A_h(z)$ is the vector

$$\mathbf{L}A_h(z) = \left(L_1 z^{\frac{\beta_1}{1-\beta_1}\frac{1-\beta_h}{\beta_h}}, L_2 z^{\frac{\beta_2}{1-\beta_2}\frac{1-\beta_h}{\beta_h}}, ..., L_H z^{\frac{\beta_H}{1-\beta_H}\frac{1-\beta_h}{\beta_h}}\right), \quad h = 1, ..., H, \tag{7}$$

the product of the employment vector $\mathbf{L}$ and the diagonal matrix $A_h(z)$ with representative diagonal entry $z^{\frac{\beta_n}{1-\beta_n}\frac{1-\beta_h}{\beta_h}}$ for $n = 1, ..., H$

We consider workers as allocated to different functions within the firm, such as management, support staff, production workers, etc. Theory suggests that there should be some complementarity between these types of job categories. We assume that within categories, individual workers differ with respect to skill, but are perfect substitutes. Skill within type depends on observable person characteristics, such as education and labor experience, as well as unobservable. Let a_{ih} represent the skill of individual i in job category h. If one views w_h as the piece wage per unit of skill in job type h, then the wage of individual i in job category h is $w_{ih} = a_{ih}w_h$, which implies that the log of individual wage rates can be written as a linear function of a worker effect and a firm effect,

$$\ln w_{ih} = \ln a_{ih} + \ln w_h(K, \mathbf{L}, p), \tag{8}$$

where the firm effect, the second term, depends on the employing firm's input composition and TFP.

Our specification (8) can be regarded as a generalization of the Abowd, Margolis, and Kramartz (1999) (henceforth AMK) wage decomposition to incorporate potentially important firm characteristics as determinants of the firm effects. Since total labor input of type h is given by

$$L_h = \sum_{i\in\mathcal{I}_h} a_{ih}, \tag{9}$$

where $\mathcal{I}_h$ is the set of workers of type h in the firm, the specification implies that labor input quantities are determined by worker heterogeneity, which may be observed or unobserved to an econometrician. In anticipation of our estimation procedure we note that fixed worker effects are generally identified from repeated observation on a given worker in different firms. Hence, our matched employer-employee data will allow us to construct better measure of firm heterogeneity in inputs than those used in other studies in the literature.

2.1 A Special Case

Consider the case where the production function is log linear in capital and a constant elasticity of substitution (CES) labor aggregate. Thus, let

$$\ln Y = \ln p + \alpha_K \ln K + \frac{\alpha_L}{\rho} \ln \left(\sum_{h=1}^{H} L_h^{\rho} \right), \quad \rho \leq 1, \tag{10}$$

where α_K and α_L represent the output elasticities with respect to capital and the labor aggregate, respectively, and the parameter ρ determines the elasticity of substitution between job categories.[4] In this case, the marginal product with respect to category h labor is proportional to output per unit of labor of category h given the appropriately defined composition of the labor aggregate. Specifically, for $h = 1, ..., H$,

$$\begin{aligned} \frac{\partial Y}{\partial L_h} &= p f_h(K, \boldsymbol{L}) = \alpha_L p K^{\alpha_L} \left(\sum_{n=1}^{H} L_n^{\rho} \right)_i^{\frac{\alpha_L}{\rho}-1} L_h^{\rho-1} \\ &= \alpha_L \frac{L_h^{\rho-1}}{\sum_{n=1}^{H} L_n^{\rho}} Y = \alpha_L \frac{L_h^{\rho}}{\sum_{n=1}^{H} L_n^{\rho}} \frac{Y}{L_h}. \end{aligned} \tag{11}$$

If bargaining power is the same across job categories, $\beta_h = \beta$ for $h = 1, ..., H$, then because $\mathbf{L}A_h(z) = z\mathbf{L}$ we have

$$\begin{aligned} \int_0^1 z^{\frac{1-\beta}{\beta}} p f_h(K, \mathbf{L}A_h(z)) dz &= \int_0^1 z^{\frac{1-\beta}{\beta}} \alpha_L \frac{(zL_h)^{\rho-1}}{\sum_{n=1}^{H} (zL_n)^{\rho}} p K^{\alpha_K} \left(\sum_{n=1}^{H} (zL_n)^{\rho} \right)^{\frac{\alpha_L}{\rho}} dz \\ &= \alpha_L \frac{L_h^{\rho}}{\sum_{n=1}^{H} L_n^{\rho}} \frac{Y}{L_h} \int_0^1 z^{\frac{1-\beta}{\beta} + \alpha_L - 1} dz \\ &= \left(\frac{\alpha_L \beta}{1 - \beta + \alpha_L \beta} \right) \frac{L_h^{\rho}}{\sum_{n=1}^{H} L_n^{\rho}} \frac{Y}{L_h}. \end{aligned} \tag{12}$$

[4] Because $L_h = \sum_{i \in \mathcal{I}_h} a_{ih}$ by definition, the symmetric specification of the labor aggregator with respect to the elements of $\mathbf{L} = (L_1, ..., L_H)$ imposes no loss of generality.

Because the Stole-Zwiebel wage per unit of h type labor is

$$\begin{aligned} w_h(K, \mathbf{L}, p) &= (1-\beta)b_h + \int_0^1 z^{\frac{1-\beta}{\beta}} p f_h(K, \mathbf{L}A_h(z))dz \\ &= (1-\beta)b_h + \frac{\beta\alpha_L}{1-\beta+\beta\alpha_L}\frac{L_h^\rho}{\sum_{n=1}^H L_n^\rho}\frac{Y}{L_h}, \quad h = 1, ..., H, \end{aligned} \tag{13}$$

the firm effect in the wage equation (8) for each category h depends only on output per unit of labor of that category and the share of the type in the labor aggregate, appropriately scaled. Furthermore, the dependence of the firm effect on firm labor productivity for each job category is increasing in the elasticity of output with respect to labor input, the worker bargaining power, and the category's share in the firm's total labor input. Although the assumption of equal bargaining power is not particularly satisfying, the integral in (6) does not have a closed form solution (as it does in (13)) when $\beta_h \neq \beta$. Still, with the help of numerical integration one can use the more general specification in empirical work.

2.2 Firm Dynamics

In this section, we relax the assumption that inputs and total factor productivity are fixed. Instead, TFP is a stochastic process and capital and labor are quasi-fixed factors of production. Formally, we consider a discrete time formulation in which $t = 1, 2, ...$ indexes periods and the TFP sequence $\{p_t\}$ is a first order Markov process. We assume differences in the timing of hiring and investment decisions, intended to reflect the fact that capital is relatively more fixed than employment. Specifically, hires and wage bargaining take place prior to production in each period, after the value of current TFP is observed, while investment decisions in period t are are realized in period $t+1$.

Although one can easily generalize the formulation to any finite number of job categories, for expositional simplicity we sketch a the basic model only for a single category. Given that hires, separations, and bargaining take place at the beginning of each period, labor input in period t is given by

$$L_t = H_t + (1-s)L_{t-1}, \tag{14}$$

where H_t represents new hires and s is the separation rate. Capital evolves according to the law of motion

$$K_{t+1} = I_t + (1-\delta)K_t, \tag{15}$$

where I_t represents gross additions to physical capital determined in period t and to be installed in period $t+1$ and δ is a fixed depreciation rate. The hiring and investment decision in period t are made contingent on L_{t-1}, K_t, and p_t. Of course, as in (4),

$$w(K_t, L_t, p_t) = (1-\beta)b + \frac{\beta\alpha}{1-\beta+\alpha\beta}\frac{p_t f(K_t, L_t)}{L_t} \tag{16}$$

is the outcome of the wage bargaining in period t which is simultaneous with the total labor input.

Suppose that the costs of hiring and gross investment in period t take the forms $c_L(v_t/L_{t-1})L_{t-1}$ and $c_K(I_t/K_t)K_t$, where $c_L(\cdot)$ and $c_K(\cdot)$ are increasing and convex. Specifically, the cost functions are increasing, convex, and homogenous of degree one in stock and gross flow. The value of the firm associated with the optimal decisions solve the Bellman equation

$$\begin{array}{rcl} V(L_{t-1}, K_t, p_t) & = & \max\limits_{v_t, I_t, L_t, K_{t+1}} \left\{ \begin{array}{c} p_t f(K_t, L_t) - w(K_t, L_t, p_t)L_t \\ -c_L\left(\frac{H_t}{L_{t-1}}\right)L_{t-1} - c_K\left(\frac{I_t}{K_t}\right)K_t \\ +\gamma E\left\{V(L_t, K_{t+1}, p_{t+1})|p_t\right\} \end{array} \right\} \\ & & \text{subject to equations (14) and (15).} \end{array} \tag{17}$$

It should be pointed out that we do not need to take a stand on the particulars of this dynamic formulation for the purposes of the estimation that follows other than the specification of what the employer knows when making the hiring and investment decisions. Hence, more complicated and possibly more realistic formulations, one that explicitly accounts for employed worker search for example, are consistent with the estimation procedure that follows so long as the timing assumptions hold.

3 Data

The empirical analysis is carried out on Danish register-based Matched Employer-Employee (MEE) panel data. We rely on three different data sources. Employer data are secured from a firm level panel with accounting and balance sheet information for the period 1995-2006. The accounting data set derives from an annual survey conducted by Statistics Denmark subject to a specific sampling scheme detailed below. Employee data are drawn from the Integrated Database for Labor Market Research (IDA), an individual level annual panel containing all individuals aged 15 through 70 in Denmark. We use the IDA files for the years 1995-2006 covered by our employer data set. IDA contains one observation for each year, for each individual. If a worker is employed in the last week of November, IDA provides information on the annual average hourly wage and the ID of the employer. In addition, IDA is rich on background information on individuals (including a unique person ID), whereas employer information is limited, consisting of firm ID, ownership structure, and industry indicators. The third and final data element is the firm-IDA integration (FIDA) that we use to link employers and employees in the last week of November of each year. Statistics Denmark uses various different sets of employer IDs for different purposes, and FIDA provides a key between the different definitions. We use FIDA to merge the detailed employer information in the firm accounting data with the detailed worker information in IDA, this way constructing a unique and very comprehensive MEE panel data set on firms and workers. In the following, we provide some more detail on each of these data sources, the construction of our MEE panel, and the sample selection rules imposed. We end the section with descriptive statistics for the resulting sample for analysis.

3.1 Data sources

3.1.1 Employer data

We obtain accounting and balance sheet information on a sample of firms from an annual survey conducted by Statistics Denmark over the period 1995-2006. Industry coverage in the survey increases over time, starting in the initial year 1995 with Manufacturing, Construction, Wholesale & Retail, and Transportation & Telecommunications. Coverage is gradually expanded until in 1999 most industries are covered, a few exceptions being Agriculture, Public Services, and parts of the financial sector (source: Statistics Denmark). As a consequence of the expansion of the industry coverage, the number of firms in the survey increases from around 4,600 in 1995 to around 8,000 in 1999. After 1999 the industry coverage and the number of included firms is stable. The unit of observation is a firm-year and the survey has a rolling panel structure where firms are selected based

November workforce	Fraction of employers sampled	Years included/ excluded
0-4 employees	0.00	-
5-9 employees	0.10	1/9
10-19 employees	0.20	2/8
20-49 employees	0.50	3/3
>49 employees	1.00	-
Previous year's turnover	**Fraction of employers sampled**	**Years included/ excluded**
>DKK 100 mill. (wholesale: >DKK 200 mill.)	1.00	-

Table 1: Sampling scheme for Statistics Denmark's accounting data survey. Source: Statistics Denmark.

on the size of their last week of November workforce. The specific sample selection rules applied are given in Table 1. If sampled, a firm is required to submit a standardized balance sheet to Statistics Denmark. Hence, we are able to compute value added, capital stock, wage bill, material inputs and firm size measures. These computations are detailed in Appendix A.

3.1.2 Employee data

We obtain data on workers from IDA for the years 1995-2006 covered by our employer data. IDA is comprised of several files with distinct types of information: Person files, establishment files, and job files. The person files contain annual information on a wide range of socio-economic variables for the entire Danish population aged 15-70. The establishment files contain annual information on all establishments with at least one employee in the last week of November in each year. The job files provide information on all jobs that are active in the last week of November in each year, including an estimate of the average hourly wage. We use education and occupation information from the person files, industry codes from the establishment files, and average hourly wage from the job files.

3.1.3 Matching the employer data with the employee data: The FIDA file

The firm ID used by Statistics Denmark in the accounting survey is different from that used for the ID of employing firm for each job in IDA. We use the link-file FIDA as a key between the different employer ID definitions. Industry coverage in FIDA follows that of the accounting data survey. Within each year, FIDA contains all last week of November employment relationships in the covered industries. It is available for the required years 1995-2006, thus enabling us to match employers and employees in the last week of November in each year to create our MEE panel.

We first merge accounting data onto the link-file FIDA using the firm IDs. All firms in the accounting data are matched to the FIDA panel. Not every observations in FIDA is matched to an observation in the accounting data, since this is survey based, whereas FIDA includes all firms in the covered industries. The actual fraction matched is about 30% of all FIDA observations. We retain all the original FIDA observations. Next, we merge the FIDA panel with the IDA person files. All observations in the FIDA panel are matched with IDA person information. Since IDA is a population data set, whereas FIDA only covers selected industries, there are observations in the IDA person files that cannot be matched with FIDA. We retain the original FIDA observations. Merging the FIDA panel with the IDA job files results in 83 percent of all FIDA observations being matched with an observation from the IDA job files. No match is possible if there was no employment relationship in the last week of November. Conversely, 84 percent of the observations in the IDA job files are matched with a FIDA observation. Still, we retain all the original FIDA observations. Finally, we match the FIDA file with the IDA establishment files. Here, we match 94 percent of all FIDA observations to an IDA establishment. Conversely, we match 99 percent of all IDA establishment observations to an observation in FIDA.

3.2 Sample selection

The following sample selection criteria are imposed. First, we focus on the three two-digit NACE code industries Manufacturing, Construction and Wholesale & Retail. Second, we retain only observations with valid education codes. Third, we drop observations with invalid occupation

	Manufacturing	Construction	Wholesale & Retail
Coverage	1995-2006	1995-2006	1995-2006
# of observations	4,578,829	1,622,060	4,036,078
# of workers	893,602	334,697	1,073,145
# of firm-years	149,377	178,969	437,116
# of firm-years w/ acc. data	29,466	12,935	18,307
Avg. frac. of obs. w/ acc data	0.77	0.40	0.30
Avg. frac. of firm-years w/ acc data	0.20	0.07	0.04

Table 2: Summary statistics on individual level analysis panels.

information and all workers classified as self-employed. Fourth, we drop all observations where either value added, capital stock, material cost, number of work hours in production, or individual wage paid is non-positive. Further, to purge the data for outliers we trim the industry specific annual value added distributions at the top and bottom 1 percent. All nominal variables (value added, capital, material cost, and wages) are inflated to 2006 levels using the implicit deflator in value added within Manufacturing. Finally, we stratify our panel into the three analysis data sets given by the industries we focus on. Table 2 provide some summary statistics on the sample sizes of the individual level analysis panels. The fact that large firms are over-sampled implies that the fraction of observations for which we have accounting data information is markedly larger than the fraction of firms for which this information is available.

Our empirical analysis is carried out in two steps (see section 4 for details on our estimation strategy). In the first step we decompose individual log-wage variance into worker effects, firm-time effects and residual individual wage dispersion. In the second step we construct a firm-level panel data to analyze dispersion in firm specific compensation and productivity. The first step of the empirical analysis is carried out on the full individual level panel data described above. It entails estimation of a “long” log-wage regression with individual worker and firm-year indicators. To ensure identification (that is, to ensure a unique solution to the resulting system of normal equations) we restrict attention to the largest group of "connected" workers and firm-years (see Abowd, Creecy and Kramarz, 2002). As it turns out, almost all our workers and firm-years are connected, and this restriction leads to a very small loss of observations (see Table 8 in Appendix B for details). The second step of our empirical strategy is based on a firm-level panel constructed from the individual level panel. In particular, this firm-level panel contains all firms-years for which we have accounting data information and for which a one-period lagged observation is available. The latter condition ensures that we can take one-period first differences at the firm level, which will play a key role in the identification of the structural parameters (again, we refer the reader to section 4 for details on identification and estimation).

3.3 Descriptive statistics

We now present a number of facts regarding the observed wage and productivity dispersion in our analysis samples. Here, productivity is measured as the firm-specific average hourly value added per FTE. We measure FTE as the number of annual full time workers that have gone into the production. To obtain an hourly productivity measure comparable to the computed firm wages we divide average annual value added by the Danish Industry Confederation's annual hours norm (hours/year).[5]

The first empirical goal of this paper is to elicit new information on the structure of dispersion in value added per worker, firm wages, and individual wages, and these distributions are therefore of particular interest to us. Table 3 contains descriptive statistics on the distributions of these variables and figure 1 plots the associated nonparametric kernel density estimates.

Table 3 and Figure 1 reveal a considerable dispersion in both value added per worker, firm specific average wages, and individual wages in Denmark. Similar findings have been reported for the US, France, and Norway using MEE data (see Lentz and Mortensen (2009) and the reference therein).

The second empirical goal of the paper is to shed new light on the relationship between value added per worker and wages. Figure 2 plots nonparametric kernel regressions of firm-specific average

[5] The monthly norm is 166.33 hours, thus making the annual norm 1,995.96 hours.

	Manufacturing	Construction	Wholesale & Retail
Hourly value added/FTE:			
Mean	257.60	217.74	246.69
S.d.	93.09	56.63	106.11
Firm specific average wage:			
Mean	198.34	190.51	177.35
S.d.	112.24	110.95	136.60
Cor. w/ hrly VA/FTE	0.46	0.47	
Individual wage:			
Mean	214.93	207.07	184.72
S.d.	125.59	111.08	133.81
Cor. w/ hrly VA/FTE	0.14	0.17	

Table 3: Moments of the distributions of value added/FTE, firms specific average wages and individual wages. All moments denoted in 2006 DKK. Value added moments and firm specific wage moments are computed across firm-years. Individual wage moments are computed across individual observations. Correlations are computed using observations where both variables are non-missing.

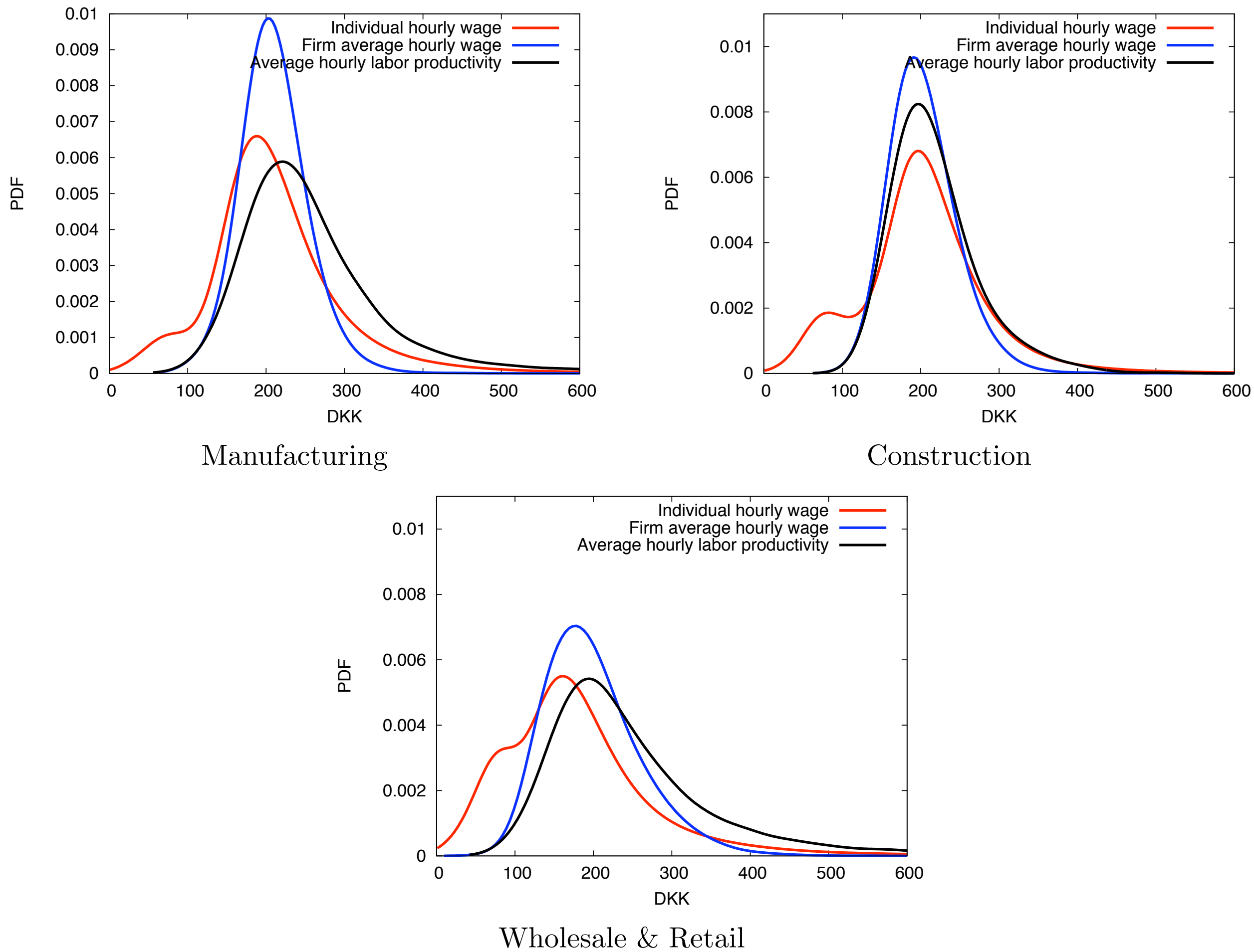


Figure 1: Empirical density of value added per FTE, firm specific average wages and individual wages. In each case densities are estimated using a kernel density estimator with Gaussian kernel and bandwidth 20 DKK.

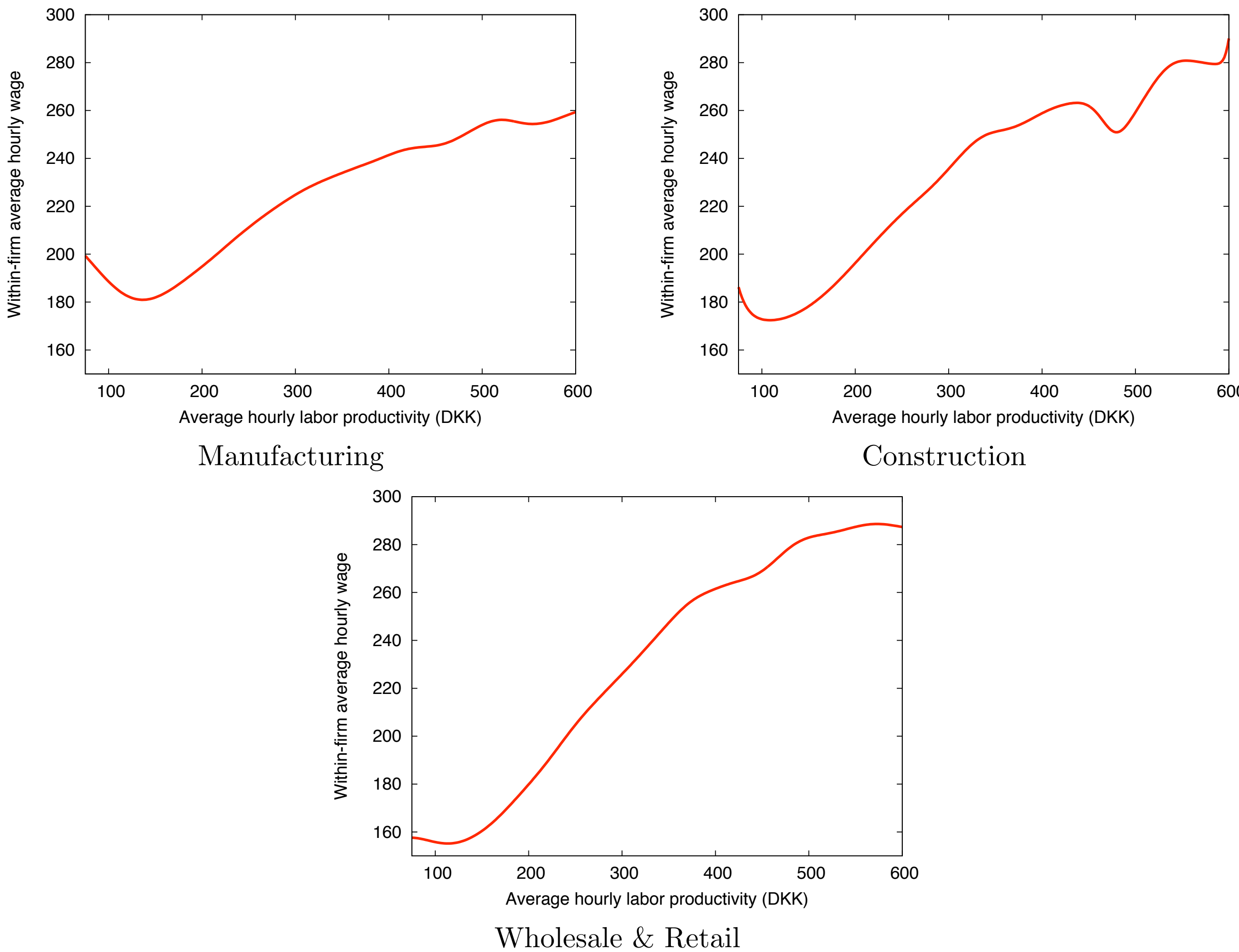


Figure 2: Nonparametric kernel regression of firm specific wage on value added per FTE. The regression is perfomed with a Gaussian kernel and bandwidth 20 DKK.

wages on value added per FTE.

Both Table 3, which reports the correlation coefficient between the two wage measures and hourly value added/FTE, and Figure 2 reveal a rather strong positive relationship between labor productivity and the average wage paid in the firm. Again, this empirical finding occurs in most other MEE datasets with both wage and productivity measures (see, e.g., Postel-Vinay and Robin, 2006, and Lentz and Mortensen, 2009).

4 Estimation

The structural model developed in section 2 gives rise to a system of equations characterizing the relationship between individual wages, firm wages and firm output (value added). Here, we describe how the system can be identified and estimated in the data described in section 3. In the following, $i \in \mathcal{I} = \{1, ..., I\}$ indexes individuals, $j \in \mathcal{J} = \{1, ..., J\}$ indexes firms, $h \in \mathcal{H} = \{1, ..., H\}$ indexes job categories, and $t \in \mathcal{T} = \{1, ..., T\}$ indexes time periods. Let $\mathrm{J} : \mathcal{I} \times \mathcal{T} \mapsto \mathcal{J}$ and $\mathrm{H} : \mathcal{I} \times \mathcal{T} \mapsto \mathcal{H}$ be the mappings indicating the employer respectively the job category of each worker in each period.

We parameterize individual ability in (8) as $a_{ih} = \exp(\psi_{ih} + \mathbf{x}'_{it}\gamma)$, where ψ_{ih} is a fixed effect specific to individual i in job category h, $\mathbf{x}_{it}$ is an M-vector of time- and job category varying individual specific characteristics observable to the econometrician, and γ is a conformable vector of parameters. Since in this case ability is time-varying, we may by slight abuse of notation write

$$\ln a_{it} = \psi_{i\mathrm{H}(i,t)} + \mathbf{x}'_{it}\gamma \tag{18}$$

for the ability of individual i in period t. Let p_{jt} denote firm j's productivity in period t, assumed to follow a first order Markov process. From (8), (13), and (10), the system forming the basis of our empirical analysis is

$$\begin{aligned}
\ln w_{it} &= \psi_{i\mathrm{H}(i,t)} + \mathbf{x}'_{it}\gamma + \ln w_{\mathrm{J}(i,t)\mathrm{H}(i,t)t} + \varepsilon^w_{it}, && (19)\\
\ln w_{jht} &= \ln\left((1-\beta)b_h + \frac{\beta\alpha_L}{1-\beta-\beta\alpha_L}\frac{L^\rho_{jht}}{\sum_{n\in\mathcal{H}} L^\rho_{jht}}\frac{Y_{jt}}{L_{jht}}\right) + \varepsilon^w_{jht}\ ,\ \ h \in \mathcal{H}, && (20)\\
\ln Y_{jt} &= \alpha_K \ln K_{jt} + \frac{\alpha_L}{\rho}\ln\left(\sum_{h\in\mathcal{H}} L^\rho_{jht}\right) + p_{jt} + \varepsilon^Y_{jt}, && (21)
\end{aligned}$$

where $(i,t) \in \mathcal{I} \times \mathcal{T}$ in the individual wage equation (19), whereas in the firm level subsystem consisting of the firm wage equation (20) and the production function (21) we have $(j,t) \in \mathcal{J} \times \mathcal{T}$. In (20) we have utilized the fact that from (13) the firm wage $w_{jht}(K_{jt}, \mathbf{L}_{jt}, p_{jt}) = w_h(\mathbf{L}_{jt}, Y_{jt})$ only depends on productivity through output. Moreover, our assumption on within-job category perfect substitutability of the supplied abilities implies that job category h labor input in the firm level subsystem (20)-(21) is $L_{jht} = \sum_{\{i\ :\ \mathrm{J}(i,t)=j,\ \mathrm{H}(i,t)=h\}} a_{it}$. We think of the error terms as capturing measurement errors and non-modeled stochastic components in individual wages, firm wages, and output. Assumption 1 states the assumed properties of ε^w_{it}, ε^w_{jht}, and ε^Y_{jt}.

Assumption 1 (Error terms) *The error terms* ε_{it}^{w}, ε_{jht}^{w} *and* ε_{jt}^{Y} *are independent of each other and of the right hand side variables in their respective equations, and each of the three is i.i.d..*

Regarding productivity p_{jt} we follow the assumptions in Ackerberg, Caves and Frazer (2006), who in turn rely on the structure introduced by Olley and Pakes (1996) and Levinsohn and Petrin (2003).[6]

Assumption 2 (Firm productivity) *Firm productivity* p_{jt} *follows a first order Markov process:* $p_{jt} = \mathrm{E}[p_{jt}|p_{jt-1}] + \nu_{jt}$, *where the productivity shocks* ν_{jt} *are i.i.d. innovations across firms and time.*

Firms decide on investments I_{jt}, hiring effort, and material inputs. We assume that investments are determined before hiring and material inputs, and that hiring is determined before material inputs. The law of motion for the capital stock is given as (see (15))

$$K_{jt+1} = I_{jt} + (1 - \delta_{jt})K_{jt}, \tag{22}$$

where I_{jt} is firm j's period t investment and δ_{jt} represents the depreciation rate (known to the firm when I_{jt} is determined). Let $\Xi_j(z)$ be firm j's information set when generic variable z is determined. Assumption 3 is crucial for our estimation procedure:

Assumption 3 (Investment behavior) *Period* t *investments* I_{jt} *are determined with full knowledge of* p_{jt-1}, *but prior to the revelation of* p_{jt}. *That is,* $p_{jt-1} \in \Xi_j(I_{jt})$ *and* $p_{jt} \notin \Xi_j(I_{jt})$.

Assumption 3 is satisfied if it takes one period to acquire and install desired capital. Our data are annual, and in the empirical analysis we take a period to be one year.

Now turn to labor input. As mentioned above, we assume that firms have decided on (and committed to) investments before they turn to the hiring decision. The labor market is frictional and firms are in general off their labor demand curve. As in Section 2.4, firms decide on hiring effort (vacancy posting) v_{jt}. Given hiring effort, labor force configuration $\mathbf{L}_{jt}$ then evolves according

[6] In estimating the system (19), (20) and (21) we shall make use of the estimation procedure developed in Ackerberg et al. (2006).

to some stochastic process that we leave unspecified. While this results in sluggish labor force adjustments, we do not think that it is appropriate to assume that firms decide on their labor force one year in advance. Hence, we impose the less restrictive Assumption 4.

Assumption 4 (Hiring behavior) *Period t hiring effort v_{jt} is determined with full knowledge of p_{jt}. That is, $p_{jt} \in \Xi_j(v_{jt})$.*

Finally, turn to firms' demand for material inputs, denoted M_{jt}. Material inputs do not enter directly in the system (19), (20) and (21) as the production function is specified in terms of value added. Nonetheless, the properties of material demand plays a crucial role in the estimation procedure. In terms of timing, material inputs M_{jt} are determined after investments I_{jt} have been decided upon and not only after hiring effort v_{jt} is set, but also after the period t labor force configuration $\mathbf{L}_{jt}$ has been fixed. We stress that the fact that the outcome of the hiring process is known when firms decide on material inputs will be crucial to our estimation procedure.

Assumption 5 (Material inputs) *Period t material inputs M_{jt} are determined with full knowledge of p_{jt}. That is, $p_{jt} \in \Xi_j(M_{jt})$. Moreover, M_{jt} is a perfectly flexible input with no dynamic effects. The firm level demand function for material inputs χ is strictly increasing in productivity p_{jt}: $M_{jt} = \chi(K_{jt}, \mathbf{L}_{jt}, p_{jt})$.*

By 'perfectly flexible' we mean that the market for material inputs is frictionless and that firms can acquire and start using material inputs immediately. By stating that M_{jt} has 'no dynamic effects' we mean that M_{jt} does not affect any of the firm's future choices.

With Assumptions 1-5 in place we can now proceed to a description of our estimation procedure. It falls in the class of Two-Step Generalized Methods of Moments estimators (see e.g. Newey and McFadden (1994) for results on consistency and asymptotic normality). Our first step makes use of equation (19) to separate worker-occupation effects from firm-occupation-time effects (i.e., separate a_{ih} from $w_h(\mathbf{L}_{jt}, Y_{jt})$). The second step then uses this decomposition to estimate the remainder of the structural parameters from the sub-system made up of equations (20) and (21). As our first step estimation produces consistent estimates if T is considered large, first step estimation errors does not affect second step consistency. To confirm the finite T properties of our method we implement a bootstrap procedure.

4.1 Step 1: Individual wages

We start by considering equation (19) describing individual wages. Note that $\ln w_{\mathrm{J}(i,t)\mathrm{H}(i,t)t}$in (19) at this point can be considered as a time-varying firm-job category effect. Writing this as $\phi_{\mathrm{J}(i,t)\mathrm{H}(i,t)t} = \ln w_{\mathrm{J}(i,t)\mathrm{H}(i,t)t}$, (19) can be expressed as

$$\ln w_{it} = \psi_{i\mathrm{H}(i,t)} + \mathbf{x}_{it}'\gamma + \phi_{\mathrm{J}(i,t)\mathrm{H}(i,t)t} + \varepsilon_{it}^{w}. \tag{23}$$

At this point it is worth considering the correlation structure among the components in (23). Our model is (purposefully) silent about the nature of the frictions that impede worker reallocation across firms and occupations (i.e., the mappings J and H), so the equilibrium we describe need not feature sorting. While sorting of workers across firms may or may not be empirically important (see Lise, Meghir and Robin (2009), de Melo (2009), and Bagger and Lentz (2009) for recent empirical work in this area), it seems likely that firms would tend to allocate good category-h workers to category h and/or that good category-h workers would self-select into category h. Job categories can be thought of as occupations. Groes, Kircher and Manovskii (2009) provide empirical evidence on sorting across occupations. In other words, $\mathrm{H} : \mathcal{I} \times \mathcal{T} \mapsto \mathcal{H}$ is likely to exhibit sorting.

Let $\ln \mathbf{w}$ be the IT-vector of stacked individual wages, $\mathbf{X}$ the $IT \times M$-matrix of stacked M-vectors $\mathbf{x}_{it}'$ of observable characteristics, and $\mathbf{e}$ the IT-vector of stacked individual errors ε_{it}^{w}. Then, upon defining $\mathbf{D}$ and $\mathbf{F}$ as the $IT \times IH$ and $IT \times JHT$ (sparse) design matrices containing the full set of interactions between worker and job category dummies and firm and job category and time dummies, we can restate the individual wage equation (19) in terms of data vectors and matrices as

$$\ln \mathbf{w} = \mathbf{X}\gamma + \mathbf{D}\psi + \mathbf{F}\phi + \mathbf{e}, \tag{24}$$

where ψ and ϕ are the IH- and JHT-vectors containing the worker-category and the firm-category-time effects, respectively. Identification of the structural parameters $(\gamma', \psi')'$ and the composite firm-occupation-time effects ϕ in (24) rely on Assumptions 1 and 6:

Assumption 6 (Invertibility) *The data matrix* $[\mathbf{X}:\mathbf{D}:\mathbf{F}]$ *has full column rank.*

Assumption 1 allows us to treat $\mathbf{e}$ as a statistical residual in a regression of individual log wages on individual-job category and firm-job category-time effects. Assumption 6 is a regularity assumption. Full column rank is easily imposed on the sub-matrix $\mathbf{X}$ by omitting collinear columns. With respect to $[\mathbf{D}:\mathbf{F}]$ we apply the identification and estimation strategy of Abowd, Creecy and Kramarz (2002). Indeed, our individual-job category and firm-job category-time effects can be treated as "worker effects" and "firm effects" in their terminology, implying that identification of the individual-occupation and the firm-occupation-time effects only is possible within connected groups of workers and firms. We select the group containing most data points for analysis and redefine $\ln\mathbf{w}$, $\mathbf{X}$, $\mathbf{D}$, and $\mathbf{F}$ to contain only the corresponding data. Under Assumptions 1 and 6 we can estimate (γ',ψ',ϕ') as

$$\begin{pmatrix}\widehat{\gamma}\\ \widehat{\psi}\\ \widehat{\phi}\end{pmatrix}=\begin{bmatrix}\mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{D} & \mathbf{X}'\mathbf{F}\\ \mathbf{D}'\mathbf{X} & \mathbf{D}'\mathbf{D} & \mathbf{D}'\mathbf{F}\\ \mathbf{F}'\mathbf{X} & \mathbf{F}'\mathbf{D} & \mathbf{F}'\mathbf{F}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{X}'\\ \mathbf{D}'\\ \mathbf{F}'\end{bmatrix}\ln\mathbf{w}. \tag{25}$$

Clearly, (25) is a GMM estimator, with each of the regressors acting as its own instrument.

At this point we have isolated a number of components of observed dispersion of individual wages, namely the effect of individual ability, $\ln\widehat{a}_{it}=\widehat{\psi}_{i\mathrm{H}(i,t)}+\mathbf{x}'_{it}\widehat{\gamma}$, the firm-occupation-time effect $\widehat{\phi}_{\mathrm{J}(i,t)\mathrm{H}(i,t)t}$, and the effect of 'noise,' $\widehat{\varepsilon}^{w}_{it}=\ln w_{it}-\widehat{\psi}_{i\mathrm{H}(i,t)}-\mathbf{x}'_{it}\widehat{\gamma}-\widehat{\phi}_{\mathrm{J}(i,t)\mathrm{H}(i,t)t}$. We now turn to a further decomposition of the firm-occupation-time effects $\widehat{\phi}$.

4.2 Step 2: Firm wage and productivity dispersion

The unit of observation in our firm level panel is a firm-occupation-year. Let firms and occupations be indexed directly by $j=1,...,J$ and $h=1,...,H$. We have $\ln w_{jht}=\widehat{\phi}_{jht}$ and consider the firm-level nonlinear system of equations given by (20)-(21),

$$\widehat{\phi}_{jht} = \ln\left((1-\beta)b_h+\frac{\beta\alpha_L}{1-\beta-\beta\alpha_L}\frac{\widehat{L}^{\rho}_{jht}}{\sum_{n\in\mathcal{H}}\widehat{L}^{\rho}_{jnt}}\frac{Y_{jt}}{\widehat{L}_{jht}}\right)+\varepsilon^{w}_{jht},\ h\in\mathcal{H} \tag{26}$$

$$\ln Y_{jt} = \alpha_K\ln K_{jt}+\frac{\alpha_L}{\rho}\ln\left(\sum_{h\in\mathcal{H}}\widehat{L}^{\rho}_{jht}\right)+p_{jt}+\varepsilon^{Y}_{jt} \tag{27}$$

where $\widehat{L}_{jht}=\sum_{\{i:\ \mathrm{J}(i,t)=j,\ \mathrm{H}(i,t)=h\}}\widehat{a}_{it}$. For fixed T, there is measurement error in $\widehat{\phi}_{jht}$, and this is accommodated by the error term ε^{w}_{jht} in (26). If T is considered large, then $\widehat{\phi}_{jht}$ and $\widehat{a}_{it}$ are consistent estimates of $\ln w_{jht}$ and a_{it}, and consistency of the second step estimator is not affected by replacing $\ln w_{jht}$ and a_{it} by the first step estimates $\widehat{\phi}_{jht}$ and $\widehat{a}_{it}$. In other words, we can ignore first step estimation errors and proceed as if $\ln w_{jht}$ and a_{it} were observed.

We develop a GMM estimator (Hansen, 1982) to estimate the structural parameters in equations (26) and (27). In so doing we recognize that production function inputs (here, capital and labor) are endogenous in relation to firm productivity ω_{jt} (see Grilliches and Mairesse (1995) for a short review of this classical problem). The empirical production function literature has identified two ways around this problem. One strand of the literature relies on instrumental variables (Blundell and Bond, 2000). Another strand uses explicit structural models in a control function approach. This approach was first developed in Olley and Pakes (1996) and extended in Levinsohn and Petrin (2003) and again in Ackerberg, Caves and Frazer (2006). The Olley and Pakes approach requires labor to be non-dynamic and completely flexible. Moreover, to construct an appropriate control function, firms' investment functions must be strictly monotonic functions of the capital stock. This makes lumpy investment behavior a problem. Levinsohn and Petrin (2003) augment the Olley and Pakes estimator by using intermediary material inputs to construct the control function, thus circumventing the problem of lumpy investments. Moreover, labor input is allowed to have dynamic effects, but must be perfectly flexible. Ackerberg et al. (2006), building on Levinsohn's and Petrin's approach, allow dynamic and nonflexible labor inputs. We apply the Ackerberg et al. (2006) estimator in this paper.

4.2.1 Estimation procedure

When β is assumed constant across workers and job categories, the firm level system involves the $H+4$ unknown parameters $(\alpha_K, \alpha_L, \rho, \beta, b_1, ..., b_H)$ to be estimated.

The firm-wage equations: By Assumption 1, ε^w_{jht} is uncorrelated with Y_{jt} and $\widehat{L}_{jht}$ for all $h \in \mathcal{H}$. This yields $2 \times H$ orthogonality conditions that we can exploit in the estimation: $\mathrm{E}[Y_{jt}\varepsilon^w_{jht}] = 0$ and $\mathrm{E}[L_{jht}\varepsilon^w_{jht}] = 0$ for all $h \in \mathcal{H}$, where the expectation is take over firms and time. These moment conditions are thus operationalized as

$$\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}} Y_{jt}\left[\widehat{\phi}_{jht} - \ln\left((1-\beta)b_h + \frac{\beta\alpha_L}{1-\beta-\beta\alpha_L}\frac{\widehat{L}^{\rho}_{jht}}{\sum_{n\in\mathcal{H}}\widehat{L}^{\rho}_{jnt}}\frac{Y_{jt}}{\widehat{L}_{jht}}\right)\right] = 0 \ , \ h\in\mathcal{H} \quad (28)$$

$$\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}} \widehat{L}_{jht}\left[\widehat{\phi}_{jht} - \ln\left((1-\beta)b_h + \frac{\beta\alpha_L}{1-\beta-\beta\alpha_L}\frac{\widehat{L}^{\rho}_{jht}}{\sum_{n\in\mathcal{H}}\widehat{L}^{\rho}_{jnt}}\frac{Y_{jt}}{\widehat{L}_{jht}}\right)\right] = 0 \ , \ h\in\mathcal{H} \quad (29)$$

Additionally, we may utilize moments based on cross-occupation correlations.

The production function: Extracting moments from the production function is a bit more involved because K_{jt} and $\mathbf{L}'_{jt}$ are likely to endogenous in relation to unobserved firm productivity p_{jt}. We circumvent this problem by implementing the production function estimator developed by Ackerberg et al (2006). Indeed, as a consequence of Assumption 5 we can 'back out' firm productivity p_{jt} from the demand function for material inputs,

$$p_{jt} = \chi^{-1}(M_{jt}, K_{jt}, \widehat{\mathbf{L}}'_{jt}), \quad (30)$$

such that the production function can be written as

$$\ln Y_{jt} = \alpha_K \ln K_{jt} + \frac{\alpha_L}{\rho}\ln\left(\sum_{h\in\mathcal{H}}\widehat{L}^{\rho}_{jht}\right) + \chi^{-1}(M_{jt}, K_{jt}, \widehat{\mathbf{L}}'_{jt}) + \varepsilon^Y_{jt}. \quad (31)$$

Of course, this equation is not directly useful for estimation as $\chi(\cdot)$ is not known. However $\chi^{-1}(\cdot)$ can be estimated as a nonparametric function of the inputs $(M_{jt}, K_{jt}, \widehat{\mathbf{L}}'_{jt})$. It is not possible to separately identify a fully nonparametric function $(M_{jt}, K_{jt}, \widehat{\mathbf{L}}'_{jt})$ and the "direct" effects of K_{jt} and $\widehat{\mathbf{L}}'_{jt}$ in (31). Hence, in terms of estimation, at this point we can only separate the structural part of the production function from the idiosyncratic part (which we interpret as measurement error):

$$\ln Y_{jt} = f_t(M_{jt}, K_{jt}, \widehat{\mathbf{L}}'_{jt}) + \varepsilon^Y_{jt}. \quad (32)$$

From equation (32) we can nonparametrically estimate $f_t(M_{jt}, K_{jt}, \widehat{\mathbf{L}}'_{jt})$.[7] Denote the estimated function $\widehat{f}_t$ and the fitted values $\widehat{f}_{jt} = \widehat{f}_t(M_{jt}, K_{jt}, \widehat{\mathbf{L}}'_{jt})$. Then

$$\widehat{p}_{jt} = \widehat{f}_{jt} - \alpha_K \ln K_{jt} + \frac{\alpha_L}{\rho}\ln\left(\sum_{h\in\mathcal{H}}\widehat{L}^{\rho}_{jht}\right). \quad (33)$$

[7] We may reduce the dimensionality of the regression by conditioning on the structural parameter ρ, such that $\ln Y_{jt} = f_t(M_{jt}, K_{jt}, \widehat{L}_{jt}) + \varepsilon^Y_{jt}$, where $\widehat{L}_{jt} = \sum_{h\in\mathcal{H}}\widehat{L}^{\rho}_{jht}$. In this case, the nonparametric regression must be updated whenever ρ is updated in the structural estimation procedure.

Now, by Assumption 2, p_{jt} is a first order Markov process: $p_{jt} = \mathrm{E}[p_{jt}|p_{jt-1}] + \nu_{jt}$, where ν_{jt} is an i.i.d. innovation. Given assumptions 3 and 4 this provides us with a set of $H+1$ moments that can be exploited in the estimation, namely $\mathrm{E}[\nu_{jt}K_{jt}] = 0$ and $E[\nu_{jt}L_{jht-1}] = 0$, for all $h \in \mathcal{H}$.

What remains is to construct expressions for ν_{jt} in terms of observable quantities. We again exploit Assumption 2. Indeed, for a given set of structural parameters $(\alpha_K, \alpha_L, \rho)$ we can construct a time series $\widehat{p}_{jt}(\alpha_K, \alpha_L, \rho)$ for each firm j. Then, we can construct a panel of $\nu_{jt} = \nu_{jt}(\alpha_K, \alpha_L, \rho)$ as the residuals from the nonparametric regression across firms and time

$$\widehat{p}_{jt}(\alpha_K, \alpha_L, \rho) = g(\widehat{p}_{jt-1}(\alpha_K, \alpha_L, \rho)) + \nu_{jt}, \tag{34}$$

that is, $\widehat{\nu}_{jt}(\alpha_K, \alpha_L, \rho) = \widehat{p}_{jt}(\alpha_K, \alpha_L, \rho) - \widehat{g}_{jt}$. Now the production function moments conditions have been operationalized as

$$\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}}(\widehat{p}_{jt}(\alpha_K, \alpha_L, \rho) - \widehat{g}_{jt})K_{jt} = 0 \tag{35}$$

$$\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}}(\widehat{p}_{jt}(\alpha_K, \alpha_L, \rho) - \widehat{g}_{jt})\widehat{L}_{jt-1} = 0\ ,\ h \in \mathcal{H} \tag{36}$$

The two sets of moments (from the firm wage equation and the production function) totals $3H+1$ moments. The system contains $H+4$ parameters (when $H > 1$). Hence, the system is in fact overidentified. Our estimation procedure falls within the framework of a 2-step GMM estimators (Newey and McFadden, 1994).

Once the structural parameters are estimated we back out p_{jt}. Hence, we have estimates for all components of the system (19), (20) and (21), including the distributions of the idiosyncratic terms ε^w_{it}, ε^w_{jht}, ε^Y_{jt}, and ν_{jt}.

4.3 Empirical results

The estimation method described above is implemented in each of the three industry samples. We first focus on the Manufacturing sector and estimate the specification with $H = 1$ occupation. Table 8 in the Appendix B provides descriptive statistics on the connectivity of our samples.

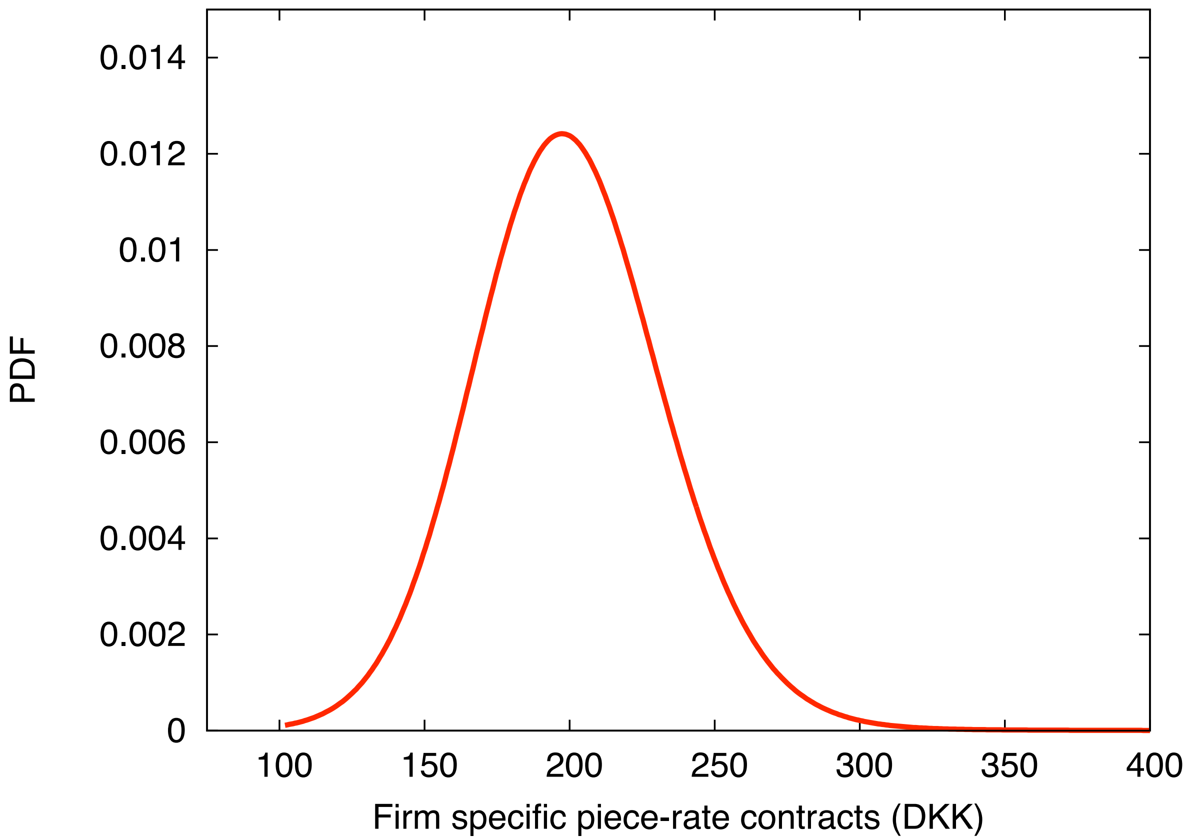
The individual wage equation is estimated in the first step. Figure 3 shows the density (across firms and time) of estimated hourly firm-time effects (hourly firm specific piece-rate wages), es-

	Manufacturing
Avg. piece-rate wage w_{jt}	219.22
S.d. piece-rate wage	32.17
Avg. quality adj. labor prod. Y_{jt}/L_{jt}	528.85
S.d. quality adj. labor prod.	192.18
Corr$(w_{jt}, Y_{jt}/L_{jt})$	0.36

Table 4: Moments and correlations of estimated hourly piece-rate wages and quality adjusted annual labor productivity (in 1000 DKK)

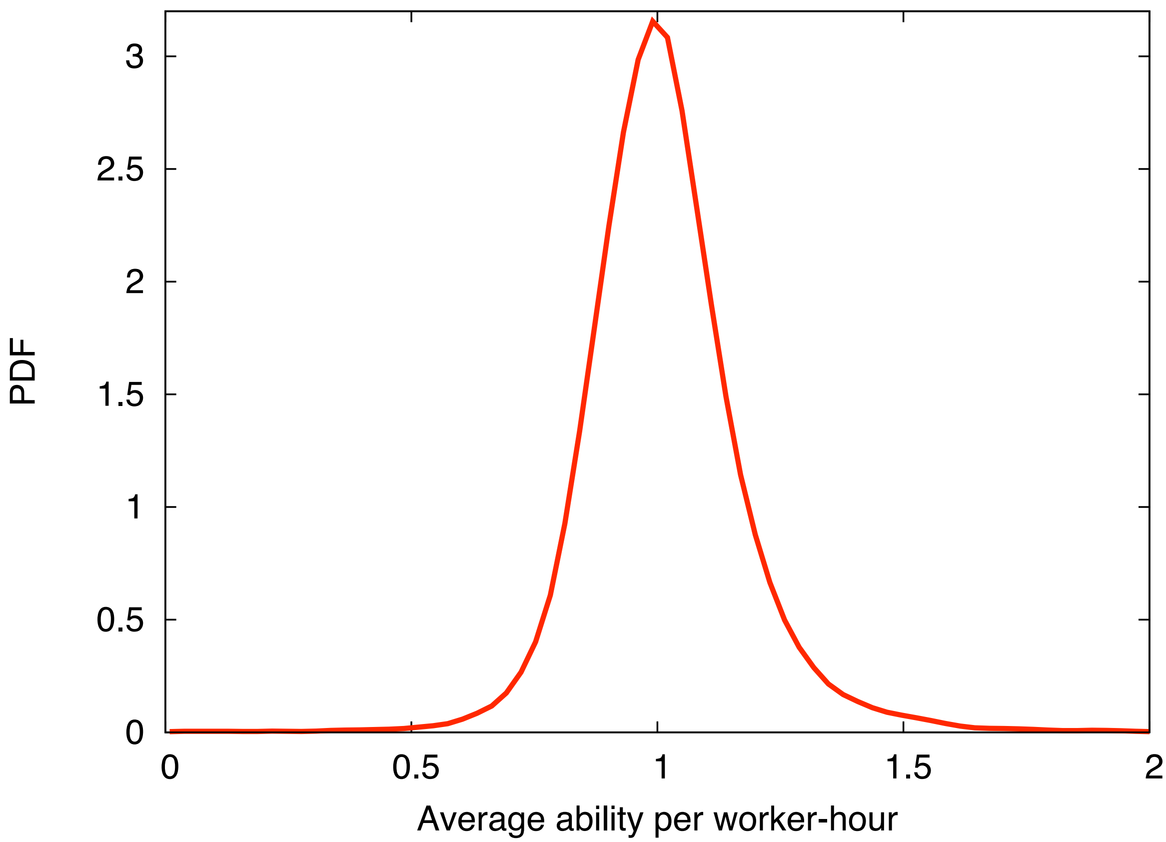
timated using a Gaussian kernel with bandwidth 20. Most of the mass is concentrated on the interval from DKK 150 to DKK 250. Figure 4 shows the similarly estimated density (across firms and time) of firm level average ability. The distribution of ability is normalized such that total ability supplied across firms in a given year equals the total number of man-hours used in production that year. That is, average ability per worker is normalized to unity. As it turns out, the within-firm average ability is also close to unity. This outcome would materialize if there were no sorting of workers across firms. Finally, Table 4 shows that the the piece rate wage is positively correlated with quality adjusted labor productivity, and that the correlation is relatively strong. In the manufacturing industry the correlation coefficient is .36. Hence, the positive correlations reported in the descriptive analysis in section 3 survive our labor quality adjustment.

The firm level system is estimated in the second step. Here, we are interested in the structural parameters $(\alpha_K, \alpha_L, \beta, b)$, as well as the productivity process for p_{jt} and the distribution of error terms, which we interpret as measurement errors. Estimation results for the Manufacturing industry appear in Table 5. We present our (preferred) estimates based on the Ackerberg et al. (2006) approach together with "naive" least squares and within-firm estimates. The least squares estimates imposes the implausible assumption that hiring and investment decisions are uncorrelated with TFP shocks. The within-firm estimates are based on the assumption that TFP is a firm specific fixed variable (i.e. $p_{jt} = p_j$). As described further above, the Ackerberg et al. (2006) estimates, on which we base our analysis and conclusions, allow investments and hiring decisions of firms to be



Manufacturing

Figure 3: Nonparametric kernel density estimates of the distributions of estimated firm-time specific wage effects (i.e. firm-time specific piece-rate wages, manufacturing sector only). The reported densities are computed using all firm-years. The densities are calculated from a Gaussian kernel and bandwidth 20



Manufacturing

Figure 4: Nonparametric kernel density estimates of the distributions of estimated average worker ability across firm-years (manufacturing sector only). The density is estimated using a (Gaussian) kernel density estimator with bandwidth 0.03. Labor input have been normalized such that average labor input in each annual cross-section is 1.

	LS	Within	ACF
Output elasticity w.r.t. K, α_K	0.1146 (0.0018)	0.1072 (0.0033)	0.3263
Output elasticity w.r.t. L, α_L	0.8854 (0.0057)	0.8928 (0.0103)	0.6896
Workers' barganing power, β			0.4154
Workers' outside option, b			205.5841

Table 5: Estimated structural parameters for firm-level wage and production function system (manufacturing sector only). At present, we do not have standard errors to report for the ACF based estimates.

dependent on stochastic TFP shocks. Relative to our preferred estimates, the least squares and within firm estimates of the output elasticity of capital is significantly downward biased, although the return to scale seems to be unaffected by the estimation method.

Focusing now on the production function estimates based on the Ackerberg et al. (2006) approach, we estimate the output elasticity of capital, α_K, to .326. This is estimate is large in magnitude, and, as already noted, larger than in other specifications not controlling for unobserved heterogeneity. The output elasticity with respect to labor input, α_L, is estimated to .690, so the results seem consistent with constant returns to scale in these two inputs. Workers' bargaining power, β, is estimated to .415, and the outside option, i.e., the value of home production per unit of ability, is estimated to 205.6, slightly below the average piece-rate wage of 219.2 from Table 5.

4.4 Variance decompositions and fit

Given the estimated structural parameters, the model naturally admits a decomposition of log ouput per worker and the individual log wages. Moreover, we may assess how well our rent sharing mechanism fits the observed positive relationship between labor productivity and within firm average wages by comparing model predictions to the relationship observed in our data.

4.4.1 Log labor productivity variance decomposition

The production function (21) may in the case $H = 1$ be written in terms of observed output per worker Y_{jt}/N_{jt} as

$$\ln\left(\frac{Y_{jt}}{N_{jt}}\right) = \alpha_K \ln\left(\frac{K_{jt}}{N_{jt}}\right) + \alpha_L \ln\left(\frac{L_{jt}}{N_{jt}}\right) + (\alpha_L + \alpha_K - 1)\ln N_{jt} + p_{jt} + \varepsilon^Y_{jt}, \tag{37}$$

and it is natural to decompose the variance of $\ln(Y_{jt}/N_{jt})$ in the following way:

$$\begin{aligned} \operatorname{Var}\left[\ln\left(\frac{Y_{jt}}{N_{jt}}\right)\right] &= \alpha_K \operatorname{Cov}\left[\ln\left(\frac{K_{jt}}{N_{jt}}\right), \ln\left(\frac{Y_{jt}}{N_{jt}}\right)\right] \\ &+ \alpha_L \operatorname{Cov}\left[\ln\left(\frac{L_{jt}}{N_{jt}}\right), \ln\left(\frac{Y_{jt}}{N_{jt}}\right)\right] \\ &+ (\alpha_L + \alpha_K - 1)\operatorname{Cov}\left[\ln N_{jt}, \ln\frac{Y_{jt}}{N_{jt}}\right] \\ &+ \operatorname{Cov}\left[p_{jt}, \ln\frac{Y_{jt}}{N_{jt}}\right] \\ &+ \operatorname{Cov}\left[\varepsilon^Y_{jt}, \ln\frac{Y_{jt}}{N_{jt}}\right]. \end{aligned} \tag{38}$$

In this setup, we may measure, say, the fraction of variation in observed output per worker $\ln(Y_{jt}/N_{jt})$ that can be accounted for by its covariation with the capital labor ratio $\ln(K_{jt}/N_{jt})$ as $\alpha_K \operatorname{Cov}[\ln(K_{jt}/N_{jt}), \ln(Y_{jt}/N_{jt})]/\operatorname{Var}[\ln(Y_{jt}/N_{jt})]$. The first two terms and the last two terms on the right hand side of (38) are easily interpreted, as they reflect the contributions of the dispersions in capital-labor ratio, average ability, productivity, and measurement error to observed output dispersion. The third term is a scale effect in the labor productivity distribution that vanishes if the production function exhibits constant returns to scale.

Table 6 reports the resulting variance decompositions for the Manufacturing industry. From the table, 41% of the variation in output per worker is attributable to cross-firm differences in the levels of capital per worker. The second largest contributor is differences in total factor productivity, accounting for 39% of the variation. A mere 4.7% of the variation stems from quality differences in the labor input. The portion left unexplained by our model is relatively low, at 15%.

That capital intensity and TFP are the main sources of output dispersion suggest that substantial gains in aggregate productivity can be achieved by reallocation workers from less capital intensive and low TFP firms to capital intensive and high TFP firms. If labor input quality had

$\text{Var}\left[\ln \frac{Y_{jt}}{N_{jt}}\right]$	0.09685
$\text{Cov}\left[\alpha_K \ln \frac{K_{jt}}{N_{jt}}, \ln \frac{Y_{jt}}{N_{jt}}\right]$	0.03967 (0.40977)
$\text{Cov}\left[\alpha_L \ln \frac{L_{jt}}{N_{jt}}, \ln \frac{Y_{jt}}{N_{jt}}\right]$	0.00454 (0.04691)
$\text{Cov}\left[(\alpha_L + \alpha_K - 1) \ln N_{jt}, \ln \frac{Y_{jt}}{N_{jt}}\right]$	0.00079 (0.00818)
$\text{Cov}\left[p_{jt}, \ln \frac{Y_{jt}}{N_{jt}}\right]$	0.03748 (0.38701)
$\text{Cov}\left[\varepsilon^Y_{jt}, \ln \frac{Y_{jt}}{N_{jt}}\right]$	0.01436 (0.14824)

Table 6: Decomposition of total log output per worker variation explained. The numbers in parentheses are fractions of log output per worker.

been the main driving force behind productivity dispersion, the gains from reallocation would be minor. Taken at face value, the production dispersion decomposition presented in Table 6 suggests that policies that stimulate labor mobility would be much more effective than policies stimulating, say, human capital investments by individual workers in generating aggregate productivity growth.

A central driver of labor mobility is workers' aspiration for higher wages. Hence, rent sharing between workers and firms may be an important mechanism in the reallocation process of workers from less to more productive firms. We now turn to an evaluation of the relative importance of rent sharing in generating wage dispersion at the level of the individual worker.

4.4.2 Decomposition of individual log wage variation explained

Like the production function, our individual level wage equation admits a natural log wage variance decomposition. Recall that individual wages w_{it} are given as

$$\ln w_{it} = \ln a_{it} + \ln w(K_{jt}, L_{jt}, p_{jt}) + \varepsilon^w_{it}. \tag{39}$$

That is, individual log wages consist of three components: A person specific effect, a firm-time specific effect stemming from rent sharing between worker and employer, and measurement errors. We may thus decompose individual log wage variance in the following way:

$$\text{Var}[\ln w_{it}] = \text{Cov}[\ln a_{it}, \ln w_{it}] + \text{Cov}[\ln w(K_{jt}, L_{jt}, p_{jt}), \ln w_{it}] + \text{Cov}[\varepsilon^w_{it}, \ln w_{it}]. \tag{40}$$

	Manufacturing
$\mathrm{Var}[\ln w_{it}]$	0.2192
$\mathrm{Cov}[\ln a_{it}, \ln w_{it}]$	0.07994 (0.68730)
$\mathrm{Cov}[\ln w(K_{jt}, L_{jt}, p_{jt}), \ln w_{it}]$	0.01583 (0.13611)
$\mathrm{Cov}[\varepsilon^W_{jt}, \ln w_{it}]$	0.02054 (0.17659)

Table 7: individual log wage variance decomposition. The numbers in parentheses are fractions of total individual log wage variation.

The decomposition (40) is easily obtained given our first step estimates and is reported in Table 7.

As was evident from Figures 3 and 4, by far the most important source of individual wage dispersion is ability heterogeneity across workers. Within the Manufacturing industry, 68% of the observed dispersion in individual log wages can be accounted for by dispersion in worker abilities. In comparison, variation in piece-rate wages across firms within Manufacturing accounts for only 14% of observed dispersion in individual log wages. The remainder, i.e. around 18% is accounted for by measurement errors. These findings are in line with other log wage decompositions reported for different countries and periods (see e.g. Abowd, Kramarz and Margolis, 1999 and Abowd, Creecy and Kramarz, 2002). Notice however that we interpret the firm effects as firm specific piece-rate wages and that we allow these firm effects to vary over time. The temporal variation of the firm effects reflects the stochastic evolution of firms' TFP, labor force and capital stock. The corresponding decomposition imposing that the firm effects are constant over time yields a very similar pattern: 68% of the variation in individual level log wages is accounted for by worker heterogeneity, whereas firm heterogeneity explains 12%. The remaining 20% is residual wage dispersion.

5 Conclusion

Our model allows for differences in input quality as well as rent sharing, both of which may potentially help explain the observations that wage and value added per worker are dispersed across firms, and that more productive firms tend to pay higher wages. Estimating the model using detailed matched employer-employee data from the manufacturing sector in Denmark, we find that both input heterogeneity and intrinsic differences in total factor productivity across firms are important for dispersion. Our wage determination mechanism fits the observed relation between wage and productivity quite well. In quantitative terms, we find that 41% of the dispersion in log output per worker within the manufactoring sector is attributable to cross-firm differences in the levels of capital per worker, while another 39% of the variation stems from intrinsic TFP differences across firms. Only a smaller portion of log output per worker (5%) is associated with quality differences in the labor input.

Our results suggest that there are major gains to reallocation of labor from less to more productive firms. The same would not be the case if the dispersion in productivity were due to differences in input quality. Rent sharing can enhance the reallocation process by inducing the wage dispersion needed to motivate worker search. However, we find that firm differences (arising through rent sharing) accounts for only 13% of individual level wage dispersion in the manufacturing sector in Denmark. Instead individual level wage dispersion appears to be driven mainly by heterogeneity in workers' abilities. We note that the relatively small contribution of firm heterogeneity to individual wage dispersion is thus consistent with an inefficient allocation of labor across firms.

It is the detailed matched employer-employee data that allow us to draw our conclusions. In addition, in analyzing these data, we have contributed to the estimation of production functions, by embedding them in a structural equation system involving worker, firm, time, and occupation effects from an individual wage decomposition, while accommodating stochastically varying factor productivity.

References

[1] Abowd, J., D. Margolis, and F. Kramarz (1999). "High Wage Workers and High Wage Firms", *Econometrica*, Vol. 67, 251-333.

[2] Abowd, J., R. Creecy, and F. Kramarz (1999). "Computing Person and Firm Effects Using

Linked Longitudinal Employer-Employee Data", Cornell University working paper.

[3] Baily, M.N., C. Hulten, and D. Campbell (1992). "Productivity Dynamics in Manufacturing Plants," *Brooking Papers on Economic Activity: Microeconomics*, 1992: 187-249.

[4] Bernard A.B., and J.B. Jensen (1995). "Exporters, Jobs, and Wages in U.S. Manufacturing: 1976-1987," *Brookings Papers on Economic Activity: Microeconomics,* 1995: 67-119.

[5] Bernard A.B., and J.B. Jensen (1999a). "Exceptional Exporter Performance: Cause, Effect, or Both?", *Journal of International Economics,* vol. 47: 1-25.

[6] Bernard A.B. and J.B. Jensen (1999b). "Exporting and Productivity," National Bureau of Economic Research Working Paper No. 7135.

[7] Bernard A.B., J. Eaton, J.B. Jensen, S. Kortum (2003). "Plants and Productivity in International Trade," *The American Economic Review*, vol. 93: 1268-1290.

[8] Blundell R. and S. Bond (2000). "GMM estimation with persistent panel data: an application to production functions", *Econometric Reviews*, vol. 19: 321-340 .

[9] Burdett, K, and D.T. Mortensen (1998) "Wage Differentials, Employer Size and Unemployment," *International Economic Review* 39: 257-273.

[10] Cahuc, P., F. Marque, and E. Wasmer (2007). "A Theory of Wages and Labor Demand with Intrafirm Bargaining and Matching Frictions," forthcoming in the *International Economic Review.*

[11] Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006), "Wage Bargaining with On-the-job Search: Theory and Evidence," *Econometrica*, vol. 74, 323-364.

[12] Foster L., J. Haltiwanger, and C.J. Krizan (2001). "Aggregate Productivity Growth: Lessons from Microeconomic Evidence," In *New Directions in Productivity Analysis,* Dean E, M. Harper, and C. Hulten, eds., University of Chicago Press.

[13] Fox, J.T., and V. Smeets (2009), "Does Input Quality Drive Measured Differences in Firm," University of Chicago working paper.

[14] Griliches Z. and J. Mairesse (1995), "Production Functions: The Search for Identification", NBER working paper No. 5067.

[15] Groes, F. N., P. Kircher and I. Manovski (2009), "The U-Shapes of Occupational Mobility", University of Pennsylvania working paper

[16] Hall, R. and P.R. Milgrom (2008): "The Limited Influence of Unemployment on the Wage Bargain," *American Economic Review,* 98:4, pp. 1653-1674.

[17] Hansen, L. P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica*, vol. 50, 1029-1054

[18] Irarrazabal, A., A. Moxnes, and K.H. Ulltveit-Moe (2009). "The Black Box of Productivity and the Exporter Premium," University of Oslo working paper.

[19] Jorgenson, D.W., F.M. Gollop, and B. Fraumeni (1987). *Productivity and US Economic Growth*, Harvard University Press.

[20] Lentz R. and D. T. Mortensen (forthcoming) "Labor Market Models of Worker and Firm Heterogeneity", *Annual Review of Economics*

[21] Levinsohn, J. and A. Petrin (2003), "Estimating Production Functions Using Inputs to Control for Unobservables", *Review of Economic Studies*, vol. 70, 317-342.

[22] Mortensen, D.T. (2009) "Wage Dispersion in the Search and Matching Model with Intra-Firm Bargaining," NBER Working Paper #....

[23] Mortensen, D. (2003), *Wage Dispersion: Why Are Similar Workers Paid Differently?*, MIT Press.

[24] Murphy, K. M. and R. H. Topel (1990): “Efficiency Wages Reconsidered: Theory and Evidence” in Y. Weiss and G. Fishelson, eds *Advances in the Theory and Measurement of Unemployment*, pp. 204-40, Macmillan.

[25] Newey W. and D. McFadden (1994), ”Large Sample Estimation and Hypothesis Testing”, in R. Engle and D. McFadden, (eds.) Handbook of Econometrics, vol. IV, 2111-2245, North Holland: Amsterdam

[26] Olley, S. and A. Pakes (1996), ”The Dynamics of Productivity in the Telecommunications Equip- ment Industry”, *Econometrica*, vol. 64, 1263-1295

[27] Pedersen, N.K. (2009). *Essays in International Trade*, PHD thesis, Northwestern University.

[28] Postel-Vinay, F. and J.-M. Robin (2006), “Microeconometric Search-Matching Models and Matched Employer-Employee Data,” in The Proceedings of the 9th World Congress of the Econometric Society, edited by R. Blundell, W. Newey, and T. Persson, Cambridge University Press.

[29] Postel-Vinay, F., J.-M. Robin (2002a), ”The Distribution of Earnings in an Equilibrium Search Model with State-Dependent Offers and Counter-Offers,” *International Economic Review,* vol. 43, 989-1016.

[30] Postel-Vinay, F., J.-M. Robin (2002b), ”Equilibrium Wage Dispersion with Heterogeneous Workers and Firms,” *Econometrica*, vol. 70, 1295-1350.

[31] Iranzo, S., Schivardi, F., and Tosetti, E. (2008). ”Skill Dispersion and Firm Productivity: An Analysis with Employer-Employee Matched Data,” *Journal of Labor Economics*, 26: 247-284.

APPENDIX

A Value added and capital stock computations

This appendix details the computations of value added, capital stock, wage bill, and firm size measures from the accounting data files (REGNSK). REGNSK is a series of annual surveys conducted by Statistics Denmark during the period 1995-2006. REGNSK is to provide an accurate description of the Danish business community and serves as an important input in the construction of the national accounts. Unfortunately, the surveys change slightly in terms of variable definitions over the sample period, and the definition of value added must be changed accordingly. We provide evidence that these changes not induce major breaks in the data suggesting that our value added definition is time consistent.

A.1 Value added

Value added is revenue (turnover) minus cost of intermediate (material) inputs. The changes in definition over the sample period affect the computation of both revenue and cost of material inputs. Below follows a list of the value added definitions we apply and a short description of each of the variables entering the definitions. We use Statistics Denmark's official definitions. All variable definitions can be also be found under `http://www.dst.dk/times` (in Danish).

1995-1998: For the initial four years value added Y is defined as follows:

$$Y = (OMS+AUER+ADR+DLG)-(KRH+KENE+KLOE+UDHL+UASI+OEEU+SEUD), \quad (1)$$

where OMS is revenue, $AUER$ is work conducted at own expense and recorded under assets, ADR is other operating revenue, and DLG is ultimo inventory minus primo inventory. On the cost side KRH is cost of intermediates, $KENE$ is cost of energy, $KLOE$ is costs of subcontractors, $UDHL$ is housing rent, $UASI$ is purchase of minor equipment, $OEEU$ is other external costs and $SEUD$ are secondary costs.

1999-2001: For the next three years, 1999-2001 we define Y as follows:

$$\begin{aligned} Y = &(OMS + AUER + ADR + DLG + 0.0079 \times TGT) \\ &- (KRH + KENE + KLOE + UDHL + UASI + UDVB + ULOL + ANEU + SEUD), \quad (2) \end{aligned}$$

where TGT is total credits, $UDVB$ is purchase of temporary employment agency services, $ULOL$ is costs of long-term leasing, and $ANEU$ is other external costs (net of secondary costs).

2002-2003: In this period the definition of value added changes to:

$$\begin{aligned} Y = {} & (OMS + AUER + ADR + DLG) \\ & - (KRH + KENE + KLOE + UDHL + UASI + UDVB + ULOL + ANEU + SEUD), \end{aligned} \tag{3}$$

such that credits again drop out of the value added definition.

2004-2006: From 2004 we apply the following value added definition:

$$\begin{aligned} Y = {} & (OMS + AUER + ADR + DLG) \\ & - (KVV + KRHE + KENE + KLOE + UDHL \\ & + UASI + UDVB + ULOL + ANEU + SEUD), \end{aligned} \tag{4}$$

where KVV is purchase of goods for resale, and $KRHE$ is cost of intermediates.

A.2 Capital stock

The accounting data allow us to measure a firm's capital stock through the book value of material assets. Alternatively, we can construct a capital stock measure from the firm's reported investments in material assets (the perpetual inventory approach).

Book value approach As for value added, we follow Statistics Denmark's official definition of the book value of material fixed assets, denoted K^{BVA}. Unlike the value added definition this does not change over the sample period. It is given by

$$K^{BVA} = AADI + GRBY + ATAM + FMAA, \tag{5}$$

where $AADI$ is operating equipment and other equipment and facilities, $GRBY$ is buildings and sites, $ATAM$ is technical equipment and machinery, and $FMAA$ is pre-paid material fixed assets and material fixed assets under construction.

Investments and the perpetual inventory approach Computing the capital stock using the perpetual inventory approach requires that we compute firm-specific investments. We compute a firm's total annual investment, denoted I, as the sum of it's investments in real properties and production equipment:

$$I = FET + TDRT, \tag{6}$$

where FET is investment in real property and $TDET$ is investment in production equipment.

Then, introducing subscript t to denote time and given an initial condition which we take to be K_0^{BVA}, we may deduce a firm's capital stock in periods $t = 1, ..., T$ from it's investments by iterating on the identity $K_t^{PIA} = K_{t-1}^{PIA} + I_t$. In doing so we need to deal with the fact that the accounting data sampling scheme implies that firms may have incomplete accounting data histories. In other words, a firm may be included in some years, while excluded in others. We handle this by "re-starting" the perpetual inventory algorithm in K_t^{BVA} for the first observation after a "gap". Hence,

$$K_t^{PIA} = \begin{cases} K_t^{BVA} & \text{if } K_{t-1}^{PIA} \text{ not available,} \\ K_{t-1}^{PIA} + I_t & \text{if } K_{t-1}^{PIA} \text{ available,} \end{cases} \tag{7}$$

where K_t^{BVA} and I_t have been inflated to 2006 levels by the implicit deflator in value added within Manufacturing.

Flexible and non-dynamic inputs Our estimation procedure relies on the presence of a flexible and non-dynamic input. here we describe how we construct two candidates for such a variable. The first is material cost or cost of intermediates (denoted KRH in 1995-2003 and $KRHE$ in 2004-2006). Note that KRH and $KRHE$ excludes cost of energy ($KENE$), cost of subcontractors ($KLOE$), housing rent ($UDHL$), purchase of minor equipment ($UASI$), other external costs ($OEEU$) and secondary costs ($SEUD$). We denote this measure of material cost based on KRH and $KRHE$ by M. That is,

$$M = \begin{cases} KRH & \text{in 1996-2003,} \\ KRHE & \text{in 2004-2006.} \end{cases} \tag{8}$$

Our second candidate variable for a flexible and non-dynamic input is energy. Energy cost is recorded in $KENE$ and we denote this variable by E. Hence,

$$E = KENE \tag{9}$$

A.3 Wage bill

The REGNSK files allow us to compute the firm's wage bill. Although we do not use the wage bill information in the empirical analysis presented in this paper (individual wages are needed), we retain the information to check the consistency between the individual wages obtained from IDA and the firm wage bills obtained from REGNSK. The wage bill W is defined as follows:

$$W = LGAG + PUDG + AUDG, \tag{10}$$

where $LGAG$ is wages and salaries, $PUDG$ is pension expenses, and $AUDG$ is other costs of social security.

A.4 Firm size

Firm size is measured in terms of a firm's work force. The REGNSK files provide two such firm size measures. The first is the annual workforce in full time equivalents (i.e., a measure of the total number of full-time employed workers used to generate the year's production). We denote this measure FTE. The other measure is a simple headcount of the workforce in the last week of November. We do not make use of this latter size measure, but retain it to check the consistency of the information in the three data sources, REGNSK, IDA, and FIDA, that are used to construct our MEE panel.

B Tables

	Manufacturing
Full sample	
Number of observations	4868912
Number of workers	959139
Number of firm-years	152371
Number of groups	5185
Number of estimable effects	1106325
Largest group of "connected" workers and firm-years	
Number of observations	4843177
Number of workers	949686
Number of firm-years	136586
Number of estimable effects	1086271

Table 8: Summary statistics on the connectedness of the analysis samples. For precise description of the concept of connectedness applied here, see Abowd, Creecy and Kramarz (2002).