

Wage and Labor Productivity Dispersion: The Roles of Total Factor Productivity, Labor Quality, Capital Intensity, and Rent Sharing*

Jesper Bagger[†] Bent Jesper Christensen[‡]
Royal Holloway Aarhus University

Dale T. Mortensen[§]
Northwestern University
Aarhus University

January 23, 2014

*We thank John Abowd, Paul Beaudry, Steve Bond, Melvyn Coles, Manolis Galenianos, John Haltiwanger, John Kennan, Francis Kramarz, Peter Kuhn, Rasmus Lentz, Iourii Manovskii, David Rivers, Jean-Marc Robin, participants at the 2009 Structural Labor Market Models and Policy Analysis conference at Sandbjerg Manor in Denmark, the 2009 Danish Econometric Society Annual Meeting, the 2010 Royal Economic Society Annual Conference, the 2010 joint LMDG/CREAM workshop at the BI Norwegian School of Management, the 2010 SOLE/EALE conference, the 2010 SED Annual Meeting, the 2010 Econometric Society World Congress, the 2011 NBER Workshop on Micro and Macro Perspectives on the Aggregate Labor Market, the 2012 SOLE conference, the 2013 EEA conference, the 2014 ASSA meeting, and seminar participants at Oxford University, Royal Holloway, Science Po, University of Essex, University of Leicester, University of Southampton, and UC Santa Barbara for helpful and constructive comments. The authors gratefully acknowledge financial support from the Danish Social Science Research Council and the Cycles, Adjustment, and Policy research unit, CAP, Department of Economics and Business, Aarhus University.

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

[‡]Department of Economics and Business, Aarhus University, Fuglesangs Alle 4, DK-8210 Aarhus V, 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

Abstract

Firm labor productivity and average wages paid by firms vary considerably and are positively correlated. These observations can be rationalized either by exogenous TFP heterogeneity in firm productivity coupled with rent sharing or by differences in capital intensity and in the quality of labor inputs. This paper ascertains the extent to which these factors provide an explanation of the observations using Danish matched employer-employee data. Using the worker fixed effect in a wage equation as a measure of worker ability, and a combination of ability and occupational composition for labor force quality, we find that TFP heterogeneity explains more of the observed labor productivity dispersion than differences in capital intensity and labor force quality in each of the industries considered, and that variation in labor force composition explains more than ability differences. Both differences in labor force quality and rent sharing are important in explaining firm level wage dispersion, whereas rent sharing is most important for the positive correlation between average firm wage and labor productivity.

Keywords: Bargaining, Rent sharing, Wage dispersion, Productivity dispersion, Worker heterogeneity, Firm heterogeneity, Matched employer-employee data

JEL codes: C33, C78, J21, J24, J31

1 Introduction

Firm heterogeneity is ubiquitous in available micro data sets. Labor productivity and the average wage paid vary substantially across firms and are positively correlated, even within narrowly defined industries.¹ In the case of (large) Danish privately owned firms, documented in this paper, the 90/10 percentile ratio for the distribution of labor productivity ranges between 2.1 and 2.9 depending on the industry considered, and the corresponding range for the distribution of average wage bill paid per employee is 1.3 to 2.0, while the cross firm correlation between the two variables ranges between .4 and .5. Similar relationships are documented for other countries and data sources.²

Two broad lines of explanation for this robust empirical observation have been brought forward. Some argue that the observed productivity and wage dispersion 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 and firm rents are shared (Krueger and Summers, 1989).³ Undoubtedly, both lines of explanation play a role, and it is important to disentangle their relative importance. If input quality and composition is the main driver of productivity dispersion across firms, then policies involving training and skill upgrading of the labor force may support aggregate growth. On the other hand, if productivity dispersion across firms primarily reflects heterogeneity in intrinsic total factor productivity (TFP), then factor reallocation is potentially output enhancing. The importance of this possibility is emphasized by Hsieh and Klenow (2009), who report 90/10 TFP ratios of approximately 5.0 for the major developing countries of China and India. Moreover, in the case of systematic productivity differences across firms, rent sharing provides the link between individual

¹As Foster, Haltiwanger, and Krizan (2001) point out, one needs to distinguish between productivity measures, both labor productivity and total factor productivity, expressed in physical units and monetary value. Throughout the paper we measure productivity in value terms. One defence is the fact that the value of output is maximized only when the values of marginal products per unit of input are equalized across firms, as Hsieh and Klenow (2009) point out, which occurs in the competitive case only in the absence of input market distortions. Of course, doing so ignores distortions attributable to imperfect competition in the market for output, but except in the case of the manufacturing industry there is little choice since unit measures of output generally do not exist. Bartelsman, Haltiwanger, and Scarpetta (2013) conduct their analysis using revenue based productivity measurements.

²See, e.g., Baily, Hulten, and Campbell (1992), Foster, Haltiwanger, and Krizan (2001), Syverson (2004), Postel-Vinay and Robin (2006), and Bartelsman, Haltiwanger, and Scarpetta (2013). A related finding from the foreign trade literature is that firms that export tend to be more productive and pay significant wage premiums, see, e.g., Bernard and Jensen (1995), Bernard, Eaton, Jensen, and Kortum (2003), Eaton, Kortum, and Kramarz (2004, 2011), and Pedersen (2009).

³Recently, formal models of rent sharing in markets with search and matching friction have been proposed and estimated. These include Burdett and Mortensen (1998), Postel-Vinay and Robin (2002), Cahuc, Postel-Vinay, and Robin (2006), Cahuc, Marque, and Wasmer (2008), Mortensen (2003, 2010), among others.

wages and firm productivity, and thus ties the dispersion in productivity to the wage distribution. Hence, as wage dispersion motivates job search (see, e.g., Christensen, Lentz, Mortensen, Neumann, and Werwatz, 2005), rent sharing is a potentially important vehicle for labor reallocation. To date there has been no quantitative assessment of the relative importance of differences in input composition and quality versus systematic TFP heterogeneity coupled with rent sharing as explanations of the cross firm differences in labor productivity and wages and their correlation.

In this paper, we offer such an assessment. Specifically, we construct a wage bargaining model that incorporates both variation in input quality and a stochastic TFP process. Input quality is determined by capital intensity, as well as average ability and composition of the labor force. In particular, we consider multiple job types or occupations, and the labor force composition of the firm is the distribution of ability adjusted labor across occupations. Our model of wage determination is composed of the firm's production function, with inputs that include capital and ability adjusted labor of the various occupations, and a wage equation that allows for rent sharing. The wage equation is based on the Stole and Zwiebel (1996) and Smith (1999) model of bilateral bargaining between an employer and each of its employees in a multi-person firm, extended to allow for search frictions and multiple labor inputs by Cahuc, Marque, and Wasmer (2008). Observed firm differences in labor productivity and wages and their correlation may then more heavily reflect differences in capital intensity and labor force quality (average ability and composition) or differences in TFP acting through the rent sharing channel, depending on parameter values in our encompassing model. For estimation, we obviously need firm data, and to measure the labor input on an ability adjusted basis, we use high quality Danish matched employer-employee (MEE) data, focusing on large firms and segmenting the data into four industries: Manufacturing; Wholesale & Retail Trade; Transport, Storage & Communications; and Real Estate, Renting & Other Business Activities, over the period 1995-2007.

Our econometric approach contributes to the the literature on estimation of production functions by embedding these in a structural equation system involving worker-occupation effects as well as firm-occupation-time effects derived from an individual level wage decomposition. The estimated individual effects from the wage decomposition are used as measures of worker ability in the construction of labor input aggregates for each firm, one for each of the four occupations we consider: Managers, Salaried workers, Skilled workers, and Unskilled workers.⁴ The production function estimates use these aggregates as labor inputs together with a capital measure and

⁴Iranzo, Schivardi, and Tosetti (2008) construct production function inputs from skill measures obtained as individual effects from a similar individual wage decomposition, but they do not allow for occupation-specific worker effects or time-and-occupation specific firm effects, nor do they explicitly model the wage setting procedure. All in all they consider a related, but separate, set of questions regarding wage and productivity dispersion.

a stochastically varying TFP shock. The estimated time-varying occupation-specific firm effects from the decomposition of individual level wages admit a structural interpretation as the firm-specific unit prices of human capital, and these are used in the estimation of a firm level wage equation arising from the proposed bargaining game.

Early attempts to take account of changes in human capital variables in the decomposition of aggregate productivity growth by Jorgenson, Gollop, and Fraumeni (1987) concluded that these were small contributors. Although the availability of MEE data with detailed worker information 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, surprisingly little research has been done on the subject. We are aware of two recent projects that exploit these data bases. Fox and Smeets (2011) use Danish MEE data and estimate Cobb-Douglas and translog production functions with quality weighted employment and capital inputs on Danish data. The worker characteristics include gender, education, experience, firm tenure, and age. Including these labor quality variables as explanations of the total labor input with weights determined by the data accounts for very little of the variance in firm productivity above and beyond that implied by simply using total firm employment as the labor input. Applying similar methods to Norwegian MEE data, Irarrazabal, Moxnes, and Ulltveit-Moe (2010) approach the issue of labor input quality from the perspective of the empirical trade literature. They find that observable worker characteristics explain about 25% of the average productivity differential between exporting and non-exporting firms. 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.

If variation in input quality is not the main driver of productivity dispersion, then this suggests an important role for TFP heterogeneity, and hence a potential for growth enhancing factor reallocation, as highlighted by Restuccia and Rogerson (2008) and Hsieh and Klenow (2009). There is a sizeable literature engaged in documenting and analyzing the extent and costs of misallocation of production factors; see e.g. Restuccia and Rogerson (2013) and the references therein.⁵ Traditionally, this literature has been concerned with the impact of misallocations on TFP measurements in a cross-country context, Hsieh and Klenow (2009) being a recent and prominent example of this. Even if our analysis does not have a cross-country element, we contribute to this literature in several ways. First, we provide measurements on the extent of misallocation with a sophisticated control for worker ability, allowing us to consider misallocation of talent. Second, we conduct our analysis on four major industries, including Manufacturing.

⁵Restuccia and Rogerson (2013) is the Editorial for a special issue of the Review of Economic Dynamics on misallocation and productivity.

Most other studies focus exclusively on Manufacturing. Third, our analysis and measurements are grounded in a structural model with a clear connection between data and our theory of misallocation.

Rent sharing would support such labor reallocation, since it implies that the productivity dispersion induces wage dispersion. With wages dispersed, there is a motive for job search. Further, rent sharing implies that productive firms pay more, so that the search process tends to reallocate workers to more productive firms. However, although the papers cited point to the importance of TFP heterogeneity, there has been no attempt to combine the production function based analysis of TFP and input quality with a model of rent sharing.

In the present paper, we do both: We use the production function to assess the roles of input quality and TFP in firm level labor productivity dispersion, and we use the wage equation to assess the roles of labor force ability and composition versus rent sharing in wage dispersion. Upon estimation of our encompassing model, we apply standard variance-covariance decompositions to quantify the relative importance of the input quality versus TFP heterogeneity and rent sharing explanations. Our production function includes a firm level labor force quality measure reflecting individual level abilities, measured using worker-occupation fixed effects, as well as the effect of occupational composition through the estimated relative productivities by occupation. We find that capital intensity and labor force quality as measured only account for about 9 and 20 percent, respectively, of the variance in log labor productivity in Manufacturing. Heterogeneity in TFP accounts for 72 percent of the labor productivity variance. In the other three industries considered, TFP variation similarly accounts for more than half of labor productivity dispersion. Capital intensity and labor force quality together never account for more than 40 percent, and within labor force quality, differences in occupational composition are more important than ability differences. In all industries, at least half of the labor productivity dispersion accounted for by TFP heterogeneity stems from persistent as opposed to transitory TFP variation, implying that persistent TFP heterogeneity by itself is more important for labor productivity dispersion than differences in labor force quality. Our results show that firms differ systematically in productivity, and not only because of different inputs, thus indicating a scope for productivity improving factor reallocation across firms.

We find that rent sharing exists in all industries. A decomposition of firm wage variance as well as a decomposition of the covariance between the firm wage and firm labor productivity is computed for the estimated model. Again, the relevant firm level labor force quality measure combines individual level abilities with the effect of occupational composition, the latter now operating through workers' estimated outside options and bargaining powers. Although cross-firm differences in this labor force

quality measure explain 12 percent of the variance of firm wages in Manufacturing, the rent sharing effect through firm labor productivity variation explains considerably more, at 33 percent, and similarly in other industries. The calculated shares rely on structural parameters from estimation of our encompassing model. For comparison, an alternative projection based decomposition leaves differences in labor force quality and rent sharing about equally important for firm wage dispersion. Furthermore, rent sharing explains 75 percent and labor force quality differences 25 percent of the covariance between firm wage and labor productivity in Manufacturing by the projection method. The picture is similar in other industries, and the relative importance of rent sharing compared to labor force quality in explaining the covariance is even greater based on the structural parameters. In sum, differences in labor force quality explain about as much of the variance in firm wages as rent sharing, but rent sharing is the most important reason for the observed positive correlation between firm wages and labor productivity. This suggests that although frictions prevent that all workers arrive at the intrinsically most productive firm, at least some incentives for labor reallocation are in place, in that job search is driven by workers chasing higher wages.

Recent literature suggests that labor markets are characterized by some degree of positive assortative matching.⁶ Our general framework imposes conditions that suffice for disentangling the roles of labor force quality and rent sharing in generating dispersion and correlation, without tying down the particular equilibrium assignment of worker types (abilities) to firm types (TFP). Hence, we do not provide any specific interpretation of our results in terms of labor market sorting, although our novel measurements and decompositions of the joint distribution of productivity and wages are broadly consistent with this possibility.

The rest of the paper is laid out as follows. The model is introduced in Section 2. Section 3 describes the construction of our matched employer-employee panel data set. Section 4 presents the estimation method and empirical results, and Section 5 reports the aforementioned decompositions. Section 6 concludes. Some further details on the data, as well as some additional empirical results, are provided in an online supplement to this paper.

2 The Model

We consider firms that employ multiple workers. In the Stole and Zwiebel (1996) and Smith (1999) bargaining model, the firm bargains with each employee as if s/he were

⁶Eeckhout and Kircher (2011), Lopes de Melo (2013), Lise, Meghir, and Robin (2013), Bagger and Lentz (2013), and Hagedorn, Law, and Manovskii (2013) present theoretical characterizations of sorting patterns in the labor market, as well as empirical evidence through structural estimation of the proposed models.

the marginal worker. The worker is assumed to have an outside option, which in the Cahuc, Marque, and Wasmer (2008) version of the model is taken to be the worker's reservation wage. In the special case of identical workers, the gross profit flow of the firm can be represented as

$$\pi(N, p) = pf(N) - \phi N,$$

where N is employment, p is the firm's factor productivity, ϕ is the wage, and $f(N)$ is the baseline production function, an increasing concave function of employment. The net value of the marginal worker to the firm is $\partial\pi/\partial N$, while the surplus value of employment to the worker is $\phi - b$, where b represents the worker's reservation wage. We introduce asymmetric 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. Hence, the outcome satisfies

$$\beta \left(pf'(N) - \phi - \frac{\partial\phi}{\partial N}N \right) = (1 - \beta) (\phi - b). \quad (1)$$

In other words, the worker receives the share β of the joint surplus. As Stole and Zwiebel (1996) show, the solution to (1), the wage bargaining outcome function, is

$$\phi(N, p) = (1 - \beta)b + p \int_0^1 z^{\frac{1-\beta}{\beta}} f'(zN) dz. \quad (2)$$

In the constant marginal product case, the wage reduces to the average of the outside option b and the value of the marginal product, $(1 - \beta)b + \beta pf'$, 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 $pf(N)/N$ and the wage,

$$\phi(N, p) = (1 - \beta)b + \frac{\beta\alpha}{1 - \beta + \alpha\beta} \left(\frac{pf(N)}{N} \right).$$

Thus, the wage is equal to the average product when workers have all the bargaining power ($\beta = 1$) and the outside option when they have none ($\beta = 0$). In general, the wage is an increasing function of both workers' share of the rent β , the labor elasticity of output α , and the outside option b .⁷

Cahuc, Marque, and Wasmer (2008) generalize the outcome of the Stole and Zwiebel bargaining problem to allow for any number H of different types of labor that are im-

⁷Of course, the bilateral bargaining outcome (2) determines the wage if and only if it is consistent with participation conditions. Mortensen (2010) shows that these do not bind in equilibrium.

perfect substitutes in general, and other quasi-fixed factors, such as capital. Given a general production function of the form

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

where p represents firm TFP, K is the capital stock, and $\mathbf{L} = (L_1, L_2, \dots, L_H)$ the vector of labor inputs, the generalization of (2) to the case of H different labor input types is

$$\phi_h(K, \mathbf{L}, p) = (1 - \beta_h)b_h + p \int_0^1 z^{\frac{1-\beta_h}{\beta_h}} f_h(K, \mathbf{L} \circ \mathbf{A}_h(z)) dz, \quad h = 1, \dots, H. \quad (3)$$

Here, β_h is the bargaining power and $pf_h = \frac{\partial Y}{\partial L_h}$ the value of marginal product of type h labor, and $\mathbf{L} \circ \mathbf{A}_h(z)$ is the vector

$$\mathbf{L} \circ \mathbf{A}_h(z) = \left(L_1 z^{\frac{\beta_1}{1-\beta_1} \frac{1-\beta_h}{\beta_h}}, L_2 z^{\frac{\beta_2}{1-\beta_2} \frac{1-\beta_h}{\beta_h}}, \dots, L_H z^{\frac{\beta_H}{1-\beta_H} \frac{1-\beta_h}{\beta_h}} \right), \quad h = 1, \dots, H,$$

the Hadamard product of the employment vector \mathbf{L} and the vector $\mathbf{A}_h(z)$ with k' th element $z^{\frac{\beta_k}{1-\beta_k} \frac{1-\beta_h}{\beta_h}}$, for $k = 1, \dots, H$. In our empirical analysis, we follow Christensen, Lentz, Mortensen, Neumann, and Werwatz (2005) and interpret labor types as four (aggregate) occupations: Managers, Salaried workers, Skilled workers, and Unskilled workers. In the rest of the paper we therefore refer to the H labor types as occupations.

We allow individuals to differ with respect to their abilities in each of the H occupations. Let a_{ih} represent the ability or skill of individual i in occupation h . For empirical tractability we restrict attention to the case where complementarities arise only between occupations; within each occupation, individual workers differ with respect to ability, but are perfect substitutes. A firm's total labor input in occupation h is thus $L_h = \int_{\Omega_h} a_{ih} di$, where Ω_h is the index set of workers employed in occupation h in the firm. In the case of heterogenous workers, we retain N to indicate the total number of (non-ability adjusted) workers in the firm. That is, N is a headcount.

If one views ϕ_h as the piece rate wage per unit of ability in occupation h , then the wage of individual i in occupation h is $w_{ih} = a_{ih}\phi_h$, which implies that the log of individual wage rates can be written as a linear function of a worker effect and a firm effect,

$$\ln w_{ih} = \ln a_{ih} + \ln \phi_h(K, \mathbf{L}, p), \quad (4)$$

where the firm effect, the second term, depends on the employing firm's input composition and TFP. Below, we generalize the theoretical framework to a dynamic setting, which makes (4) amenable to empirical analysis using MEE panel data. In this extended framework, (4) implies that worker ability and firm level piece rate wages are identified as the worker and firm fixed effects in a panel regression model of log wages.

Thus, specification (4) can be regarded as a generalization of the Abowd, Kramarz, and Margolis (1999) individual log wage regression model to explicit endogenous formation of the firm effects, and we interpret the firm effect in the equation as the firm level price of human capital per ability unit.

2.1 A Tractable Case

Our preferred empirical specification involves a Cobb-Douglas production function with a log linear labor aggregator $\prod_{h=1}^H L_h^{\gamma_h}$ such that the elasticity of substitution between occupations is unity,

$$Y = pK^{\alpha_K} \left(\prod_{h=1}^H L_h^{\gamma_h} \right)^{\alpha_L}, \quad (5)$$

where α_K and α_L , respectively, represent the capital and labor aggregate elasticities of output, and γ_h is a measure of the relative productivity of occupation h , with $\sum_{h=1}^H \gamma_h = 1$. In this case, $pf_h = \frac{\partial Y}{\partial L_h} = \gamma_h \alpha_L \frac{Y}{L_h}$ so that

$$pf_h(K, \mathbf{L} \circ \mathbf{A}_h(z)) = \gamma_h \alpha_L \left(\prod_{k=1}^H z^{\gamma_k \alpha_L \frac{\beta_k}{1-\beta_k} \frac{1-\beta_h}{\beta_h}} \right) z^{-1} \frac{Y}{L_h}, \quad (6)$$

and equation (3) takes the simple form

$$\phi_h(K, \mathbf{L}, p) = (1 - \beta_h) b_h + \left(\frac{\alpha_L \gamma_h \frac{\beta_h}{1-\beta_h}}{1 + \sum_{k=1}^H \alpha_L \gamma_k \frac{\beta_k}{1-\beta_k}} \right) \frac{Y}{L_h}. \quad (7)$$

Thus, the firm average wage is

$$\bar{w}(K, \mathbf{L}, N, p) = \sum_{h=1}^H \frac{\phi_h L_h}{N} = \frac{L}{N} \sum_{h=1}^H (1 - \beta_h) b_h \frac{L_h}{L} + \left(\frac{\sum_{h=1}^H \alpha_L \gamma_h \frac{\beta_h}{1-\beta_h}}{1 + \sum_{h=1}^H \alpha_L \gamma_h \frac{\beta_h}{1-\beta_h}} \right) \frac{Y}{N}, \quad (8)$$

with N denoting total (non-ability adjusted) labor force size, $L = \sum_{h=1}^H L_h$ total ability input, and b_h now interpreted as a worker's reservation (or outside) piece rate wage. The first term in the firm level wage function reflects labor force quality, appropriately combining the average ability and occupational composition of the firm's labor force. Firms that employ workers of higher average ability $\frac{L}{N}$, and firms allocating greater fractions $\frac{L_h}{L}$ of total ability input to occupations with high outside options $(1 - \beta_h) b_h$, pay higher wages, consistent with the argument in Murphy and Topel (1990). The second term on the right hand side represents a rent sharing effect, as in Krueger and Summers (1989), in our encompassing model a linear function of measured labor

productivity, $\frac{Y}{N}$, with common slope across firms.⁸ Given rent sharing, dispersion in average productivity across firms induces wage dispersion.

In (8), the relative importance of labor force quality versus rent sharing in the firm wage structure depends on the structural parameters of our model. If $\beta_h = 0$ for all h , so that workers' bargaining power vanishes, composition and ability differences are the only source of firm wage variation. As β_h increases, for a given distribution of average productivity, the relative importance of rent sharing for dispersion increases. Indeed, as β_h approaches unity, we have the limiting case $\bar{w} = \frac{Y}{N}$: The average wage paid out is given by average productivity, regardless of labor force quality differences. In Section 5, we use the firm wage equation to analyze firm level wage variation and the covariance between firm wages and measured productivity.

2.2 Firm Dynamics

In this section, we relax the assumption that inputs and TFP are fixed. Instead, TFP is a stochastic process and labor and capital are quasi-fixed factors of production. Formally, we consider a discrete time formulation in which $t = 1, 2, \dots$ indexes periods and the TFP sequence $\{p_t\}$ is a first order Markov process. In the empirical analysis we shall take p_t to follow a linear $AR(1)$ process. We assume that labor reallocation, i.e., hiring and separations, and investment in capital take place after current TFP is observed, and that investments in period t materialize within the same period. Wage bargaining takes place after labor reallocation and investment decisions. As a period in our empirical model is a year, this accommodates the possibility of a time-to-build of less than a year. There is nothing salient about this particular timing assumption. We proceed this way only because our empirical analysis favors specifications where both labor and capital are allowed to be potentially endogenous in relation to contemporaneous TFP (see Section 4). We could easily adapt a different structure, e.g., with investment decisions in period t realized in period $t + 1$, implying a time-to-build of at least one year.

Although one can generalize the formulation to any finite number of occupations, for expositional simplicity we sketch 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

$$L_t = H_t + (1 - s)L_{t-1}, \quad (9)$$

⁸A similar additive decomposition of the firm average wage into a labor quality and a rent sharing term applies in the case of a CES labor aggregator and common bargaining powers across occupations. In our empirical work below, we cannot reject the Cobb-Douglas restriction on the CES labor aggregator when estimating the production function.

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

$$K_t = I_t + (1 - \delta)K_{t-1}, \quad (10)$$

where I_t represents gross additions to physical capital determined in period t and δ is a fixed depreciation rate. The hiring and investment decisions in period t are made contingent on L_{t-1} , K_{t-1} , and p_t . Of course, as in (7), the piece rate wage is the outcome of the wage bargaining in period t which is simultaneous with total labor and capital input.

Suppose that the costs of hiring and gross investment in period t take the forms $c_L(\frac{H_t}{L_{t-1}})L_{t-1}$ and $c_K(\frac{I_t}{K_{t-1}})K_{t-1}$, where $c_L(\cdot)$ and $c_K(\cdot)$ are positive, increasing and convex functions. This is implied, e.g., if the cost functions are increasing, convex, and homogenous of degree one in stock and gross flow. Then the value of the firm associated with the optimal decisions solves the Bellman equation

$$V(L_{t-1}, K_{t-1}, p_t) = \max_{H_t, I_t} \left\{ p_t f(K_t, L_t) - \phi(K_t, L_t, p_t) L_t - c_L\left(\frac{H_t}{L_{t-1}}\right) L_{t-1} - c_K\left(\frac{I_t}{K_{t-1}}\right) K_{t-1} + \Lambda E[V(L_t, K_t, p_{t+1}) | p_t] \right\} \quad (11)$$

subject to equations (9) and (10), where $\Lambda \in (0, 1)$ is a discount factor. For purposes of the estimation that follows, we do not need to take a stand on the particulars of this dynamic formulation, other than the specification of what the employer knows when making the hiring and investment decisions. Hence, more complicated and possibly more realistic formulations, for example explicitly accounting for employed worker search, 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 information for the period 1995-2007. 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 1995-2007 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

Table 1: Sampling scheme for accounting data

Nov. workforce	Fraction sampled	Years in/out
0-4	0	—
5-9	.10	1/9
10-19	.20	2/8
20-49	.50	3/3
> 49	1.00	—

Note: In addition, firms that generated revenue exceeding DKK 100 mill. (Wholesale: DKK 200 mill.) in the preceding year are always sampled. Source: Statistics Denmark.

and the ID of the employer. In addition, IDA is rich on background information on individuals (including a unique person ID), whereas employer information in IDA is limited, consisting of firm ID, ownership information, 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, this way constructing a comprehensive MEE panel data set. In the following, we provide some more details 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 analysis sample.

3.1 Data Sources

3.1.1 Employer Data

Accounting data are available for the period 1995-2007. Industry coverage increases over time, starting in the initial year 1995 with Manufacturing and parts of Wholesale & Retail Trade.⁹ Coverage gradually expands until by 1999 most industries are included. The survey from which the accounting data originate has a rolling panel structure where firms are selected based on the size of their last week of November workforce in the previous year. All firms with at least 50 employees are surveyed, and some with less. The detailed 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, measures of the firm's capital stock and various other relevant measurements, including annual workforce measured in Full Time Equivalent (FTE) workers. Our online supplementary material contains details on these computations. We refer to these data as the employer data.

⁹We classify industries according to NACE rev. 1 (see EUROSTAT, 1996).

3.1.2 Employee Data

We retrieve data on workers from IDA for the years 1995-2007 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. We use education and occupation information from the person files, industry code from the establishment files,¹⁰ and average hourly wage from the job files. We refer to these data as the employee data.

3.2 Analysis Data

Merging the employer and the employee data using FIDA yields an MEE panel with 24 million observations on 3.7 million workers and 550,00 firms (some of which with accounting data information). Our empirical analysis utilizes both an individual level panel and an associated firm level panel. Both panels are extracted from the MEE panel. Here, we provide a short description of the selection criteria. Details on the merging procedure, and on the selection of the analysis data, are relegated to the online supplementary material.

3.2.1 Individual Level Panel

First, we restrict attention to the following industries (NACE rev. 1 section and years in the sample in parentheses): Manufacturing (D, 1995-2007), Wholesale & Retail Trade (G, 1999-2007), Transport, Storage & Communications (I, 1999-2007), and Real Estate, Renting & Business Activities (K, 1999-2007). To verify the robustness of our results, we analyze the four largest sub-industries within Manufacturing in the online supplementary material.

Second, we discard observations on workers aged below 18 and observations with invalid or missing education information, and define four occupational categories according to ISCO-88 grouping: Managers, Salaried workers, Skilled workers, and Unskilled workers. Workers with missing occupation data are assigned to an occupation based on length of education. The occupational categories are consistent with Christensen, Lentz, Mortensen, Neumann, and Werwatz (2005).

Third, we discard observations with missing or non-positive wages and firm-years

¹⁰Each firm has one or more establishments. We define the industry of a firm as that of its largest establishment.

with non-missing and non-positive value added, hours, or capital stock. The annual industry and occupation specific distributions of individual level wages and the (non-employment weighted) annual industry-specific distributions of hourly value added are trimmed at the top and bottom 1%. All nominal variables are detrended using the NACE rev. 1 section-specific implicit deflator in hourly value added.

Fourth, we retain only firm-years where at least two workers are employed in each of the four occupations. This selection criterion excludes small firms, leading to a considerable loss in the number of observations in the individual level panel. The loss of information from the employer data is moderate since these are already large firms, due to the sampling scheme for the accounting data survey outlined in Table 1.

Fifth, to ensure identification of the individual level log wage regression with worker-occupation and firm-year-occupation indicators, we select the largest group of “connected” workers and firm-years in each occupation and industry.¹¹ Intuitively, a set of workers and firm-years is connected when it contains all the workers who were ever linked to any of the firm-years in the set, and all the firms-years to which any of the workers in the set was ever linked. In contrast, when two sets of workers and firm-years are not connected (within an occupation and industry), no firm-year in the first set has ever been linked to a worker in the second, nor has any worker in the first set ever been linked to a firm-year in the second.¹² The vast majority of workers and firm-years are connected in one large group in each occupation and industry.

Table 2 provides basic summary statistics on the resulting individual level panel. Manufacturing is the largest of the four industries, in terms of number of observations, workers, and firm-years. In terms of firm-years, Wholesale & Retail Trade is second largest, followed by Real Estate, Renting & Business Activities, and the smallest industry by this measure is Transport, Storage & Communications.

3.2.2 Firm Level Panel

The firm level panel is constructed from the individual level panel in the following way. First, we select only firm-years for which we have accounting data information. Second, we select only firm-years that are represented in all four occupations.¹³ Third, we select only firm-years that are part of a sequence of at least three consecu-

¹¹The identification of this type of two-way error component regression is studied by Abowd, Creedy, and Kramarz (2002) in the case where the error components are a worker-specific permanent effect and a firm-specific permanent effect. In the absence of time-varying observable covariates, their identification result carries over to our context where the error components are a worker-occupation effect and a firm-year-occupation effect, with identification established within each occupation and industry separately.

¹²This description of the notion of connectedness is a slight paraphrasing of the description given in Abowd, Creedy, and Kramarz (2002, p. 3).

¹³This amounts to selecting firm-years that for each occupation belong to the largest group of connected workers and firm-years, such that each worker-occupation effect and firm-occupation-time effect is identified.

Table 2: Summary Statistics for Individual Level Panel

	Manu- facturing	Wholesale & Retail Trade	Transport, Storage & Communi- cations	Real Estate, Renting & Business Activities
Years	1995-2007	1999-2007	1999-2007	1999-2007
Managers				
Observations	366,541	93,329	53,606	240,559
Workers	94,020	37,313	19,774	84,037
Firm-years	15,979	4,659	951	4,253
Salaried workers				
Observations	486,144	209,262	151,504	195,797
Workers	123,315	76,229	44,962	78,748
Firm-years	17,023	7,714	1,466	4,441
Skilled workers				
Observations	2,034,107	595,564	427,624	225,943
Workers	478,527	256,643	140,765	123,263
Firm-years	18,597	8,326	1,671	4,647
Unskilled workers				
Observations	415,988	176,662	115,497	201,905
Workers	175,976	97,388	60,417	116,626
Firm-years	15,807	6,326	1,433	3,582

Note: Unit of observation is a worker-year. Number of observations, workers, and firm-years are reported for the largest set of connected workers and firm-years (cf. Abowd, Creedy, and Kramarz, 2002), by industry and occupation.

Table 3: Summary Statistics for Firm Level Panel

	Manu- facturing	Wholesale & Retail Trade	Transport, Storage & Communi- cations	Real Estate, Renting & Business activities
Years	1995-2007	1999-2007	1999-2007	1999-2007
Number of firm-years	9,962	2,332	408	1,288
Managers/firm-year	31.77	27.92	66.48	107.34
Salaried workers/firm-year	42.35	56.92	203.10	75.79
Skilled workers/firm-year	177.16	159.60	455.60	191.75
Unskilled workers/firm-year	36.02	48.13	105.93	99.68
Workers/firm-year	287.30	292.57	831.11	374.57

Note: Unit of observation is a firm-year. The firm level panel is constructed by aggregating individual level information from the individual level panel (see Table 2) to the firm level subject to selection criteria described in the text.

tive firm-years, thus allowing the construction of lagged differences at the firm level. Table 3 summarizes the sizes of the final firm level analysis panels. From the table, we are primarily dealing with firms with large workforces.¹⁴ Industry size in terms of firm-years is ordered as before, with Manufacturing largest and Transport, Storage & Communications smallest. In terms of size of firms, the ordering of industries is roughly opposite, in each occupation with most workers per firm-year in Transport, Storage & Communications, followed by Real Estate, Renting & Business Activities (except that firms here employ most Managers across all industries), and with firms of about equal size in Manufacturing and Wholesale & Retail Trade. We refer to Manufacturing as the largest and Wholesale & Retail Trade as the second largest industry, keeping in mind the smaller average firm size.

3.3 Descriptive Statistics

We now present a number of facts regarding the firm level relation between wages, productivity, labor, and capital in our analysis samples. Firm level wages are computed as the firm level average of the individual wages, denoted \bar{w} . Labor productivity (hourly) and capital intensity (hourly) are computed by dividing total value added

¹⁴Comparison of the average number of FTE workers per year reported in Table 3 to the corresponding average number of FTE workers in each of the four industries (source: Statistics Denmark) reveals that the share of workers included in our analysis sample (out of the industry total) is about .55 for Manufacturing and about .25 for the other three industries. These differences in coverage may be explained by differences in industry structure (e.g., fewer small firms in Manufacturing), implying that the sample selection criteria imposed here impact the industries differently.

Table 4: 90/10 Ratios for Labor Productivity, Average Wage, and Capital Intensity

	Manu- facturing	Wholesale & Retail Trade	Transport, Storage & Communi- cations	Real Estate, Renting & Business Activities
Labor productivity	2.06	2.49	2.89	2.93
Average wage	1.42	1.65	1.33	1.99
Capital intensity	11.25	17.06	94.20	84.14

Note: Unit of observation is a firm-year.

Y and total capital stock K from the employer data by total hours,

$$N = 12 \times 166.33 \times FTE,$$

where 166.33 hours is the monthly hours norm for jobs under the Danish Industry Confederation and FTE is the reported workforce size from the employer data. Thus, N indicates a non-ability adjusted or headcount measure of employment.

Figure 1 plots kernel-smoothed density estimates of the distributions of labor productivity, firm level average wage and capital intensity for the four industries considered. It is evident that there is considerable dispersion in all three variables. The distribution of labor productivity tends to be slightly to the right of the distribution of firm level wages, and more dispersed. Capital intensity exhibits the highest dispersion among the three variables. Table 4 quantifies the dispersion in terms of 90/10 ratios. All ratios are between 1.3 and 2.0 for wages, between 2.0 and 3.0 for labor productivity, and higher for capital intensity.¹⁵

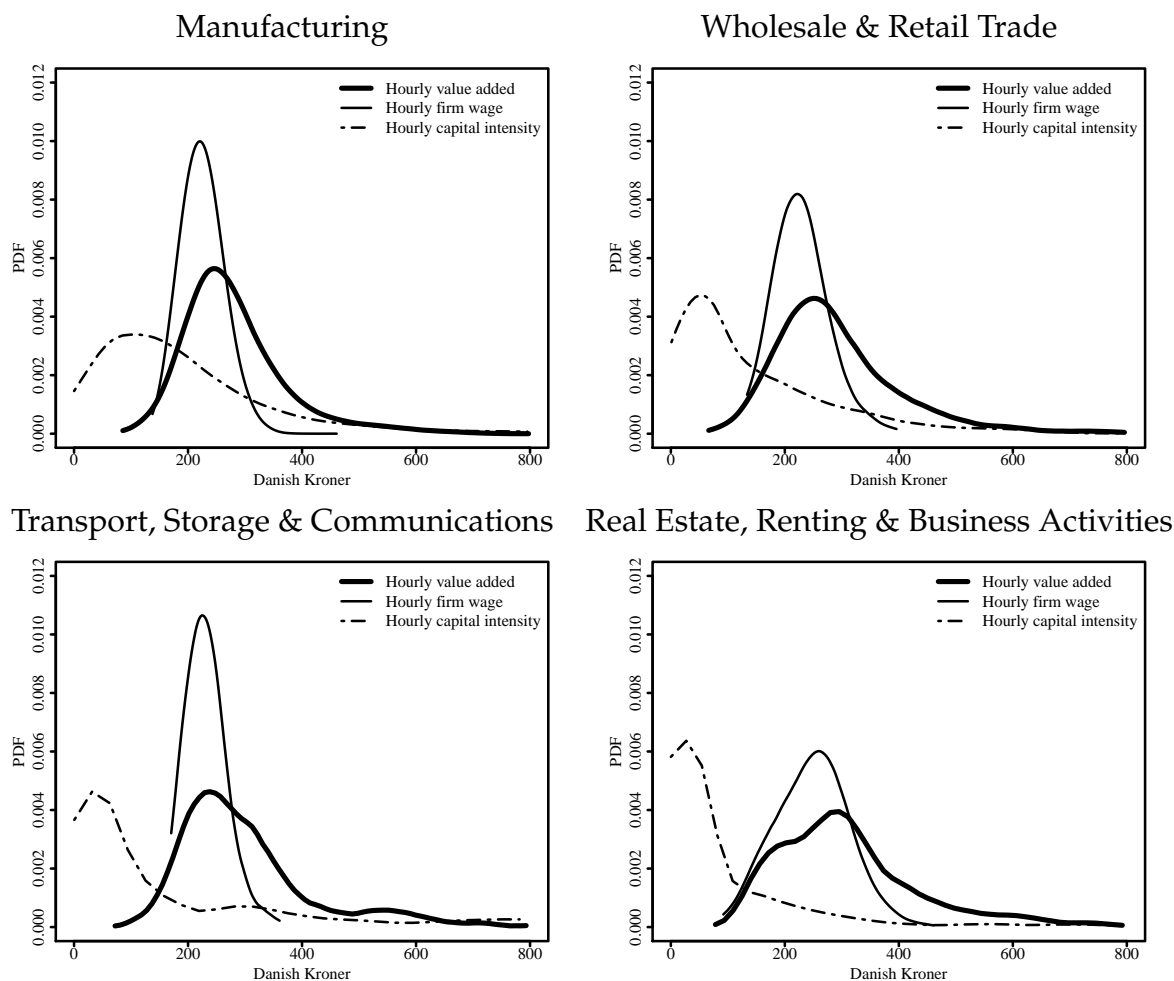
Turning to the relation between wage and labor productivity, Figure 2 shows non-parametric regressions of firm level wages on productivity, along with superimposed parametric (linear) regressions.¹⁶ There is a clear positive relation between a firm's productivity and the average wage it pays to its employees. Moreover, the relation is close to linear and precisely estimated where the data are dense.¹⁷ To further quantify the relations among the variables, Table 5 reports the correlations between average firm wage \bar{w} , labor productivity $\frac{Y}{N}$, and capital intensity $\frac{K}{N}$. The correlation between wages and productivity is narrowly tied down between .40 and .51. This range is about six times wider for the correlation between capital intensity and productivity, with capital potentially most important in Transport, Storage & Communications.

¹⁵The online supplementary material to this paper documents that the dispersion persists when we consider sub-industries within Manufacturing.

¹⁶We do not report standard errors in the figures, but all estimated parameters are strongly significant.

¹⁷See the productivity and wage densities in Figure 1, and the 95% confidence regions (shaded) in Figure 2.

Figure 1: Distributions of Labor Productivity, Average Wage, and Capital Intensity



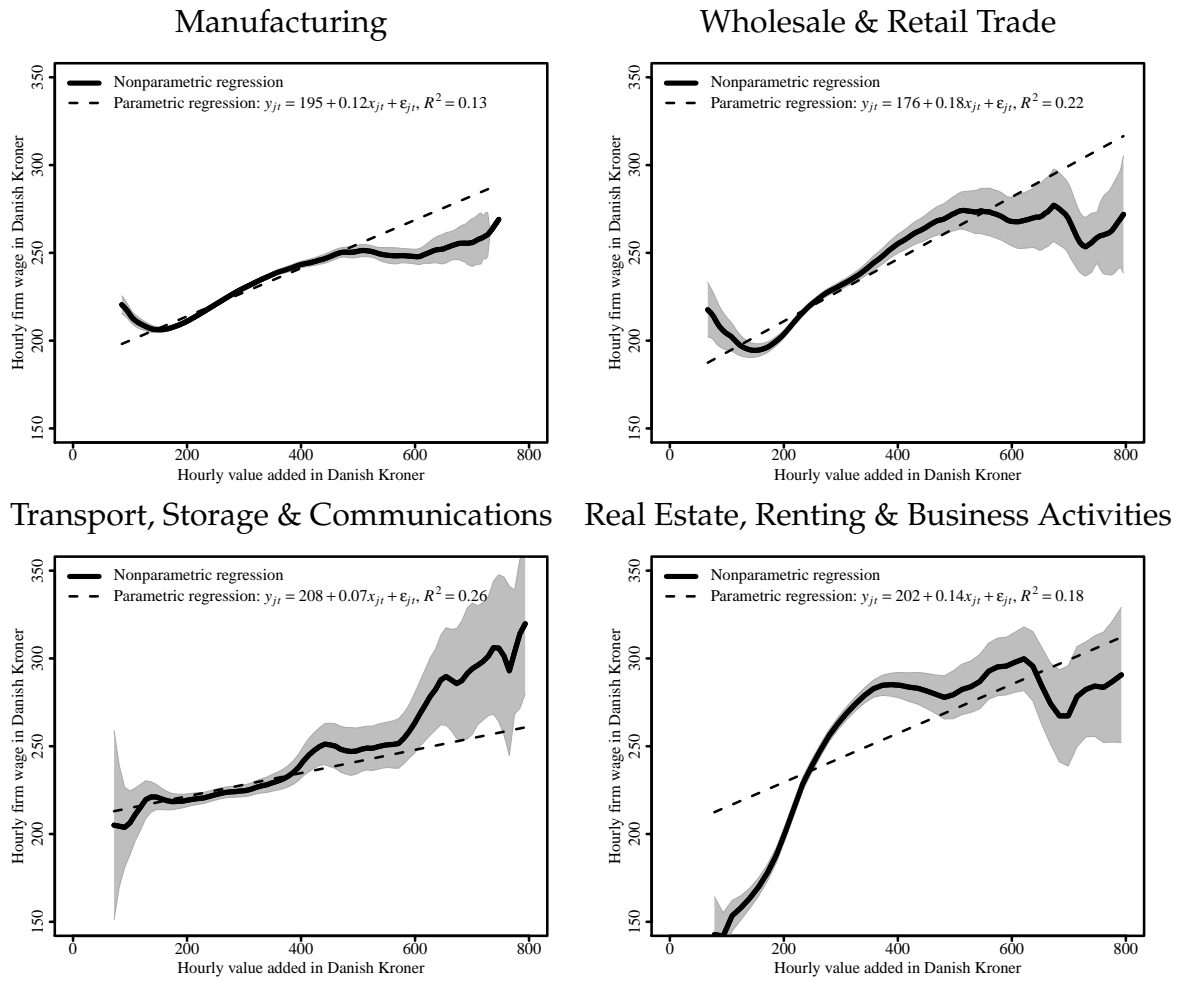
Note: Unit of observation is a firm-year. The densities of labor productivity and firm level average wage are estimated using the Epanechnikov kernel with a bandwidth of 25. The density of hourly capital intensity is estimated using a bandwidth of 35.

Table 5: Correlations Between Labor Productivity, Average Wage, and Capital Intensity

	Manu- facturing	Wholesale & Retail Trade	Transport, Storage & Communi- cations	Real Estate, Renting & Business Activities
$Cor\left(\bar{w}, \frac{Y}{N}\right)$.399	.470	.510	.423
$Cor\left(\frac{K}{N}, \frac{Y}{N}\right)$.185	-.005	.588	.182

Note: Unit of observation is a firm-year.

Figure 2: Nonparametric Regressions of Average Wage on Labor Productivity



Note: Unit of observation is a firm-year. The local polynomial regressions are estimated using the Epanechnikov kernel with a bandwidth of 25. Shaded regions represent estimated 95% pointwise bootstrap confidence bands.

Taken together, the data reveal that more productive firms co-exist with less productive firms, and that high-paying firms co-exist with low-paying firms. Moreover, firm level wages and productivity are positively related. These are well-known empirical facts that have been documented for different data sets, countries, and time-periods (see, e.g., Mortensen, 2003 and Lentz and Mortensen, 2010). We now try to understand and quantify the economic mechanisms driving these relations, using the model from Section 2.

4 Empirical Analysis

Our model gives rise to an equation governing individual level wages, an equation governing firm level average labor productivity, and, for each occupation, an equation governing firm level piece rate wages.¹⁸ This system is taken to the data by conducting the estimation in the corresponding three steps: First, an individual level wage regression; second, a firm level production function; and third, for each occupation a firm level piece rate wage equation.

Given the panel structure of our data, we index individuals by $i = 1, \dots, I$, firms by $j = 1, \dots, J$, occupations by $h = 1, \dots, H$, and time by $t = 1, \dots, T$. Writing $J(i, t)$ and $H(i, t)$ for the functions indicating the firm respectively occupation that individual i is employed in at time t , the MEE data consist of these two functions, along with individual level wages, w_{it} , and firm level value added and capital, Y_{jt} and K_{jt} . Both the individual and firm level panels extracted from the MEE data are unbalanced, i.e., not all workers and firms are observed in every time period.

4.1 First Step: Individual Level Wage Equation

At the individual level, the wage is the product of individual ability and the piece rate wage per ability unit of the employing firm. We allow occupation-specific individual (log) abilities to be subject to shocks of a nature to be specified below, i.e., $\ln a_{iht} = \ln a_{ih} + \varepsilon_{iht}^a$. The individual level (log) wage equation therefore reads

$$\ln w_{it} = \ln \phi_{J(i,t)H(i,t)t} + \ln a_{iH(i,t)} + \varepsilon_{iH(i,t)t}^a \quad (12)$$

which for estimation purposes is conveniently viewed as decomposing individual log wages into time and occupation specific firm effects $\xi_{jht} \equiv \ln \phi_{jht}$, occupation specific worker effects $\zeta_{ih} \equiv \ln a_{ih}$, and the residual terms ε_{iht}^a unexplained by the former two. We impose the following identifying assumption on (12):

¹⁸Individual level wages are given by (4), and the production function by (5), whereas occupation-specific firm level piece rate wages are given by (3) in the general case, and (7) in the Cobb-Douglas case.

Assumption A1 $E[\varepsilon_{iH(i,t)t}^a | i, t, J(i, t), H(i, t)] = 0$, for all $i = 1, \dots, I$ and $t = 1, \dots, T$.

Assumption A1 rules out that individuals change employer or occupation based on ε_{iht}^a . We interpret ε_{iht}^a as idiosyncratic shocks to individual ability, measurement error, or a combination of both, thus allowing estimation of the (many) parameters of (12) by OLS. In particular, identification of the parameter vector (made up of all the ζ_{jht} 's and ζ_{ih} 's in (12)) requires that $\varepsilon_{iH(i,t)t}^a$ is orthogonal to the design matrix consisting of firm-year-occupation and worker-occupation indicators, and that the design matrix is of full rank. The former condition is met under Assumption A1. The latter condition is also met subject to a normalization of a worker or a firm-year effect within each occupation, by taking the individual level analysis panel as the largest set of connected workers and firm-years within each occupation (see Abowd, Creecy, and Kramarz, 2002).¹⁹

There is a large empirical literature documenting the existence of firm fixed effects in wages, see, e.g., Abowd, Kramarz, and Margolis (1999). Our theoretical framework provides a structural foundation for these effects in strategic bilateral bargaining theory. The endogenous firm fixed effects reflect TFP shocks p_{jt} , worker reallocation \mathbf{L}_{jt} , and investment decisions K_{jt} , and are therefore time varying.

From the estimated ζ_{jht} 's and the ζ_{ih} 's we can construct firm-year-occupation specific ability adjusted labor inputs L_{jht} . Indeed, since $\widehat{\zeta}_{ih} \equiv \ln \widehat{a}_{ih}$, we construct estimates of L_{jht} as

$$\widehat{L}_{jht} = \sum_{i=1}^I \omega_h \exp(\widehat{\zeta}_{ih}) \mathbf{1}\{J(i, t) = j \wedge H(i, t) = h\}, \quad (13)$$

for $j = 1, \dots, J$, $h = 1, \dots, H$, and $t = 1, \dots, T$, where $\mathbf{1}\{\cdot\}$ is the indicator function and ω_h (re-)normalizes the grand mean of ability in each occupation to unity. Estimation errors on individual level ability are likely to be averaged out in this aggregation to firm level labor inputs. Estimates of piece rate wages are constructed in a similar manner:

$$\widehat{\phi}_{jht} = \omega_h^{-1} \exp(\widehat{\zeta}_{jht}). \quad (14)$$

4.1.1 Results

In the first step, individual-occupation and firm-year-occupation effects are obtained from the individual level log wage equation (12), for representativeness estimated on the full individual level panel. This includes workers employed by firms for which we do not observe accounting data. These firms are therefore not included in the firm

¹⁹We exploit the sparsity of the design matrix to circumvent the dimensionality problem and solve the system of normal equations associated with (12) using a sparse matrix conjugate gradient algorithm.

level panel, and so are not used in the second and third step of the our estimation procedure. We use firms without accounting data in this first step estimation to improve the precision of the estimated worker effects, and hence also of the firm effects used in the second and third step. However, for purposes of comparison and interpretation, results reported in the following on distributions and decompositions relating to the individual level wage regression pertain to the subsample of firms included in the firm level panel. The corresponding results for the full individual level panel are available on request.

Figure 3 depicts the distribution of piece rate wages (14) across firm-years, by occupation and industry. Piece rate wages, like firm level wages, exhibit considerable dispersion across firm-years. From Figure 3, Managers are paid higher and more dispersed wages per ability unit than other occupations, followed by Salaried workers, in all four industries. The very fact that there is dispersion in piece rate wages, i.e., in unit prices of human capital, represents a violation of the “Law of One Price,” and so, not surprisingly, labor input quality is not the only driver of the observed joint distribution of productivity and wages; misallocation also plays a role. Mortensen (2003) cites further evidence to this effect.

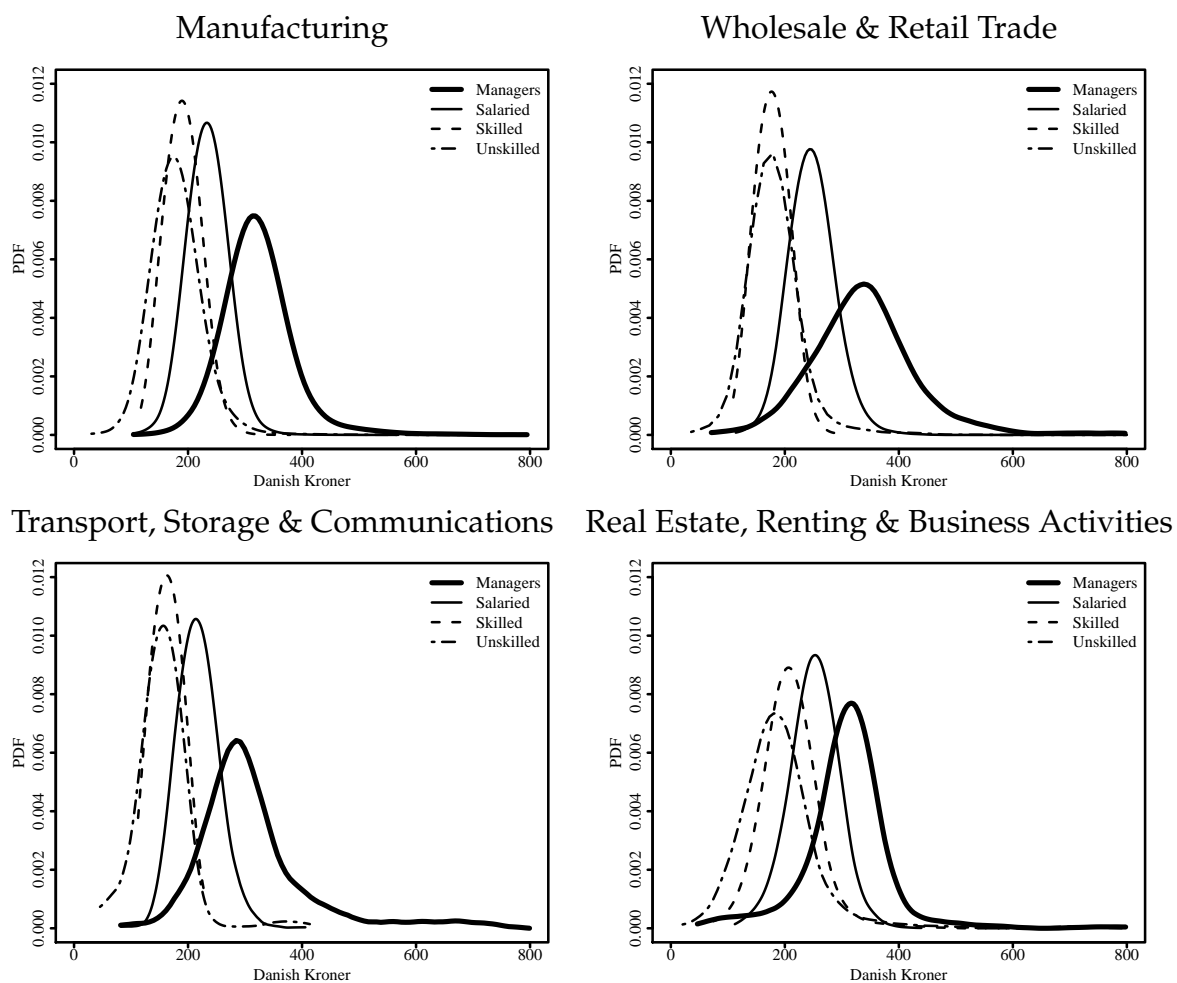
Figure 4 exhibits the distribution of ability adjusted labor input (13) across firm-years, by occupation and industry. This production input also varies considerably across firm-years. From the figure, the Manufacturing industry employs relatively more Skilled labor and Transport, Storage & Communications relatively less Management, compared to other industries. We return to the estimated individual level wage regression when we further analyze the dispersion in individual log wages in Section 5.

The individual level wage equation (12) represents a log-linear decomposition of individual wages into a firm-occupation-year specific log piece rate wage, a worker-occupation specific permanent ability, and a residual term which we may interpret as a transitory disturbance to worker ability. Hence, we obtain the following decomposition of the variance of individual level log wages:

$$\begin{aligned}
 \text{Var}(\ln w_{it}) = & \underbrace{\text{Cov}(\ln w_{it}, \ln \phi_{J(i,t)H(i,t)t})}_{\text{Log piece rate wage}} \\
 & + \underbrace{\text{Cov}(\ln w_{it}, \ln a_{iH(i,t)})}_{\text{Worker log ability}} + \underbrace{\text{Cov}(\ln w_{it}, \varepsilon_{iH(i,t)t}^a)}_{\text{Residual}}. \quad (15)
 \end{aligned}$$

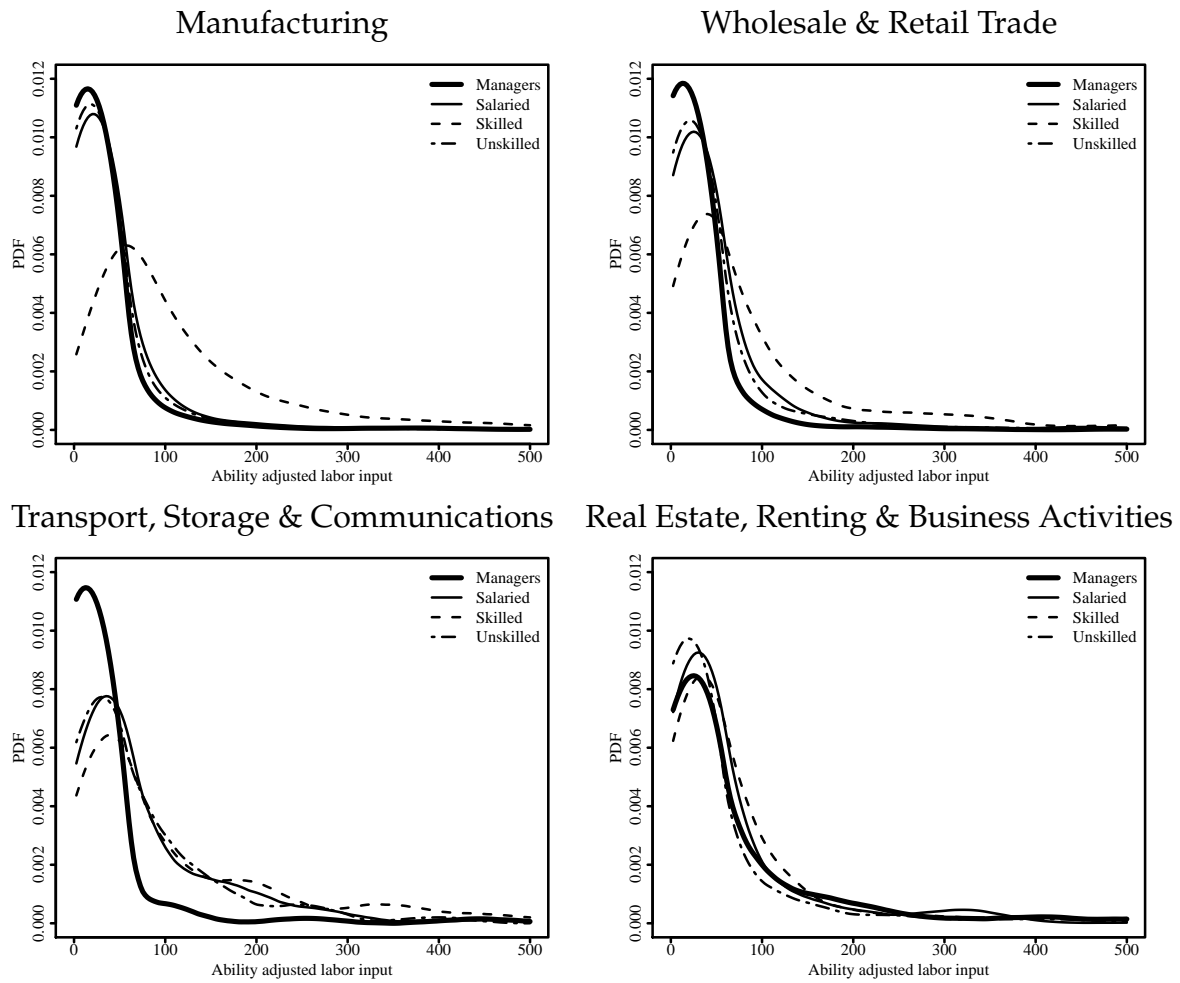
Since Abowd, Kramarz, and Margolis (1999), numerous studies have estimated log wage equations with two-sided heterogeneity and reported decompositions similar to (15). This paper contributes to this literature by placing the decomposition within a rigorous wage setting game and by allowing firm heterogeneity to vary over time.

Figure 3: Distributions of Occupation Specific Piece Rate Wages



Note: Unit of observation is a firm-year. The densities are estimated using the Epanechnikov kernel with a bandwidth of 25.

Figure 4: Distributions of Occupation Specific Ability Adjusted Labor Input



Note: Unit of observation is a firm-year. The densities regressions are estimated using the Epanechnikov kernel with a bandwidth of 25.

Table 6: Individual Level Log Wage Variance Decomposition

	Manu- facturing	Wholesale & Retail Trade	Transport, Storage & Communi- cations	Real Estate, Renting & Business Activities
Managers				
$Var(\ln w_{it})$.120	.194	.159	.155
Worker log ability	78.9%	87.1%	88.4%	80.4%
Log piece rate wage	9.2%	7.7%	3.9%	9.3%
Residual	11.9%	5.1%	7.7%	10.4%
Salaried workers				
$Var(\ln w_{it})$.077	.124	.098	.134
Worker log ability	79.6%	83.9%	79.7%	75.9%
Log piece rate wage	8.3%	7.8%	6.6%	16.9%
Residual	12.1%	8.2%	13.7%	7.2%
Skilled workers				
$Var(\ln w_{it})$.063	.108	.066	.240
Worker log ability	60.1%	85.7%	75.8%	78.8%
Log piece rate wage	18.4%	2.5%	10.5%	14.9%
Residual	21.5%	11.9%	13.7%	6.3%
Unskilled workers				
$Var(\ln w_{it})$.096	.127	.105	.201
Worker log ability	69.3%	84.7%	77.2%	71.1%
Log piece rate wage	20.5%	8.4%	14.6%	20.3%
Residual	10.2%	6.9%	8.2%	8.6%

Note: The decomposition is given by (15).

Our decompositions of individual level log wage dispersion are reported in Table 6. We report the component in percentage of total variance in log individual wages $Var(\ln w_{it})$. The lion's share of observed dispersion in individual log wages is explained by permanent differences in worker ability (varying between 60 and 88 percent across industries and occupations). Within industry, permanent differences in worker ability explain the greatest portion of wage dispersion among Managers.²⁰ Differences in time varying firm level log piece rate wages explain between 3 and 20 percent of the variance in individual log wages. Within industry, Unskilled workers exhibit the highest share explained by these firm-year effects, with shares above those attributable to residual or transitory ability differences. For other occupations, transitory ability may explain more than firm effects, but never more than permanent ability, and the general tendency is for the piece rate (firm-year) share to increase as we move down the education-skill hierarchy. Qualitatively, this pattern, with worker (ability) effects most important for the highest occupations and firm (piece rate) effects for the lowest, is consistent with the findings of Postel-Vinay and Robin (2002) using French data.

Looking across industries, Manufacturing and Real Estate, Renting & Business Activities have about the same portion of wage dispersion explained by piece rate (firm-year) effects, and this is higher than the corresponding portion in the other two industries. For each occupation, the portion of wage dispersion explained by permanent ability differences tends to be relatively low in Manufacturing and relatively high in Wholesale & Retail Trade.

We end the discussion of the individual level wage equation with a word of caution: The conclusions reviewed above in relation to the relative importance of worker and firm heterogeneity are derived from a structural model of wage setting in which a worker's previous labor market search history does not affect the wage earned. This assumption justifies the implication that the estimated worker effect in the wage equation reflects the worker's productivity in the firm. As Postel-Vinay and Robin (2002) point out, this assumption is invalid in a job ladder in which wages can be bid up when an employed worker finds an alternative employment opportunity. Then, as in their model, the worker fixed effect in the wage equation includes a return to search and, consequently, is likely to be biased up as a measure of the worker's productive human capital. As we shall argue further below, this issue might have implications for the quantitative results in this paper, but not the qualitative conclusions we draw.

²⁰In Manufacturing, nearly identical portions are explained for Managers and Salaried workers, about 80 percent.

4.2 Second Step: Firm Level Production Function

The second step of the estimation procedure is concerned with the estimation of the firm level production function. Our empirical analysis favors a production function specification with elasticity of substitution of unity between occupations, i.e., a Cobb-Douglas labor aggregator. Although at times we refer to specifications involving the less restrictive CES aggregator, we therefore present the main estimation equations imposing the Cobb-Douglas labor aggregator. In this case, the empirical version of the (log) production function based on (5) with L_{jht} replaced by its estimate \widehat{L}_{jht} from (13) is

$$\ln Y_{jt} = \alpha_K \ln K_{jt} + \alpha_L \sum_{h=1}^H \gamma_h \ln \widehat{L}_{jht} + p_{jt}. \quad (16)$$

Difficulties in estimating production function parameters are well known: Firm-level inputs K_{jt} and L_{jt} are likely to be chosen by the firm in response to realizations of TFP p_{jt} . Dealing with this issue requires imposing some structure. We consider TFP processes on the following form:

Assumption A2 *The TFP process is $p_{jt} = p_j + v_{jt}$ where $v_{jt} = \eta v_{jt-1} + \varepsilon_{jt}^p$, with $E[\varepsilon_{jt}^p] = 0$ for all j, t and $E[\varepsilon_{jt}^p \varepsilon_{js}^p] = 0$ for all j and $s \neq t$.*

This allows for both a time-invariant firm fixed effect p_j and an $AR(1)$ component that may be more or less persistent, depending on the value of η . The TFP innovations ε_{jt}^p are idiosyncratic, and if $\eta = 0$ they are purely transitory.

With respect to the timing of events in the model, we adhere to the following structure (consistent with the dynamic environment described in Section 2): The TFP innovation ε_{jt}^p is realized at the beginning of period t . Following this, labor reallocates across firms, and investments in capital are made. Given the set of (potential) workers available after labor reallocation, the firm now negotiates wage contracts in accordance with the procedure outlined in Section 2. Finally, production takes place and payments are made. Hence, from an econometric point of view, labor and capital are both potentially endogenous, i.e., correlated with contemporaneous TFP innovations. We did consider variations with capital predetermined, i.e., orthogonal to contemporaneous TFP innovations, and our results did not change much, but our empirical analysis nonetheless favored specifications with both capital and labor endogenous.

Besides the timing assumption, we do not impose further restrictions on labor reallocations. In particular, we do not assume that firms attain the optimal labor input combination: Although they condition hiring and separation decisions on current TFP innovations, labor market frictions imply that reallocation is imperfect and sluggish.

Assumption A2 allows quasi-differencing the production function to express the

unobserved TFP innovations as functions of data and unknown parameters,

$$\begin{aligned} \varepsilon_{jt}^p(\theta, p_j) = & \ln Y_{jt} - \eta \ln Y_{jt-1} - \alpha_K \ln K_{jt} + \eta \alpha_K \ln K_{jt-1} \\ & - \alpha_L \sum_{h=1}^H \gamma_h \ln \widehat{L}_{jht} + \eta \alpha_L \sum_{h=1}^H \gamma_h \ln \widehat{L}_{jht-1} - (1 - \eta)p_j, \end{aligned} \quad (17)$$

where $\theta = (\alpha_K, \alpha_L, \gamma_1, \dots, \gamma_H, \eta)'$.

Using (17) and the assumed timing of events we can construct orthogonality conditions that can be used to identify and estimate the production function parameters via GMM. Indeed, restricting $p_j = p$, we have $E \left[\varepsilon_{jt}^p(\theta, p) \mathbf{Z}_{jt} \right] = \mathbf{0}$, where \mathbf{Z}_{jt} is a vector of instrumental variables in the information set at $t - 1$. If instead all $p_j, j = 1, \dots, J$, are left free, the standard procedure would be to difference (17) once more and consider conditions of the form $E \left[\Delta \varepsilon_{jt}^p(\theta) \mathbf{Z}_{jt} \right] = \mathbf{0}$, where $\Delta \varepsilon_{jt}^p(\theta, p_j) = \varepsilon_{jt}^p - \varepsilon_{jt-1}^p = \Delta \varepsilon_{jt}^p(\theta)$ depends on data and θ , only, and \mathbf{Z}_{jt} in this case is a vector of instrumental variables in the information set at $t - 2$. The levels version nevertheless remains a relevant specification because $AR(1)$ persistence ($\eta > 0$) with separate TFP innovations ε_{jt}^p across firms allows capturing much the same empirical features in the data as separate fixed effects p_j . The approach of quasi-differencing to isolate exogenous variation follows the dynamic panel data literature, see, e.g., Holtz-Eakin, Newey, and Rosen (1988), Arellano and Bond (1991) and Blundell and Bond (1998). Olley and Pakes (1996), Levinsohn and Petrin (2003), Akerberg, Caves, and Frazer (2006) and Gandhi, Navarro, and Rivers (2011) present alternative ways of dealing with endogenous inputs in production functions.

In our empirical work, we “collapse” the system of orthogonality conditions by averaging across firms and time. Thus, our implementation of the conditions $E \left[\varepsilon_{jt}^p(\theta, p) \mathbf{Z}_{jt} \right] = \mathbf{0}$ reads

$$\frac{1}{M - J} \sum_{j=1}^J \sum_{t=2}^{T_j} \varepsilon_{jt}^p(\theta, p) \mathbf{Z}_{jt} = \mathbf{0}, \quad (18)$$

with $M = \sum_{j=1}^J T_j$ and T_j the number of time periods available for firm j .²¹ Averaging only across firms would yield one such vector of moment restrictions for each time period and thus a large number of over-identifying restrictions, a strategy we do not pursue (see Roodman (2009) on problems with instrumental variables proliferation under such an approach).

We consider instrumentation by lagged endogenous variables in levels. Both labor and capital are treated as endogenous, and when $p_j = p$ is imposed, \mathbf{Z}_{jt} contains functions of $\ln K_{j,t-1}$, $\ln L_{j,t-1}$, and $\ln Y_{j,t-1}$. Notice that these instruments are valid even if capital and labor inputs are in fact predetermined. We use the minimum valid

²¹To avoid clutter, it is suppressed in this notation that the panel is unbalanced, with gaps in many firm level time series.

lag order only in order to maximize both the size of the data set used, and the power of the instruments. When p_j is left free and the moment conditions involve $\Delta \varepsilon_{jt}^p(\theta)$, one observation per firm is lost, and we construct \mathbf{Z}_{jt} by lagging the endogenous variables twice. The exact elements in the vector of instrumental variables are specified when we present the empirical results below.

4.2.1 Results

For purposes of testing down from a more general specification in our empirical work, we consider a constant elasticity of substitution (CES) labor aggregator, $(\sum_{h=1}^H \gamma_h \widehat{L}_{jht}^\rho)^{1/\rho}$, with elasticity of substitution between occupations of $\epsilon = 1/(1 - \rho)$. Thus, our most general production function specification is

$$\ln Y_{jt} = \alpha_K \ln K_{jt} + \frac{\alpha_L}{\rho} \ln \left(\sum_{h=1}^H \gamma_h \widehat{L}_{jht}^\rho \right) + p_{jt}, \quad (19)$$

with (17) augmented accordingly. The log linear, or Cobb-Douglas, labor aggregator special case corresponds to the testable null of unit elasticity, $\epsilon = 1$ (i.e., $\rho = 0$). The CES labor aggregator is consistent with the theoretical framework developed in Section 2, but does not in general deliver a simple closed form solution for the piece rate wage equation, unless workers' bargaining power is the same across occupations.

Least squares estimation We start with a brief presentation of estimates obtained from least squares estimation of the production function. This naive empirical strategy treats capital and labor inputs as orthogonal to the TFP innovation ε_{jt}^p , and also imposes $p_{jt} = p + \varepsilon_{jt}^p$, i.e., $\eta = 0$. It is included here for comparison purposes.

With CES labor aggregator, the production function (19) is nonlinear in parameters (in particular, ρ) and is estimated by Nonlinear Least Squares (NLS). With Cobb-Douglas labor aggregator, it is given as (16), amenable to estimation by ordinary (linear) regression (OLS). Estimates pertaining to both specifications are presented in Table 7. Here, we leave the returns to scale in capital and aggregate labor input unrestricted, i.e. $\alpha_K + \alpha_L$ need not be unity.

In the bottom panel of Table 7 we report the p -values from F -tests of the restrictions $\epsilon = 1$ (p_{CD}), $\alpha_K + \alpha_L = 1$ (p_{CRTS}), and the joint hypothesis $\epsilon = 1$ and $\alpha_K + \alpha_L = 1$ ($p_{CD+CRTS}$). The tests involving the restriction $\epsilon = 1$ are only relevant for the specification with a CES labor aggregator. Empirically, we find that ϵ often approaches infinity, preventing us from computing the tests on ϵ in all industries except Wholesale & Retail Trade. In this case we cannot reject $\epsilon = 1$ and $\alpha_K + \alpha_L = 1$. Moreover, at a 5% significance level we only reject a constant returns to scale technology in Real Estate, Renting & Business Activities, where nonetheless $\alpha_K + \alpha_L = .94$. Numerically,

we found that the NLS objective functions were rather flat in ρ , and that α_K and α_L were quite insensitive to this.

Clearly, ϵ is not well identified in this NLS specification, and we focus attention on the columns imposing $\epsilon = 1$ (estimated and labeled by OLS in Table 7). Overall, the parameter estimates appear reasonable. Across industries, Wholesale & Retail Trade gets the lowest point estimate of α_K , the capital input elasticity, at .08, with the highest, at .16, in Transport, Storage & Communications, consistent with these industries having the lowest respectively the highest correlation between labor productivity and capital intensity, see Table 5. The capital and overall labor input elasticity estimates α_K and α_L are relatively precisely estimated and sum to nearly unity in each industry, indicating that the data are consistent with the constant returns to scale restriction (the lowest p -value for this, again in Real Estate, Renting & Business Activities, is 4%). Throughout, the lowest of the four occupation-specific relative productivities is that of Unskilled workers, γ_4 . In Real Estate, Renting & Business Activities, the γ_h s are ordered according to the standard education-skill hierarchy, with Managers highest, whereas in the other three industries, this ordering is reversed among the three other occupations besides the Unskilled. In the unrestricted NLS estimation, the standard ordering is obtained in all industries, except that the two highest occupations are switched in Manufacturing and the two lowest in Wholesale & Retail Trade. This suggests that the reduction to Cobb-Douglas labor aggregator is not without issues, after all, based on the least squares estimation.

GMM estimation We now turn attention to results of GMM estimation of the production function parameters, relaxing the restrictions behind the NLS and OLS estimation to allow for potentially endogenous capital and labor choices on the part of firms, and allowing for $\eta \in (-1, 1)$, i.e., potentially persistent firm-year effects in the TFP processes. We restrict $p_j = p$ for all $j = 1, 2, \dots, J$. Hence, our empirical orthogonality conditions are exactly on the form (18). Notice that the absence of time-invariant firm fixed effects in the production function is perfectly consistent with an individual level wage equation with time-invariant worker fixed effects and firm fixed effects that are time-varying (in our model, due to firm-year effects in TFP, and changing input composition).²²

Consider first the general case with a CES labor aggregator, and where the returns to scale in capital and aggregated labor input is unrestricted. That is, the production function is given by (19). Recall that $\sum_{h=1}^H \gamma_h = 1$. Hence, in the CES case there are $H + 3$ parameters to be estimated in $\theta = (\alpha_K, \alpha_L, \gamma_1, \dots, \gamma_H, \eta, \rho)'$. With four occupations in

²²We also compared to estimates treating p_j as free, replacing ϵ_{jt}^p by $\Delta \epsilon_{jt}^p$ in the orthogonality conditions (18) and modifying \mathbf{Z}_{jt} accordingly. In these cases, we obtained much less stable results, inflated standard errors, and unrealistically low estimated returns to scale in production.

Table 7: Production Function Parameters–NLS and OLS Estimation

	Manu- facturing		Wholesale & Retail Trade		Transport, Storage & Communi- cations		Real Estate, Renting & Business Activities	
	NLS	OLS	NLS	OLS	NLS	OLS	NLS	OLS
α_K	.102 (.007)	.115 (.008)	.066 (.012)	.082 (.013)	.152 (.028)	.155 (.026)	.124 (.015)	.113 (.018)
α_L	.910 (.010)	.885 (.013)	.928 (.020)	.907 (.022)	.886 (.059)	.875 (.055)	.820 (.024)	.827 (.031)
γ_1	.309 (.019)	.178 (.012)	.368 (.035)	.208 (.021)	.511 (.085)	.170 (.055)	.408 (.027)	.374 (.030)
γ_2	.346 (0.016)	.239 (.012)	.300 (.022)	.291 (.020)	.202 (.049)	.290 (.056)	.260 (.029)	.274 (.036)
γ_3	.177 (.006)	.453 (.014)	.154 (.017)	.341 (.019)	.193 (.039)	.428 (.046)	.197 (.024)	.197 (.030)
γ_4	.168 (.007)	.131 (.008)	.179 (.016)	.159 (.016)	.093 (.031)	.111 (.040)	.134 (.015)	.156 (.026)
ϵ	∞	1.000	12.686 (13.592)	1.000	∞	1.000	∞	1.000
# observations	9,962	9,962	2,332	2,332	408	408	1,288	1,288
p_{CD}	—	—	.390	—	—	—	—	—
p_{CRTS}	.074	.966	.682	.496	.360	.438	.006	.040
$p_{CD+CRTS}$	—	—	.674	—	—	—	—	—

Note: The constant term is not reported.

the data, we therefore estimate eight parameters in (θ, p) .

We employ an over-identified system of orthogonality conditions, treating both labor and capital as endogenous. We use lagged endogenous variables as instruments, but do not use lagging as a source of over-identification. Instead, we use 1 , $\ln L_{jt-1}$, $\ln K_{jt-1}$, $\ln Y_{jt-1}$, $\prod_{h=1}^H \ln L_{jht-1}$, $\ln L_{jt-1} \cdot \ln Y_{jt-1}$, and $\ln K_{jt-1} \cdot \ln Y_{jt-1}$ as instruments. With $H = 4$, this yields thirteen orthogonality conditions, and hence five over-identifying restrictions. We consequently use a two-step GMM estimator, where the first step is used for estimation of the optimal weight matrix, and report Hansen's J -test of validity of the over-identifying restrictions (Hansen, 1982).

Results appear in Table 8, columns (1), (3), (5) and (7) (first of two columns for each industry). Parameter estimates are sensible, with capital input elasticities α_K ranging from .05 to .18, again with Wholesale & Retail Trade lowest and Transport, Storage & Communications highest. In Manufacturing, the capital input elasticity is estimated to .10. The GMM estimates of α_K are slightly below the naive NLS estimates of Table 7, except that for Transport, Storage & Communications. Taken together, the unrestricted capital and aggregate labor input elasticities α_K and α_L indicate that the technology for each industry is not far from constant returns to scale. The GMM estimates of the relative productivities of the different occupations, γ_h , $h = 1, \dots, H$, exhibit a clear pattern: Managerial labor is more productive than Salaried labor, Salaried labor is more productive than Skilled labor, etc. As in the NLS estimation, there are only two exceptions: Again, Managers and Salaried workers are switched in Manufacturing, and now the Unskilled are not the least productive in Real Estate, Renting & Business Activities. Point estimates of the elasticity of substitution between any two occupational labor inputs, $\epsilon = 1/(1 - \rho)$, are very imprecise and therefore vary considerably across industries. Finally, the $AR(1)$ coefficient η from the firm level TFP processes is directly estimated in the GMM procedure, and the estimates imply strong persistence, but not unit roots, with η ranging from .65 in Transport, Storage & Communications to .81 in Real Estate, Renting & Business Activities.

The lower panel of Table 8 report p -values for Hansen's J -test (p_J), and Wald tests of the Cobb-Douglas labor aggregator restriction $\epsilon = 1$ (p_{CD}), the constant returns to scale restriction $\alpha_K + \alpha_L = 1$ (p_{CRTS}), and the joint Cobb-Douglas and CRTS hypothesis, $\epsilon = 1$ and $\alpha_K + \alpha_L = 1$ ($p_{CD+CRTS}$). At a 5% level, we cannot reject any of the restrictions (validity of over-identifying restrictions, CRTS or Cobb-Douglas labor aggregator) in any of the industries. This is in part because the elasticity of substitution in the CES labor aggregator is poorly identified in our data, a finding that is consistent with our naive least squares estimation in Table 7.

Given the outcomes of the tests on the general model, we proceed to estimate a production function imposing constant returns to scale and a Cobb-Douglas labor aggregator on a just-identified system of orthogonality conditions. We still treat both

labor and capital as endogenous. Since ρ is no longer estimated, there are now seven parameters in (θ, p) . The seven instruments used are $1, \ln L_{jt-1}, \ln K_{jt-1},$ and $\ln Y_{jt-1}$. This is our preferred specification, in part because it represents the most parsimonious use of parameters and instrumental variables consistent with the data.²³

Results are reported in Table 8, columns (2), (4), (6) and (8). In terms of the capital input elasticities α_K , the restrictions imposed lead to a strengthening of the observed pattern, with the highest coefficient, again that in Transport, Storage & Communications, now even higher, at .20, and the lowest, still in Wholesale & Retail Trade, considerably lower, now at .01. The estimate in Manufacturing is nearly unchanged, at .09. As in the least squares case, the relative productivities implied by the Cobb-Douglas labor aggregator are not as clearly ranked as in the more general model, except that Skilled labor gets higher coefficients than Unskilled, $\gamma_3 > \gamma_4$, throughout. Indeed, it also gets higher coefficients than Salaried labor, $\gamma_3 > \gamma_2$, i.e., a reversal among these two types, compared to the expected outcome. Manufacturing and Wholesale & Retail Trade exhibit the same pattern as in the OLS estimation, with the Unskilled least productive and the three higher occupations in reverse order, with Skilled labor most productive. In Transport, Storage & Communications, Skilled labor is similarly the most productive type. The high estimated relative productivity per ability unit of Skilled labor may be related to the fact that this category by the data definitions in the Danish case consists of relatively highly specialized blue collar workers, whereas Salaried labor consists of white color office workers in non-managerial positions, cf. Christensen, Lentz, Mortensen, Neumann, and Werwatz (2005), and the reversal similarly matches our findings below on the outside options by occupation.

With Cobb-Douglas labor aggregator and constant returns to scale imposed, the estimated TFP process exhibits stronger persistence, with autocorrelation coefficients ranging between .75 and .82, but still no unit root. The strong persistence implies that some of the same empirical phenomena are picked up as if using permanent firm effects (i.e., p_j free).

Based on the tests carried out on the more general model and the fact that the parameter estimates appear reasonable, this production function, with constant returns to scale and Cobb-Douglas labor aggregator, estimated on a just-identified system of orthogonality conditions, constitutes our preferred specification. The remainder of the analysis in this paper is carried out using these estimates (i.e. columns (2), (4), (6) and (8) in Table 8). Note that the Cobb-Douglas case offers the additional benefit of a closed form solutions for the firm wage equation, even when workers' bargaining powers differ by occupation.

²³In our specification search we also find that we get the most stable results with respect to both point estimates and estimated standard errors using exactly identified systems.

Table 8: Production Function Parameters–GMM Estimation

	Manu- facturing		Wholesale & Retail Trade		Transport, Storage & Communi- cations		Real Estate, Renting & Business Activities	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
α_K	.096 (.013)	.089 (.015)	.053 (.030)	.006 (.039)	.175 (.067)	.198 (.077)	.120 (.034)	.054 (.066)
α_L	.924 (.022)	.911 (.015)	.949 (.056)	.994 (.039)	.870 (.181)	.802 (.077)	.848 (.067)	.946 (.066)
γ_1	.275 (.069)	.144 (.052)	.334 (.188)	.066 (.167)	.424 (.706)	.016 (.605)	.440 (.094)	.417 (.605 p)
γ_2	.426 (.077)	.354 (.054)	.324 (.076)	.412 (.122)	.286 (.304)	.014 (1.001)	.168 (.200)	.130 (1.453)
γ_3	.165 (.075)	.433 (.036)	.255 (.200)	.486 (.089)	.249 (.313)	.822 (.394)	.155 (.105)	.281 (.484)
γ_4	.134 (.047)	.068 (.040)	.087 (.065)	.035 (.086)	.040 (.183)	.148 (.512)	.237 (.088)	.172 (.382)
ϵ	54.778 (1375.2)	1.000	1.886 (3.452)	1.000	10.886 (170.07)	1.000	164,154 (1.42×10^{10})	1.000
η	.725 (.015)	.753 (.012)	.740 (.022)	.771 (.025)	.649 (.053)	.744 (.152)	.812 (.037)	.823 (.066)
# observations	8,521	8,521	1,907	1,907	335	335	1,058	1,058
p_J	.060		.168		.176		.055	
p_{CD}	.969		.797		.954		1.000	
p_{CRTS}	.121		.941		.714		.516	
$p_{CD+CRTS}$.212		.901		.772		.804	

Note: Constant term not reported. Columns (1), (3), (5) and (7) contain estimates obtained under CES labor aggregator and unrestricted returns to scale in capital and aggregated labor. Columns (2), (4), (6) and (8) contain estimates obtained under the more parsimonious specification with a Cobb-Douglas labor aggregator and constant returns to scale in capital and aggregated labor. The reported standard errors are not corrected for the first step estimation of individual abilities.

4.3 Third Step: Firm Level Piece Rate Wage Functions

The third step of our procedure is the estimation of the firm-level occupation-specific piece rate wage equations. These apply to piece rate wages and ability adjusted labor inputs that are not directly observed in the data, but inferred from the individual level wage equation as outlined above (see (13) and (14)). To break the deterministic link between piece rate wages and marginal productivity stipulated by the model, thus making the model empirically implementable, we amend each piece rate wage function with an error-term ε_{jht}^ϕ . The piece rate wage equations based on (7) then read

$$\phi_{jht} = (1 - \beta_h)b_h + \left(\frac{\alpha_L \gamma_h \frac{\beta_h}{1 - \beta_h}}{1 + \sum_{k=1}^H \alpha_L \gamma_k \frac{\beta_k}{1 - \beta_k}} \right) \frac{Y_{jt}}{L_{jht}} + \varepsilon_{jht}^\phi \quad (20)$$

for $h = 1, 2, \dots, H$, in the Cobb-Douglas labor aggregator case. Thus, the following assumption on $(\varepsilon_{j1t}^\phi, \dots, \varepsilon_{jHt}^\phi)$ is of interest:

Assumption A3 $E \left[\varepsilon_{jht}^\phi \mid \frac{Y_{jt}}{L_{j1t}}, \dots, \frac{Y_{jt}}{L_{jHt}} \right] = 0$, for all $j = 1, \dots, J$, $h = 1, \dots, H$, and $t = 1, \dots, T$.

Under Assumption A3, ε_{jht}^ϕ may be interpreted as a tremble to the bargaining process. Alternatively, it may be regarded as a reduced form residual, containing variation in piece rate wages that are left unexplained by our model. In either case, as we infer the unobserved piece rate wages ϕ_{jht} from the individual level wage equation (see (14)), we may think of ε_{jht}^ϕ as containing the estimation errors resulting from this association, as well.

Given estimates of ability adjusted labor inputs, we may define $\hat{q}_{jht} \equiv \frac{Y_{jt}}{L_{jht}}$ for the purpose of estimation of the piece rate wage equations. The resulting system is linear in levels (see (20)),

$$\phi_{jht} = \psi_{0h} + \psi_{1h} \hat{q}_{jht} + \varepsilon_{jht}^\phi, \quad h = 1, \dots, H, \quad (21)$$

and may under Assumption A3 be considered as a seemingly unrelated regressions system (Zellner, 1962). Our structural model implies specific structure on the reduced form parameters $(\psi_{01}, \dots, \psi_{1H})$ in terms of the structural parameters, namely, the outside options b_1, \dots, b_H and the bargaining power parameters β_1, \dots, β_H . Thus, upon estimation of the $2H$ reduced form parameters, the structural parameters may be recovered as

$$\beta_h = \frac{\psi_{1h}}{\psi_{1h} + \alpha_L \gamma_h (1 - \sum_{k=1}^H \psi_{1k})} \quad \text{and} \quad b_h = \frac{\psi_{0h}}{1 - \beta_h}, \quad (22)$$

for $h = 1, \dots, H$, given estimates of the production function parameters $(\alpha_L, \gamma_1, \dots, \gamma_H)$ from the second step.

To assess the uncertainty about the structural parameter estimates, we might in principle calculate standard errors using the delta-method, thus accounting for error stemming from ε_{jht}^ϕ (trembles, sampling error, estimation error in piece rates) in the estimation of the reduced form, and transformation from reduced form to structural parameters. Such an assessment would require a further fitted regressor adjustment, along the lines of Pagan (1984), since the ability adjusted labor input vectors $\widehat{\mathbf{L}}_{jt}$ and hence \widehat{q}_{jht} in (21) are estimates from the first step analysis. However, because these estimates depend non-linearly on coefficients from the individual level wage equation, and because the estimates from the second step of $(\alpha_L, \gamma_1, \dots, \gamma_H)$ in (22) depend on $\widehat{\mathbf{L}}_{jt}$, too, the exact form of the adjustment would be complicated. Further, any errors-in-variable problem and resulting violation of Assumption A3 would induce attenuation bias toward zero in ψ_{1h} , and thus potentially β_h .

All in all, uncertainty in \widehat{q}_{jht} is likely to dominate that stemming from ε_{jht}^ϕ and estimation of $(\alpha_L, \gamma_1, \dots, \gamma_H)$ since we really use population data, so we consider instead reverse regression,

$$\widehat{q}_{jht} = \vartheta_{0h} + \vartheta_{1h}\phi_{jht} + \varepsilon_{jht}^q, \quad h = 1, \dots, H,$$

following, e.g., Friedman and Schwartz (1982). Reverse regression estimates of ψ_{0h} and ψ_{1h} are then given by $\widehat{\psi}_{0h}^r = -\frac{\widehat{\vartheta}_{0h}}{\widehat{\vartheta}_{1h}}$ and $\widehat{\psi}_{1h}^r = \frac{1}{\widehat{\vartheta}_{1h}}$. The true values of the reduced form parameters belong to the intervals within the Frisch bounds given by

$$\psi_{0h} \in \left(\psi_{0h} - \psi_{1h}^{-1} E[q_{jht}] \frac{\text{Var}(\varepsilon_{jht}^\phi)}{\text{Var}(q_{jht})}, \psi_{0h} + \psi_{1h} E[q_{jht}] \frac{\text{Var}(\widehat{\varepsilon}_{jht}^\phi)}{\text{Var}(q_{jht}) + \text{Var}(\widehat{\varepsilon}_{jht}^\phi)} \right),$$

and

$$\psi_{1h} \in \left(\psi_{1h} - \psi_{1h} \frac{\text{Var}(u_{jht})}{\text{Var}(\widehat{q}_{jht})}, \psi_{1h} + \psi_{1h}^{-1} \frac{\text{Var}(\varepsilon_{jht}^\phi)}{\text{Var}(q_{jht})} \right),$$

for $h = 1, \dots, H$. The lower and upper bounds on ψ_{0h} are consistently estimated by $\widehat{\psi}_{0h}^r$ and $\widehat{\psi}_{0h}^d$, with r and d indicating reverse respectively direct regression, and the bounds on ψ_{1h} by $\widehat{\psi}_{1h}^d$ and $\widehat{\psi}_{1h}^r$.²⁴ We report bounds on structural parameters by backing them out at the reduced form bounds using the relations (22) with interval midpoints serving as point estimates.

²⁴This form of the bounds is for error-ridden measurements $\widehat{q}_{jht} = q_{jht} + u_{jht}$ where q_{jht} is the true variable, the measurement error u_{jht} is uncorrelated with q_{jht} and ε_{jht}^ϕ , and all variables are treated as uncorrelated across equations h for simplicity. The consistency result extends to more general forms of the bounds than those exhibited, e.g., for correlation across equations, Leamer (1987), although more complicated cases require modified estimators, e.g., if u_{jht} and ε_{jht}^ϕ are correlated, Krasker and Pratt (1986).

4.3.1 Results

Estimates of the structural parameters of the piece rate wage functions are reported in Table 9. Generally, the reported ranges for the bargaining powers β_h are compatible with those reported in other relevant studies (e.g., Cahuc, Postel-Vinay, and Robin, 2006). By the point estimates, the received patterns are very clear. Within each industry, the outside options (per unit of ability adjusted labor supply) b_h are ordered according to the usual education-skill hierarchy, except that they are slightly better for Skilled than for Salaried workers. This is consistent with our results on the relative productivities in the production function estimation above, and the findings by Christensen, Lentz, Mortensen, Neumann, and Werwatz (2005) that Skilled workers face lower search costs than Salaried workers, possibly because these relatively highly specialized blue collar workers participate in well connected occupational networks. In fact, from Table 9, in the two largest industries, Manufacturing and Wholesale & Retail Trade, the outside options of Salaried workers are hardly any better than those of the Unskilled, and the bargaining power β_2 of Salaried workers is indeed the lowest across occupations—point estimates at .22 and .13, respectively. The two other industries have fewer but larger firms, and here, Managers and Salaried workers have higher bargaining power than the other groups, possibly in excess of .5 for both occupations within Transport, Storage & Communications, based on the bounds.

Comparing across industries, workers in Manufacturing have bargaining powers between those of workers in the same occupation in other industries, the exception again being Skilled labor, with β_3 considerably higher in Manufacturing than in other industries. In terms of outside options, those in Wholesale & Retail Trade dominate those in Manufacturing, and also those in Transport, Storage & Communications, except that here, Managers find the best outside options across industries. The last industry, Real Estate, Renting & Business Activities, offers very favorable outside options, the best of all for Skilled and Salaried workers, and the second-best for Managers. Again, this is the pattern by point estimates. The bounds do have bite, though: Throughout, they rule out that workers have all the bargaining power ($\beta_h = 1$), and in most cases, they also rule out zero bargaining power or outside option. They also rule out, e.g., that blue collar (Skilled and Unskilled) workers in Transport, Storage & Communications have as high bargaining power as Managers—intervals do not overlap. Less formal comparisons reveal that these blue collar workers by the Frisch bounds cannot be facing as favorable outside options as those indicated by the point estimate for Managers in the same industry. The same is true in Real Estate, Renting & Business Activities, whereas the bounds contain all the point estimates when looking across industries within each occupation, Skilled workers in Real Estate, Renting & Business Activities being the only exception. In this sense, market opportunities are more simi-

Table 9: Piece Rate Wage Function Parameters

	Manu- facturing	Wholesale & Retail Trade	Transport, Storage & Communi- cations	Real Estate, Renting & Business Activities
Managers				
Bargaining power β_1	.308 [.019,.596]	.373 [.047,.699]	.608 [.356,.861]	.093 [.007,.180]
Outside option b_1	158.4 [0.0,316.8]	171.2 [0.0,342.4]	214.3 [0.0,428.7]	208.6 [115.2,302.0]
Salaried workers				
Bargaining power β_2	.219 [.002,.435]	.131 [.001,.261]	.485 [.063,.906]	.241 [.015,.466]
Outside option b_2	116.0 [0.0,232.0]	128.6 [10.8,246.4]	115.1 [0.0,230.2]	168.7 [86.9,250.6]
Skilled workers				
Bargaining power β_3	.363 [.006,.720]	.139 [.006,.272]	.051 [.003,.100]	.071 [.012,.130]
Outside option b_3	120.7 [51.2,190.3]	132.3 [88.9,175.7]	128.2 [95.6,160.8]	173.5 [143.6,203.5]
Unskilled workers				
Bargaining power β_4	.287 [.015,.560]	.365 [.117,.613]	.141 [.000,.281]	.070 [.010,.133]
Outside option b_4	115.7 [53.5,177.8]	127.2 [67.1,187.2]	77.6 [0.0,155.2]	107.7 [32.4,183.1]

Note: The reported estimates are conditional on the production function parameter estimates reported in columns (2), (4), (6) and (8) in Table 8. Numbers presented in square brackets are bounds obtained using reverse regression techniques.

lar within occupation than within industry, especially for higher occupations.

5 Dispersion and Misallocation

Observed (i.e. non-ability adjusted) average labor productivity and firm-level average wages vary considerably and are strongly positively correlated between firms, within industries. See Figures 1 and 2, and Tables 4 and 5 for documentation. These observations may reflect either labor input heterogeneity, or factor misallocation, i.e. dispersion in marginal products across firms.²⁵ Our estimated model gives rise to a novel set of measurements on this issue. We first consider the distribution of average labor productivity, and then wages.

If dispersion in observed average labor productivity reflects misallocation of pro-

²⁵We remind the reader that throughout this paper we measure marginal products in value terms. We focus our analysis on labor inputs as we have detailed measures of labor input quality (individual ability and occupation) and on labor compensation. Such measurements are not available for capital input.

duction factors, including labor, there is scope for output enhancing (labor) reallocation. This motivates our analysis of the distribution of average labor productivity. Modern labor markets are characterized by large flows of workers between labor market states and jobs, the Danish labor market being no exception [REFERENCE HERE?]. However, labor market flows are necessary, but not sufficient, to realize efficiency gains through reallocation. To enhance efficiency, flows must direct workers towards firms with high marginal productivity. Labor flows and the resulting reallocation is a product of workers' job search, which, at least to a first approximation, is motivated by aspirations for higher wages. Hence, labor flows are more likely to be efficiency enhancing if wages, or more specifically in the context of our model, the firm-specific price of human capital, is closely linked to the firm's marginal product of labor. This motivates our analysis of firm-level wages.

5.1 Productivity Dispersion

It is natural to measure labor input heterogeneity at the firm level using the average labor input per worker, $\frac{L_{jt}^*}{N_{jt}}$, where

$$L_{jt}^* = \prod_{h=1}^H L_{jht}^{\gamma_h} \quad (23)$$

is aggregate ability adjusted labor input. We consider the following variance-covariance decomposition of observed log average labor productivity $\ln \frac{Y_{jt}}{N_{jt}} = \ln \frac{L_{jt}^*}{N_{jt}} + \ln \frac{Y_{jt}}{L_{jt}^*}$:

$$\begin{aligned} \text{Var} \left(\ln \frac{Y_{jt}}{N_{jt}} \right) &= \underbrace{\text{Cov} \left(\ln \frac{Y_{jt}}{N_{jt}}, \ln \frac{L_{jt}^*}{N_{jt}} \right)}_{\text{Labor input heterogeneity}} \\ &\quad + \underbrace{\text{Cov} \left(\ln \frac{Y_{jt}}{N_{jt}}, \ln \left[\frac{K_{jt}}{L_{jt}^*} \alpha_K \frac{Y_{jt}}{K_{jt}} + \sum_{h=1}^H \frac{L_{jht}}{L_{jt}^*} \alpha_L \gamma_h \frac{Y_{jt}}{L_{jht}} \right] \right)}_{\text{Misallocation}}. \end{aligned} \quad (24)$$

The first term in (24) is variation in observed log average labor productivity arising from labor input heterogeneity: *Ceteris paribus*, firms with higher per worker aggregate labor input $\frac{L_{jt}^*}{N_{jt}}$ has higher observed labor productivity, $\frac{Y_{jt}}{N_{jt}}$. The second term is the variance in observed log labor productivity that would prevail with per worker aggregate labor input equalized across firms, i.e. when there is no variation in $\frac{L_{jt}^*}{N_{jt}}$. Our estimated model features a homogenous of degree one Cobb-Douglas production function, and hence, $\frac{Y_{jt}}{L_{jt}^*} = \frac{K_{jt}}{L_{jt}^*} \alpha_K \frac{Y_{jt}}{K_{jt}} + \sum_{h=1}^H \frac{L_{jht}}{L_{jt}^*} \alpha_L \gamma_h \frac{Y_{jt}}{L_{jht}}$. Here, $\alpha_K \frac{Y_{jt}}{K_{jt}}$ and $\alpha_L \gamma_h \frac{Y_{jt}}{L_{jht}}$ are the marginal productivities of capital and occupation- h labor, respectively. An efficient allocation

of productive resources equalizes these across firms, and also equalizes relative (to total labor input L_{jt}^*) inputs, $\frac{K_{jt}}{L_{jt}^*}, \frac{L_{jht}}{L_{jt}^*}, \dots, \frac{L_{jHt}}{L_{jt}^*}$, across firms.²⁶ Hence, the second term in (24) captures the variation in observed log labor productivity that is accounted for by misallocated production factors.

Table 10 reports the variance of observed log labor productivity, $Var\left(\ln \frac{Y_{jt}}{N_{jt}}\right)$, and the percentage shares of labor input heterogeneity and factor misallocation in generating the observed variance.²⁷ From Table 10 labor input heterogeneity account for at most 26 percent of the dispersion in observed log average labor productivity. The decomposition is quantitatively similar in Manufacturing Wholesale & Retail Trade, and in Real Estate, Renting & Business Activities, where labor input heterogeneity comprises about 22, 21 and 26 percent of the variance in observed average labor productivity, respectively. In Transport, Storage & Communications and Real Estate, Renting & Business Activities, the contribution from labor input heterogeneity is small, but negative, at about -9 percent. In this industry, firms with high observed average labor productivity $\frac{Y_{jt}}{N_{jt}}$ employ a workforce with lower per worker ability $\frac{L_{jt}^*}{N_{jt}}$.²⁸ The limited role played by labor input heterogeneity implies that misallocation of production factors is the primary reason for dispersion in observed (log) average labor productivity, accounting for 78, 79, 109, and 74 percent of the variance in $\ln \frac{Y_{jt}}{N_{jt}}$ across firms. This finding is consistent with existing empirical results in (see e.g. Jorgenson, Gollop, and Fraumeni, 1987, Fox and Smeets, 2011, and Irarrazabal, Moxnes, and Ulltveit-Moe, 2010 as reviewed in our introduction) and with observations in Bartelsman, Haltiwanger, and Scarpetta (2013, footnote 5), where an alternative, more robust, labor productivity measure with some quality adjustment for labor input, and working hours, is highly correlated with their basic measure of labor productivity, suggesting a limited scope for labor heterogeneity in driving the distribution of observed labor productivity, $\frac{Y_{jt}}{N_{jt}}$.

Our preferred production function specification may be written as

$$\ln Y_{jt} = \alpha_0 + \alpha_K \ln \frac{K_{jt}}{L_{jt}^*} + \ln L_{jt}^* + v_{jt} \quad (25)$$

²⁶Let r and w_h be the competitive price of capital and occupation- h wage, respectively. Then, $K_{jt} = \alpha_K \frac{Y_{jt}}{r}$, $L_{jht} = \alpha_L \gamma_h \frac{Y_{jt}}{w_h}$, and $L_{jt}^* = \alpha_L Y_{jt} \prod_{h=1}^H \left(\frac{\gamma_h}{w_h}\right)^{\gamma_h}$. It follows that $\frac{K_{jt}}{L_{jt}^*} = \frac{\alpha_K}{\alpha_L} \prod_{h=1}^H \left(\gamma_h \frac{r}{w_h}\right)^{-\gamma_h}$ and $\frac{L_{jht}}{L_{jt}^*} = \gamma_h \prod_{k=1}^H \left(\gamma_k \frac{w_h}{w_k}\right)^{-\gamma_h}$, both of which are constant across firms.

²⁷This interpretation is strictly speaking only valid when both covariance terms in (24) are positive.

²⁸Employing higher ability workers *ceteris paribus* leads to higher labor productivity in our framework. However, there are multiple sources of productivity variation across firms (capital, labor and TFP), and we cannot impose the *ceteris paribus* condition on the decomposition (24). Hence, a negative covariance between $\frac{Y_{jt}}{N_{jt}}$ and $\frac{L_{jt}^*}{N_{jt}}$ is consistent with our framework and reflects confounding effects coming from capital inputs and TFP.

Table 10: Log Labor Productivity Variance Decomposition—Labor Input Heterogeneity versus TFP and Misallocation

	Manu- facturing	Wholesale & Retail Trade	Transport, Storage & Communi- cations	Real Estate, Renting & Business Activities
$Var \left(\ln \frac{Y_{jt}}{N_{jt}} \right)$.088	.134	.192	.178
Labor input heterogeneity	21.9%	20.7%	−9.4%	25.7%
Misallocation	78.1%	79.3%	109.4%	74.3%

Note: The decomposition is given by (24).

where L_{jt}^* is given by (23), and $v_{jt} = \eta v_{jt-1} + \varepsilon_{jt}^p$. Using (25), we now dissect the estimated (nondegenerate) distributions of log marginal products across firms. Consider first the log marginal product of capital, $\ln(\alpha_K \frac{Y_{jt}}{K_{jt}})$. The production function (25) admits the following variance decomposition:

$$\begin{aligned}
 Var \left(\ln \frac{Y_{jt}}{K_{jt}} \right) &= \underbrace{\alpha_L Cov \left(\ln \frac{Y_{jt}}{K_{jt}}, -\ln \frac{K_{jt}}{L_{jt}^*} \right)}_{\text{Capital-labor ratio}} \\
 &\quad + \underbrace{\eta Cov \left(\ln \frac{Y_{jt}}{K_{jt}}, v_{jt-1} \right)}_{\text{TFP, persistent}} + \underbrace{Cov \left(\ln \frac{Y_{jt}}{K_{jt}}, \varepsilon_{jt}^p \right)}_{\text{TFP, transitory}}. \quad (26) \\
 &\hspace{15em} \underbrace{\hspace{15em}}_{\text{TFP, total}}
 \end{aligned}$$

The first term in (26) is the variance in the marginal product of capital accounted for by variation in capital-labor ratios across firms. *Ceteris paribus*, a firm with a high capital-labor ratio, has a low marginal productivity of capital. Hence, dispersion in capital-labor ratios translates into dispersion in marginal products. The second and third terms in (26) is the variance accounted for by TFP shocks. A positive TFP shock increases productivity of all factors, including capital. If production factors do not reallocate efficiently, TFP shocks generate dispersion in marginal products. Our specification allow us to consider separately transitory TFP shocks, and their cumulative, or persistent, effect.

Consider next the log marginal product of labor in occupation h , $\ln(\alpha_L \gamma_h \frac{Y_{jt}}{L_{jt}})$. The

following variance decomposition is of interest:

$$\begin{aligned}
\text{Var} \left(\ln \frac{Y_{jt}}{L_{jht}} \right) &= \underbrace{\alpha_K \text{Cov} \left(\ln \frac{Y_{jt}}{L_{jht}}, \ln \frac{K_{jt}}{L_{jt}^*} \right)}_{\text{Capital-labor ratio}} + \underbrace{\text{Cov} \left(\ln \frac{Y_{jt}}{L_{jht}}, -\ln \frac{L_{jht}}{L_{jt}^*} \right)}_{\text{Labor composition}} \\
&\quad + \underbrace{\eta \text{Cov} \left(\ln \frac{Y_{jt}}{L_{jht}}, v_{jt-1} \right)}_{\text{TFP, persistent}} + \underbrace{\text{Cov} \left(\ln \frac{Y_{jt}}{L_{jht}}, \varepsilon_{jt}^p \right)}_{\text{TFP, transitory}}. \quad (27) \\
&\hspace{15em} \underbrace{\hspace{15em}}_{\text{TFP, total}}
\end{aligned}$$

The interpretations of the terms labeled capital-labor ratio, and TFP (persistent, transitory, and total) are analogous to similarly labeled terms in (26). With H occupations, labor composition across occupations matter. As employment is shifted towards occupation h , the marginal product of labor in occupation h diminishes. Hence, labor composition matters: Dispersion in $\frac{L_{jht}}{L_{jt}^*}$, the ratio of occupation- h labor input L_{jht} to aggregate labor L_{jt}^* , generates dispersion in the marginal product of labor in occupation h .

It is important to note that the decompositions (26) and (26) should be interpreted in an accounting sense, and not in a causal sense. Fundamentally, all variation in the marginal product of capital arise from firms' failure to adapt to TFP shocks. Notwithstanding this issue, (26) and (26) are still of interest. On the one hand, if most of the variation in marginal products is comprised of variation in capital-labor ratios and/or labor composition as opposed to variation in TFP (again, in an accounting sense), it indicates that firms respond to TFP shocks, but fail to reach the efficient input levels. This is consistent with a situation where firms face no or relatively low fixed adjustment costs on capital and/or labor, but where adjustment is imperfect, due to e.g. search frictions or convex variable adjustment costs. On the other hand, if TFP shocks accounts for most of the variation in marginal products, it indicates a lack of response to the shocks, which is consistent with a situation where fixed adjustment costs are relatively high.

Table 11 reports $\text{Var}(\ln \frac{Y_{jt}}{K_{jt}})$, and $\text{Var}(\ln \frac{Y_{jt}}{L_{jht}})$ for the four occupations, Managers ($h = 1$), Salaried workers ($h = 2$), Skilled workers ($h = 3$), and Unskilled workers ($h = 4$), as well as the contribution of the covariance components in (26) and (27) in percent of the variance of the relevant log marginal product. The first panel in Table 11 consider dispersion in the log marginal product of capital according to (26). In an accounting sense, we see that this dispersion is almost exclusively driven by differences in capital-labor ratios across firms, accounting for 93, 96, 88, and 97 percent in Manufacturing, Wholesale and Retail Trade, Transport, Storage and Communications, and Real Estate, Renting and Business Activities, respectively. TFP shocks therefore

Table 11: Log Marginal Productivity Variance Decomposition

	Manu- facturing	Wholesale & Retail Trade	Transport, Storage & Communi- cations	Real Estate, Renting & Business Activities
Capital, $Var\left(\ln\frac{Y_{jt}}{K_{jt}}\right)$.944	1.314	2.439	2.802
Capital-labor ratio	92.5%	96.0%	87.6%	97.3%
TFP, total	7.5%	4.0%	12.4%	2.7%
TFP, ηv_{jt-1}	3.4%	.9%	6.7%	.7%
TFP, ε_{jt}^p	4.1%	3.1%	5.7%	2.0%
Managers, $Var\left(\ln\frac{Y_{jt}}{L_{j1t}}\right)$.574	.620	.761	1.065
Capital-labor ratio	3.9%	.1%	.3%	1.3%
Labor composition	77.2%	85.3%	104.8%	74.5%
TFP, total	18.9%	14.7%	-5.1%	24.2%
TFP, ηv_{jt-1}	10.6%	7.4%	-5.1%	14.9%
TFP, ε_{jt}^p	8.3%	7.3%	.0%	9.3%
Salaried workers, $Var\left(\ln\frac{Y_{jt}}{L_{j2t}}\right)$.458	.774	.944	.768
Capital-labor ratio	4.6%	.2%	1.9%	1.9%
Labor composition	63.4%	73.1%	114.0%	76.2%
TFP, total	32.0%	26.7%	-15.9%	21.9%
TFP, ηv_{jt-1}	17.9%	16.1%	-13.6%	14.4%
TFP, ε_{jt}^p	14.1%	10.6%	-2.3%	7.5%
Skilled workers, $Var\left(\ln\frac{Y_{jt}}{L_{j3t}}\right)$.296	.636	.680	1.148
Capital-labor ratio	1.1%	.0%	23.9%	1.6%
Labor composition	66.7%	67.8%	13.9%	71.4%
TFP, total	32.0%	32.2%	62.2%	27.0%
TFP, ηv_{jt-1}	16.4%	19.1%	39.6%	18.8%
TFP, ε_{jt}^p	15.6%	13.1%	22.6%	8.2%
Unskilled workers, $Var\left(\ln\frac{Y_{jt}}{L_{j4t}}\right)$.986	1.053	1.977	2.445
Capital-labor ratio	.1%	-.1%	6.7%	1.6%
Labor composition	95.9%	98.4%	76.9%	85.9%
TFP, total	4.0%	1.7%	16.4%	12.5%
TFP, ηv_{jt-1}	1.9%	-.2%	9.9%	9.1%
TFP, ε_{jt}^p	2.1%	1.9%	6.5%	3.4%

Note: The decompositions are given by (26) for capital and (27) for labor and is based on the production function estimates reported in columns (2), (4), (6) and (8) in Table 8.

accounts for very little of the variance in the marginal product of capital across firms, namely 7, 4, 12, and 3 percent, respectively, in Manufacturing, Wholesale and Retail Trade, Transport, Storage and Communications, and Real Estate, Renting and Business Activities. In all industries, both transitory and persistent TFP induces dispersion in the marginal product of capital.

Consider next the decompositions of the occupation-specific log marginal product of labor in the second to the fifth panel in Table 11, according to (27). The decompositions are qualitatively similar across occupations. First, we see that heterogeneity in capital-labor ratios, the main source of dispersion in the marginal product of capital, contributes only little to dispersion in the marginal products of labor. The contribution from this source of variation varies from a slight negative contribution at $-.1$ percent, for Unskilled workers in Wholesale & Retail Trade to a high 24 percent for Skilled workers in Transport, Storage & Communications.²⁹ However, the latter number is a distinct outlier, with the second highest contribution being 7 percent for Unskilled workers in Transport, Storage & Communications, and the third highest being 4 percent for Managers in Manufacturing.

Instead, labor composition accounts for the largest share of the dispersion in marginal products of labor across firms. This is true for all occupations and in all industries, except for Skilled workers in Transport, Storage & Communications, where TFP variation (TFP, total) is the main source. That one industry-occupation combination aside, the shares of the variance in $\ln \frac{Y_{jt}}{L_{jht}}$ that the decompositions (27) attribute to variation in labor composition ranges between 63 percent (Salaried workers in Manufacturing) and 114 percent (Salaried workers in Transport, Storage & Communications).

Finally, consider the role of TFP. Its total contribution to dispersion in marginal product of labor, in general, across occupations and industries, clearly exceeds its contribution to dispersion in the marginal product of capital. Apart from four outliers, Unskilled workers in Manufacturing, Managers and Salaried workers in Transport, Storage & Communications, and Unskilled workers in Wholesale & Retail Trade where TFP, total contributes 4, -5 , -16 and $-.2$ percent, respectively, TFP shocks accounts for between 13 (Unskilled workers in Real Estate, Renting & Business Activities) and 62 percent (Skilled labor in Transport, Storage & Communications) of the dispersion in log marginal products of occupation-specific labor. In general, TFP variation is the second-largest contributor to dispersion in marginal products of labor. TFP shocks has both transitory and persistent effect on the dispersion, with the persistent effects typically being slightly more important in magnitude.

²⁹We reiterate that the decompositions in Table 11 should not be interpreted in a causal sense. Higher capital-labor ratios causes higher marginal products of labor. Hence, *ceteris paribus*, there is a positive covariance between marginal products of labor and capital-labor ratios. The observed negative covariance in the case of Unskilled workers in Wholesale & Retail Trade arise because we cannot impose the *ceteris paribus* assumption in the decomposition.

Summing up, in an accounting sense, according to (26) and (27), the dispersion in marginal products of capital comes almost exclusively from variation in capital-labor ratios across firms, with very little direct effect from TFP shocks. This is consistent with a situation where it is relatively cheap for firms to adjust their input choices in response to TFP shocks, but it is difficult to get it right, from an efficiency point of view. With respect to dispersion of marginal products of labor, it turns out the labor composition accounts for the lion's share of the dispersion, with some direct, albeit limited, role for TFP shocks. Again, this pattern is consistent with a situation where there are relatively low fixed adjustment costs, but other types of frictions, or variable adjustment costs, prevent the input markets from attaining the efficient allocation of productive resources across firms.

5.2 Incentives for Labor Reallocations

Dispersion in marginal products of labor and rent sharing implies dispersion in the price of human capital across firms. Hence, as already documented, wages here has a firm-specific component. In our data, this component accounts for 3-20 percent of the observed dispersion of wages across workers, see Table 6. While this suggests some incentives are in place for efficient reallocation through job search, we now explicitly quantify this incentive using the estimated model.

Since $\ln w_{it} = \ln \phi_{jht} + \ln a_{ih} + \varepsilon_{iht}^a$ with ε_{iht}^a being idiosyncratic shocks to individual abilities, cf. (12) with $j = J(i, t)$ and $h = H(i, t)$, the expected wage gain for a worker from reallocating from firm j to firm j' is $\ln \phi_{j'ht} - \ln \phi_{jht}$. Reproducing (20) as

$$\ln \phi_{jht} = \ln \left[(1 - \beta_h) b_h + \left(\frac{\frac{\beta_h}{1 - \beta_h}}{1 + \sum_{k=1}^H \alpha_L \gamma_k \frac{\beta_k}{1 - \beta_k}} \right) \alpha_L \gamma_h \frac{Y_{jt}}{L_{jht}} + \varepsilon_{jht}^\phi \right] \quad (28)$$

we see that this return depends on differences between the marginal products, $\alpha_L \gamma_h \frac{Y_{jt}}{L_{jht}}$ and $\alpha_L \gamma_h \frac{Y_{j't}}{L_{j'ht}}$, and differences in the error terms, ε_{jht}^ϕ and $\varepsilon_{j'ht}^\phi$. The error term was added in the empirical analysis and is intended to capture errors-in-variables arising from our multi-step estimation procedure, and factors other than labor productivity influencing the firm-level price of human capital, cf. Assumption A3 and the accompanying discussion in section 4. We are here seeking to isolate and quantify the incentives for workers to seek out firms with high marginal product of labor, and it is appropriate to hold the error term $\varepsilon_{j'ht}^\phi$ constant in our computations. Accordingly, we fix ε_{jht}^ϕ at its expected value, i.e. we take $\varepsilon_{jht}^\phi = 0$.³⁰

³⁰This discussion suggests an interest in quantifying how much of the variation in the piece rate wages can be accounted for by differences in marginal products across firms, and how is left for the residual. Notice however that our estimates of b_h and β_h are midpoints of intervals obtained using

Now, let $MPL_h^{(x)}$ be the x th percentile in the firm-level distribution of marginal products of occupation- h labor. Consider a worker's expected wage return from moving from a firm with marginal product of labor at the Q th percentile (in the firm-level distribution) to a firm where it is at the P th percentile, assuming $P > Q$, and denote this return $R_h^w(Q, P)$. From (28) and the above discussion, we have

$$R_h^w(Q, P) = \ln \left[(1 - \beta_h)b_h + \left(\frac{\frac{\beta_h}{1-\beta_h}}{1 + \sum_{k=1}^H \alpha_L \gamma_k \frac{\beta_k}{1-\beta_k}} \right) MPL_h^{(P)} \right] - \ln \left[(1 - \beta_h)b_h + \left(\frac{\frac{\beta_h}{1-\beta_h}}{1 + \sum_{k=1}^H \alpha_L \gamma_k \frac{\beta_k}{1-\beta_k}} \right) MPL_h^{(Q)} \right]. \quad (29)$$

These returns are easily computed from the estimated model. In Table 12 we report returns for inter-quartile mobility in the distribution of marginal products, i.e. we take $(P, Q) \in \{(25, 75), (25, 50), (50, 75)\}$. The Table also reports the returns to reallocation in marginal products, i.e.

$$R_h^{MPL}(Q, P) = \ln MPL_h^{(P)} - \ln MPL_h^{(Q)}, \quad (30)$$

the log P/Q ratio of the marginal product of labor, $\alpha_L \gamma_h \frac{Y_{jt}}{L_{jht}}$. Consider first the returns in terms of marginal products (the first column for each industry in Table 12. We see that reallocating a worker from a firm on the 25th percentile to a firm on the 75th percentile, results in large productivity gains on the margin. Across occupations and industries, the log 90/10 ratios range from 1.3 (skilled workers, Manufacturing) to 4.1 (Unskilled workers, Real Estate, Renting & Business Activities). Hence, moving a worker from a firm on the 10th percentile to one on the 90th percentile, would in general more than double the worker's marginal product, in some cases much more. Looking within industries, between occupations, we see that the largest gains from reallocation, in terms of marginal products, appear among Unskilled workers. No single occupation stands out as having the smallest gains from reallocation. In Manufacturing, it is among Skilled workers, in Wholesale & Retail Trade it is among Managers, in Transport, Storage & Communications it is among Skilled workers, and finally, in Real Estate, Renting & Business Activities it is among Salaried workers. Comparing industries within the same occupation, we observe the smallest gains from marginal reallocations in Manufacturing, followed by Wholesale & Retail Trade (with Salaried workers being the exception). The largest gains appear in Real Estate, Renting & Busi-

reverse regression techniques. We can therefore not recover residuals with the relevant residual properties, i.e. uncorrelated with marginal products, $\frac{Y_{jt}}{L_{jht}}$. Indeed, any set of residuals obtained confounds variation due to errors-in-variables and proper residual variation.

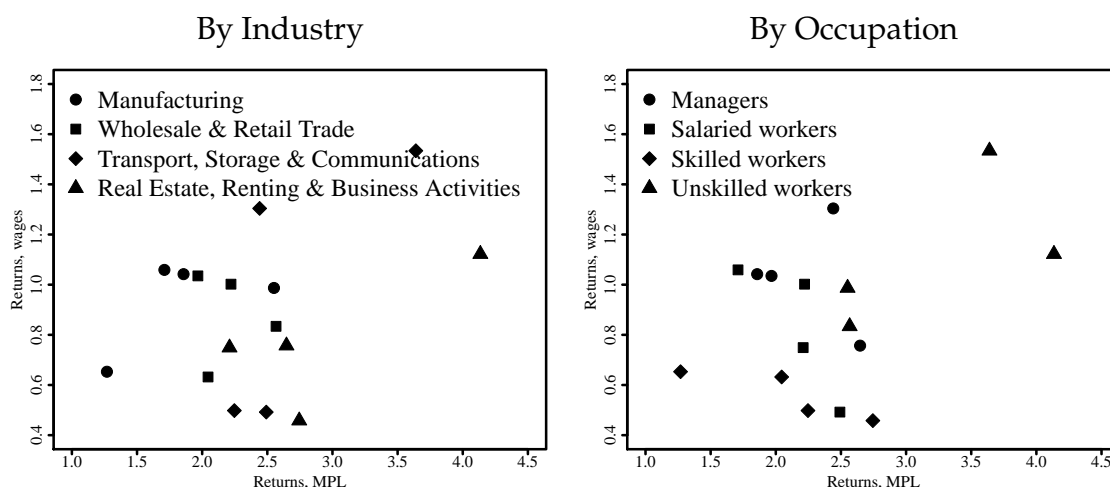
ness Activities (with the exception being Salaried workers). Considering the gains from reallocating a worker from a firm on the 10th percentile in the distribution of marginal products to a firm on the 50th, and from a firm on the 50th to a firm on the 90th, we note that, with some exceptions, the distribution of marginal products is slightly skewed to the left, resulting in a long right tail. This makes the returns slightly larger for transitions at the right tail, with $(Q, P) = (50, 90)$, as compared with transitions in the left tail, with $(Q, P) = (10, 50)$. Overall, while large in magnitude, the reported returns to reallocation in terms of marginal products, is another manifestation of the fact, documented in the previous subsection, that marginal products exhibit substantial dispersion across firms in our data.

What are the wage returns for workers from moving to firms with higher marginal products? Looking at the second column for each industry in Table 12, and focusing again on the gains from reallocating a worker from a firm on the 25th percentile to a firm on the 90th percentile, we also see large wage gains. Indeed, the returns range between .5 (Skilled workers, Real Estate, Renting & Business Activities) and 1.5 (Unskilled workers, Wholesale & Retail Trade). Hence, on the margin, moving a worker from firm on the 10th percentile in the distribution of marginal products, to one on the 90th percentile, lead to at least a 50 percent wage gain for that worker, and typically more, depending on the occupation and industry considered. There is no clear pattern as to which occupations fact the largest wage returns from reallocation. In Manufacturing and Wholesale & Retail Trade, it is Managers and Salaried workers. In Transport, Storage & Communications and Real Estate, Renting & Business Activities it is Unskilled workers. Looking across industries, we see that the wage returns to reallocation are largest in Manufacturing, the exception being Unskilled workers where the return is highest in Transport, Storage & Communications and Real Estate. Overall, the reported wage returns are large, and indicate that the market do provide strong monetary incentives for workers to seek out employment in firms with high marginal products.

We can put the wage returns in Table 12 in contexts by comparing them to estimates of the return to education. This comparison is especially attractive in our setup, as it relates directly to the different policy prescriptions associated with the distinction between wage and productivity dispersion generated from input heterogeneity and misallocation that we have alluded to several times in this paper: If (labor) input heterogeneity were the main driver of observed wage and productivity dispersion, schooling and training is likely to be output enhancing. If misallocation is the prime driver, reallocation is likely to lead to growth. [WE NEED SOME NUMBERS HERE - PERHAPS THERE IS A HANDBOOK CHAPTER WE CAN USE FOR SOME US NUMBERS, AND PERHAPS HELENA HAS SOME FOR THE CASE OF DK?]

Our framework has no predictions as to whether we should expect a positive, nega-

Figure 5: Returns from Reallocation: Returns in Marginal Products against Return in Wages



Note: The plotted numbers are also reported in Table 12.

tive, or any relationship between the returns from reallocation in terms marginal products and in terms of wages. Nonetheless, given our focus on the incentives put in place in the labor market for workers to reallocate efficiently, the relationship between the two returns in is interesting from a descriptive point of view. In Figure 5 we plot the wage returns, (29), against the returns in marginal products, (30). We consider only the return from reallocating a worker from a firm on the 10th percentile to a firm on the 90th percentile in the firm-level distribution of marginal products, i.e. we plot returns for $(Q, P) = (10, 90)$ as reported in Table 12. In the left panel of Figure 5, we mark the datapoint by industries, and in the right panel we mark them by occupations. Overall, Figure 5 reveals a positive relationship between the gains from reallocation in terms of productivity and the associated wage gains. When grouping the data by industry, we might still detect a positive relationship within each industry, even if this is based on only four datapoints for each industry (see left panel of Figure 5). When grouping the data by occupation (right panel of Figure 5) the picture is even murkier, even if one might detect a negative relationships in most occupations.

5.3 Firm Level Wage Dispersion

Figure 1 documented considerable dispersion in firm level average wages. Our model points to the effects of the quality (in terms of ability and composition) of a firm's labor force and of the productivity of the firm through rent sharing for understanding this

Table 12: Returns from Reallocation

		Manu- facturing	Wholesale & Retail Trade	Transport, Storage & Communi- cations	Real Estate, Renting & Business Activities
		R_h^{MPL}	R_h^{MPL}	R_h^{MPL}	R_h^{MPL}
		R_h^w	R_h^w	R_h^w	R_h^w
Managers					
Transition: $(Q, P) = (25, 75)$.947	.528	1.018	1.166	1.460
Transition: $(Q, P) = (25, 50)$.461	.230	.484	.586	.669
Transition: $(Q, P) = (50, 75)$.486	.298	.534	.581	.791
Salaried workers					
Transition: $(Q, P) = (25, 75)$.851	.518	1.206	1.422	1.044
Transition: $(Q, P) = (25, 50)$.376	.208	.528	.592	.505
Transition: $(Q, P) = (50, 75)$.475	.310	.678	.829	.539
Skilled workers					
Transition: $(Q, P) = (25, 75)$.621	.300	1.052	1.227	1.495
Transition: $(Q, P) = (25, 50)$.259	.114	.498	.575	.733
Transition: $(Q, P) = (50, 75)$.362	.186	.554	.651	.762
Unskilled workers					
Transition: $(Q, P) = (25, 75)$	1.333	.503	1.355	1.999	2.357
Transition: $(Q, P) = (25, 50)$.673	.203	.612	.916	1.244
Transition: $(Q, P) = (50, 75)$.660	.300	.743	1.083	1.113

Note: The return to marginal products $R_h^{MPL} = R_h^{MPL}(Q, P)$ is given by (30). The wage return $R_h^w = R_h^w(Q, P)$ is given by (29). Computations are based on the production function estimates reported in columns (2), (4), (6) and (8) in Table 8, and in Table 9.

phenomenon. From the piece rate wage equation (20), we obtain the firm level wage

$$\bar{w}_{jt} = \frac{L_{jt}}{N_{jt}} \sum_{h=1}^H (1 - \beta_h) b_h \frac{L_{jht}}{L_{jt}} + \left(\frac{\sum_{k=1}^H \alpha_L \gamma_k \frac{\beta_k}{1 - \beta_k}}{1 + \sum_{k=1}^H \alpha_L \gamma_k \frac{\beta_k}{1 - \beta_k}} \right) \frac{Y_{jt}}{N_{jt}} + \bar{\varepsilon}_{jt}^\phi, \quad (31)$$

where $\bar{\varepsilon}_{jt}^\phi = \sum_{h=1}^H \varepsilon_{jht}^\phi \frac{L_{jht}}{N_{jt}}$. As in (8), the first term, $\frac{L_{jt}}{N_{jt}} \sum_{h=1}^H (1 - \beta_h) b_h \frac{L_{jht}}{L_{jt}}$, is the relevant firm level labor force quality measure, combining average ability $\frac{L_{jt}}{N_{jt}}$ and composition effects. The second term on the right hand side of (31) is the rent sharing term, linear in firm level labor productivity. The two are not independent, since a firm that hires more productive workers thereby obtains higher labor productivity. They are not perfectly correlated, either, since our productivity decomposition has revealed that high productivity may reflect other things than a high quality work force, in particular, TFP and capital intensity. The third term on the right hand side of (31) represents variation in firm wages that are unexplained by our model.

The firm level wage equation thus gives rise to the decomposition of firm level wage dispersion

$$\begin{aligned} \text{Var}(\bar{w}_{jt}) = & \underbrace{\sum_{h=1}^H (1 - \beta_h) b_h \text{Cov} \left(\bar{w}_{jt}, \frac{L_{jt}}{N_{jt}} \cdot \frac{L_{jht}}{L_{jt}} \right)}_{\text{Labor force quality}} \\ & + \underbrace{\left(\frac{\sum_{k=1}^H \alpha_L \gamma_k \frac{\beta_k}{1 - \beta_k}}{1 + \sum_{k=1}^H \alpha_L \gamma_k \frac{\beta_k}{1 - \beta_k}} \right) \text{Cov} \left(\bar{w}_{jt}, \frac{Y_{jt}}{N_{jt}} \right)}_{\text{Rent sharing}} + \underbrace{\text{Cov}(\bar{w}_{jt}, \bar{\varepsilon}_{jt}^\phi)}_{\text{Residual}} \quad (32) \end{aligned}$$

into H occupation-specific labor force quality contributions, the contribution to firm wage dispersion from productivity dispersion through rent sharing, and a residual term.

Table 13 reports the variance decompositions. The numbers in the top portion of the table are based on the structural parameters from Table 8 (columns 2, 4, 6, and 8) for the production function parameters $(\alpha_L, \gamma_1, \dots, \gamma_4)$ from the second step and Table 9 for the piece rate wage function parameters $(\beta_1, b_1, \dots, \beta_4, b_4)$ from the third step of our estimation procedure. In Manufacturing, rent sharing accounts for 33 percent of the observed variance in firm wages, and differences in labor quality for 12 percent. The portion of cross firm wage dispersion accounted for by labor force quality and rent sharing taken together, at 45 percent, is quite substantial, e.g., in comparison with the low R^2 s common in worker level wage regressions, where observable characteristics that are supposed to account for productivity differences typically explain no more than 30 percent of the variation, cf. Mortensen (2003). In the other three indus-

Table 13: Firm Wage Variance Decomposition

	Manu- facturing	Wholesale & Retail Trade	Transport, Storage & Communi- cations	Real Estate, Renting & Business Activities
$Var(\bar{w}_{jt})$	951.5	1,965.8	890.6	3,874.6
Relative contributions, structural parameters				
Labor force quality	.120	.146	-.139	.016
Managers	.114	.117	.056	.425
Salaried workers	.110	.264	.079	.081
Skilled workers	-.039	-.183	-.155	-.251
Unskilled workers	-.066	-.052	-.120	-.238
Rent sharing	.330	.207	.309	.129
Residual	.550	.646	.830	.855
Relative contributions, projection parameters				
Labor force quality	.172	.481	.054	.411
Managers	.075	.207	.016	.220
Salaried workes	.073	.350	.016	.012
Skilled workers	.001	-.066	.009	.074
Unskilled workers	.022	-.009	.013	.105
Rent sharing	.120	.097	.237	.105
Residual	.709	.422	.709	.484

Note: The decomposition is given by (32).

tries, rent sharing similarly accounts for a greater share of firm wage variance than differences in labor force quality, although the two together account for less than in Manufacturing. As in the decomposition of labor productivity dispersion in Table ??, some negative contributions appear when disaggregating labor force quality by occupation, reflecting that shifting the labor force composition toward lower occupations reduces average wage, the effect again being sufficiently strong in the smallest industry, Transport, Storage & Communications, for the overall contribution of labor force quality to turn out negative.

The linearity of firm wage \bar{w}_{jt} in $(\frac{L_{j1t}}{N_{jt}}, \dots, \frac{L_{j4t}}{N_{jt}}, \frac{Y_{jt}}{N_{jt}})$ in (31) suggests an interest in the alternative decomposition obtained simply by projecting wages on this set of five variables, as consistent with Assumption A3. The bottom portion of Table 13 presents such alternative decompositions, using regression coefficients instead of structural parameter estimates in (32). In Manufacturing and Transport, Storage & Communications, the contributions from dispersion in labor force quality and rent sharing together now account for about 30 percent of the observed variation in firm wages. In the other two industries, this figure exceeds 50 percent. Compared to the structural decompositions, the fraction of variance accounted for increases in all industries except Manufacturing. Within the portion accounted for, the fraction due to differences in labor force quality is 59 percent in Manufacturing, with 41 percent due to the rent sharing effect. In the other industries, these fractions are 83, 19, and 65 percent for labor force quality, and 17, 81, and 35 for rent sharing. In the first three industries, most of the wage variation attributed to labor force quality is due to Managers and Salaried labor. In the last industry, Real Estate, Renting & Business Activities, Managers also contribute most, although in this case Salaried labor the least, among the four occupations.

All in all, the results show that our model accounts for a relatively large portion of firm wage variation, and that within this portion, both differences in the quality of firms' work forces and the effect of productivity dispersion on wages through rent sharing are important contributors. The structural parameters suggest that rent sharing is more important than labor force quality differences for firm wage variation, whereas the projection indicates that it is slightly less important. We take these results as evidence that, even if the labor market is frictional and sustains intrinsic productivity differences across firms, it does direct labor search toward more productive use of labor resources.

5.4 The Covariance Between Productivity and Wages

The final set of decomposition results relates to the covariance between productivity and wages (see Figure 2). Our model embeds two mutually consistent explanations of

the observed positive relation. From (31), it follows that

$$\begin{aligned}
Cov\left(\bar{w}_{jt}, \frac{Y_{jt}}{N_{jt}}\right) &= \underbrace{\sum_{h=1}^H (1 - \beta_h) b_h Cov\left(\frac{L_{jt}}{N_{jt}} \cdot \frac{L_{jht}}{L_{jt}}, \frac{Y_{jt}}{N_{jt}}\right)}_{\text{Labor force quality}} \\
&\quad + \underbrace{\left(\frac{\sum_{k=1}^H \alpha_L \gamma_k \frac{\beta_k}{1 - \beta_k}}{1 + \sum_{k=1}^H \alpha_L \gamma_k \frac{\beta_k}{1 - \beta_k}}\right) Var\left(\frac{Y_{jt}}{N_{jt}}\right)}_{\text{Rent sharing}} + \underbrace{Cov\left(\bar{\varepsilon}_{jt}^\phi, \frac{Y_{jt}}{N_{jt}}\right)}_{\text{Residual}}. \quad (33)
\end{aligned}$$

The first term on the right hand side of (33) represents the covariance between wages and average productivity that is induced by ability and compositional differences in the workforce across firms: If more productive firms have higher quality work forces, they also pay higher wages. The second term on the right hand side of (33) arises out of rent sharing. If average productivity is distributed across firms, rent sharing implies that more productive firms pay higher wages, irrespective of the quality of their labor force.

Table 14 reports the covariance decompositions. The reported shares are all relative to $Cov\left(\bar{w}_{jt}, \frac{Y_{jt}}{N_{jt}}\right)$. As in the previous table, the top portion is based on structural parameter estimates, and the bottom portion on a projection, implying that in the latter, the final term on the right hand side of (33) vanishes identically by construction. Based on the structural estimates, rent sharing accounts for 72 percent or more of the covariance between firm wage and labor productivity in all industries, and differences in labor force quality account for 16 percent or less. The projection decomposition implies that in Manufacturing, the rent sharing effect comprises 75 percent and labor force quality differences 25 percent of the covariance. In all but one industry, rent sharing motives are behind more than half of the observed covariance. The exception is Wholesale & Retail Trade, but even here, rent sharing does account for 44 percent of the covariance between wage and productivity. Again, this share is higher in other industries, e.g., 91 percent in Transport, Storage & Communications, and even higher based on the structural parameters (top portion of Table 14).

Within the share of the covariance due to labor force quality differences, the pattern is the same as for total wage dispersion: The greatest portion is explained by differences in Management and Salaried labor in all four industries based on the structural parameters, and in the first three industries based on the projection. In the last industry, Real Estate, Renting & Business Activities, Management similarly contributes the most toward the covariance, albeit Salaried labor the least, based on the projection.

The positive relation between firm-level wages and labor productivity observed in Figure 2 thus derives to a large extent from rent sharing in a frictional labor market.

Table 14: Decomposition of the Covariance Between Firm Wage and Labor Productivity

	Manu- facturing	Wholesale & Retail Trade	Transport, Storage & Communi- cations	Real Estate, Renting & Business Activities
$Cov\left(\bar{w}_{jt}, \frac{Y_{jt}}{N_{jt}}\right)$	1,101.4	2,451.8	3,504.2	4,962.9
Relative contributions, structural parameters:				
Labor force quality	.156	.156	-.227	.063
Managers	.141	.162	.151	.427
Salaried workers	.199	.287	.046	.091
Skilled workers	-.129	-.291	-.155	-.155
Unskilled workers	-.055	-.002	-.269	-.300
Rent sharing	2.075	.939	1.185	.723
Residual	-1.232	-.095	.042	.213
Relative contributions, projection parameters:				
Labor force quality	.247	.561	.091	.412
Managers	.092	.287	.044	.222
Salaried workers	.132	.380	.009	.013
Skilled workers	.004	-.105	.009	.045
Unskilled workers	.018	-.0003	.029	.132
Rent sharing	.753	.439	.909	.588

Note: The decomposition is given by (33).

Again, this is evidence that the labor market, albeit frictional, provides some incentives for workers to reallocate from less to more productive firms.

6 Conclusion

We have studied a model for which differences in TFP, labor input quality, and capital intensity as well as rent sharing are allowed to explain the observations that wage and value added per worker are dispersed across firms, and that more productive firms tend to pay higher wages. The investigation is motivated by the difference in policy implications of the two alternative mechanisms behind the observations. If the observed productivity and wage dispersion reflects differences in the composition and quality of labor and capital inputs, then an increased growth target calls for investment in education and skill upgrading. On the other hand, if there is intrinsic TFP heterogeneity, then labor reallocation may generate growth, and if rents are shared, the dispersion in productivity spills over into wages, and motivates job search as an incentive for reallocation.

Estimating the model using detailed MEE data from Denmark, we find that TFP, labor force quality, capital intensity, and rent sharing all play important roles in wage and productivity dispersion. Labor force quality is measured using an appropriate combination of ability through (occupation specific) worker fixed effects from the wage equation and firm level occupational work force composition through occupation specific relative productivities or outside options and bargaining powers. From our empirical results, a constant returns to scale production function with capital, log linear aggregator over ability adjusted occupational labor inputs, and a persistent stochastic TFP process is not rejected by the data. Most of the variation in labor productivity in all industries considered is attributable to TFP differences, with only smaller fractions associated with labor force quality and in most cases still smaller fractions with capital intensity differentials. Within the portion of productivity dispersion attributable to differences in labor force quality, firm differences in average ability of workers employed are less important than differences in the occupational composition of the labor force.

Regardless the source of average productivity dispersion, its presence generates a rent sharing motive, implying that more productive firms pay higher wages. We find that rent sharing exists and accounts for between 10 and 33 percent of the variation in firm level wages, depending on industry and estimation method—structural versus projection. Differences in labor force quality accounts for less based on the structural parameters, and for comparable fractions based on the alternative decomposition. Overall, both rent sharing and labor force quality are important contributors.

Further, rent sharing is the most important reason for the observed positive relation between firm wages and labor productivity, although labor force quality variation plays a role, too. The portion of both wage dispersion and covariance with productivity accounted for by labor force quality is driven mainly by the higher occupations in the education-skill hierarchy. This is so even though the lower occupations exhibit individual level wage dispersion that is to a greater extent driven by piece rate wages (firm-year-occupation effects) and less by ability (worker-occupation effects) than for higher occupations.

Altogether, our results show that observed cross firm productivity differentials to a great extent reflect inherent firm differences, thus indicating a scope for productivity improving worker reallocation between firms. Given our rich measurements of labor quality and detailed information on labor compensation, our focus is on potential output enhancing labor reallocations. The importance of rent sharing in generating wage dispersion and the observed positive covariance between wages and labor productivity at the firm-level imply that the wage distribution to a considerable degree reflect firm productivity differences. This provides at least some incentive for reallocation through job search. That is, even if the labor market is frictional and sustains intrinsic productivity differences across firms, it does direct labor search toward more efficient use of labor resources.

Finally, one important caveat is in order. We mentioned earlier that our interpretation of worker fixed effects from the individual level wage regression is questionable if a worker's previous labor market search history affects the wage as in Postel-Vinay and Robin (2002). In this case, the worker fixed effects are likely to overstate the workers' productive human capital. As workers seek out firms that pay a premium and are more likely to remain once one is found, their model is quite consistent with the possibility that high wage firms have more workers with favorable search experience, on the one hand, but that the firm labor input adjusted to reflect the measured fixed effects has little impact on labor productivity, on the other. In our measurements, this particular mechanism would imply an upward bias in the shares of the covariance between firm wage and labor productivity attributable to rent sharing. Nevertheless, given the magnitudes of our estimates, with shares at 75 percent or higher in Manufacturing, and other shares at 44 percent or higher, there is ample evidence that rent sharing does play a role.

References

ABOWD, J. M., R. H. CREECY, AND F. KRAMARZ (2002): "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," Discussion Pa-

- per TP-2002-06, U.S. Census Bureau, LEHD Program.
- ABOWD, J. M., F. KRAMARZ, AND D. N. MARGOLIS (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67, 251–333.
- ACKERBERG, D., K. CAVES, AND G. FRAZER (2006): "Structural Identification of Production Functions," Discussion paper, UCLA Working Paper.
- ARELLANO, M., AND S. BOND (1991): "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58, 277–297.
- BAGGER, J., AND R. LENTZ (2013): "An Equilibrium Model of Wage Dispersion and Sorting," Discussion paper, Royal Holloway, University of London.
- BAILY, M. N., C. HULTEN, AND D. CAMPBELL (1992): "Productivity Dynamics in Manufacturing Plants," *Brookings Papers on Economic Activity. Microeconomics*, pp. 187–267.
- BARTELSMAN, E., J. C. HALTIWANGER, AND S. SCARPETTA (2013): "Cross-Country Differences in Productivity : The Role of Allocation and Selection," *American Economic Review*, 103(1), 305–334.
- BERNARD, A. B., J. EATON, J. B. JENSEN, AND S. KORTUM (2003): "Plants and Productivity in International Trade," *The American Economic Review*, vol. 93, 1268–1290.
- BERNARD, A. B., AND J. B. JENSEN (1995): "Exporters, Jobs, and Wages in U.S. Manufacturing: 1976-1987," *Brookings Papers on Economic Activity: Microeconomics*, pp. 67–119.
- BLUNDELL, R., AND S. BOND (1998): "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87, 115–143.
- BURDETT, K., AND D. T. MORTENSEN (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39(2), 257–73.
- CAHUC, P., F. MARQUE, AND E. WASMER (2008): "Intrafirm Wage Bargaining in Matching Models: Macroeconomic Implications and Resolution Methods with Multiple Labor Inputs," *International Economic Review*, 49(3), 943–972.
- CAHUC, P., F. POSTEL-VINAY, AND J.-M. ROBIN (2006): "Wage Bargaining with On-the-Job Search: Theory and Evidence," *Econometrica*, 74(2), 323–364.
- CHRISTENSEN, B. J., R. LENTZ, D. T. MORTENSEN, G. NEUMANN, AND A. WERWATZ (2005): "On-the-Job Search and the Wage Distribution," *Journal of Labor Economics*, 23(1), 31–58.

- EATON, J., S. KORTUM, AND F. KRAMARZ (2004): "Dissecting Trade: Firms, Industries, and Export Destinations," *The American Economic Review, Papers and Proceedings*, 94, 150–154.
- (2011): "An Anatomy of International Trade: Evidence from French Firms," *Econometrica*, 79(5), 1453–1498.
- ECKHOUT, J., AND P. KIRCHER (2011): "Identifying Sorting - In Theory," *Review of Economic Studies*, 78(3), 872–906.
- EUROSTAT (1996): *NACE Rev. 1: Statistical Classification of Economic Activities in the European Union*.
- FOSTER, L., J. C. HALTIWANGER, AND C. J. KRIZAN (2001): *Aggregate Productivity Growth. Lessons from Microeconomic Evidence*, chap. 9, pp. 303–372, *New Developments in Productivity Analysis*. National Bureau of Economic Research, Inc.
- FOX, J., AND V. SMEETS (2011): "Does Input Quality Drive Measured Differences in Firm Productivity?," *International Economic Review*, 52(4), 961–989.
- FRIEDMAN, M., AND A. J. SCHWARTZ (1982): *Monetary Trends in the United States and the United Kingdom: The Relation to Income, Prices, and Interest Rates, 1867-1975*. Chicago University Press.
- GANDHI, A., S. NAVARRO, AND D. RIVERS (2011): "On the Identification of Production Functions: How Heterogeneous is Productivity?," Discussion paper, University of Western Ontario.
- HAGEDORN, M., T. H. LAW, AND I. MANOVSKII (2013): "Identifying Sorting," Discussion paper, University of Pennsylvania.
- HANSEN, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50(4), 1029–105.
- HOLTZ-EAKIN, D., W. NEWEY, AND H. S. ROSEN (1988): "Estimating Vector Autoregressions with Panel Data," *Econometrica*, 56(6), 1371–1395.
- HSIEH, C.-T., AND P. J. KLENOW (2009): "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 124(4), 1403–1448.
- IRANZO, S., F. SCHIVARDI, AND E. TOSETTI (2008): "Skill Dispersion and Firm Productivity: An Analysis with Employer-Employee Matched Data," *The Journal of Labor Economics*, 26(2), 247–285.

- IRARRAZABAL, A., A. MOXNES, AND K. H. ULLTVEIT-MOE (2010): "Heterogeneous Firms or Heterogeneous Workers? Implications for Exporter Premia and the Gains from Trade," Discussion paper, University of Oslo.
- JORGENSEN, D., F. GOLLOP, AND B. FRAUMENI (1987): *Productivity and US Economic Growth*. Harvard University Press.
- KRASKER, W., AND J. PRATT (1986): "Bounding the Effects of Proxy Variables on Regression Coefficients," *Econometrica*, 54, 641–656.
- KRUEGER, A. B., AND L. H. SUMMERS (1989): "Efficiency Wages and the Wage Structure," *NBER Working Paper 1952*.
- LEAMER, E. (1987): "Errors in Variables in Linear Systems," *Econometrica*, 55, 893–909.
- LENTZ, R., AND D. T. MORTENSEN (2010): "Labor Market Models of Worker and Firm Heterogeneity," *Annual Review of Economics*, 2, 577–602.
- LEVINSOHN, J., AND A. PETRIN (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70(2), 317–341.
- LISE, J., C. MEGHIR, AND J.-M. ROBIN (2013): "Mismatch, Sorting and Wage Dynamics," Discussion paper, UCL.
- LOPES DE MELO, R. (2013): "Firm Wage Differentials and Labor Market Sorting: Firm Wage Differentials and Labor Market Sorting: Reconciling Theory and Evidence," Discussion paper, University of Chicago.
- MORTENSEN, D. T. (2003): *Wage Dispersion. Why Are Similar Workers Paid Differently?* The MIT Press.
- (2010): "Wage Dispersion in the Search and Matching Model," *The American Economic Review Papers and Proceedings*, 100, 1–10.
- MURPHY, K. M., AND R. H. TOPEL (1990): *Efficiency Wages Reconsidered: Theory and Evidence* pp. 204–40, *Advances in the Theory and Measurement of Unemployment*. Macmillan.
- OLLEY, G. S., AND A. PAKES (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64(6), 1263–1297.
- PAGAN, A. (1984): "Econometric Issues in the Analysis of Regressions with Generated Regressors," *International Economic Review*, 25, 221–247.
- PEDERSEN, N. (2009): "Essays in International Trade," Ph.D. thesis, Northwestern University.

- PISSARIDES, C. (2000): *Equilibrium Unemployment Theory*. MIT Press.
- POSTEL-VINAY, F., AND J.-M. ROBIN (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 70(6), 2295–2350.
- (2006): *Microeconomic Search-Matching Models and Matched Employer-Employee Data*, chap. 11, pp. 279–310, The Proceedings of the 9th World Congress of the Econometric Society. Cambridge University Press: Cambridge, UK.
- RESTUCCIA, D., AND R. ROGERSON (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, 11(4), 707–720.
- (2013): "Misallocation and Productivity," *Review of Economic Dynamics*, 16, 1–10.
- ROODMAN, D. (2009): "A Note on the Theme of Too Many Instruments," *Oxford Bulletin of Economics and Statistics*, 71(1), 135–158.
- SMITH, E. (1999): "Search, Concave Production and Optimal Firm Size," *Review of Economic Dynamics*, 2(2), 456–471.
- STOLE, L. A., AND J. ZWIEBEL (1996): "Intra-Firm Bargaining under Non-Binding Contracts," *The Review of Economic Studies*, 63(3), 375–410.
- SYVERSON, C. (2004): "Product Substitutability and Productivity Dispersion," *Review of Economics and Statistics*, 86(2), 534–550.
- ZELLNER, A. (1962): "An Efficient Method of Estimating Seemingly Unrelated Regression Equations and Tests for Aggregation Bias," *Journal of the American Statistical Association*, 57, 348–368.