QED

Queen's Economics Department Working Paper No. 1200

Wages And Seniority When Coworkers Matter: Estimating A Joint Production Economy Using Norwegian Administrative Data

Christopher Ferrall
Queen's University

Kjell G. Salvanes
Norwegian School of Economics

Erik Ø. Sørensen
Norwegian School of Economics

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

4-2009

WAGES AND SENIORITY WHEN COWORKERS MATTER: ESTIMATING A JOINT PRODUCTION ECONOMY USING NORWEGIAN ADMINISTRATIVE DATA

Christopher Ferrall

Department of Economics

Queen's University

ferrallc@queensu.ca

Kjell G. Salvanes

Department of Economics

Norwegian School of Economics

Kjell.Salvanes@nhh.no

Erik Ø. Sørensen

Department of Economics

Norwegian School of Economics

sameos@gmail.com

April 1, 2009

Abstract:

We develop an equilibrium model of wages and estimate it using administrative data from Norway. Coworkers interact through a task-assignment model, and wages are determined through multi-lateral bargaining over the surplus that accrues to the workforce. Seniority affects wages through workplace output and relative bargaining power. These channels are separately identified by imposing equilibrium restrictions on data observing all workers within workplaces. We find joint production is important. Seniority affects bargaining power but is unproductive. We reinterpret gender and firm-size effects in wages in light of the rejection of linearly separable production.

JEL Classification: D2, J3, J24, L25, J7

Research support from the Social Science Research Council of Canada and the Norwegian Research Council is gratefully acknowledged. Earlier versions have benefitted from comments at several conferences and seminars. Access to the data on individual workers and firms used here requires a research agreement with Statistics Norway. Stata and Ox programs used to construct the sample and to carry out the estimation are available from the authors.

1. Introduction

This paper develops an equilibrium model of joint production designed to explain wages in matched data sets. The primary question we address is an old one: what is the relationship between seniority (tenure) and wages? We shed new light on this question by allowing seniority and wages to be related through two channels. The first is the usual productivity channel: seniority is a proxy for accumulated firm-specific skills. The second is a distributive channel: seniority affects a worker's bargaining power relative to coworkers when distributing the surplus created by the workplace. The combination of an equilibrium model and rich data allows us to separate these channels. Using matched data for Norway in 1997 on a large sub-sample of all workplaces we estimate the model to disentangle the absolute and relative affects of seniority and to re-interpret estimates of models that assume coworkers do not interact.

Before the arrival of widely available administrative data matching workers with their employers (Haltiwanger et al. 1999) most labor market research used data on individuals. When an individual's coworkers are unobserved it is natural to model output as linearly separable across workers within a firm. The assumption of linearly separable workers has also been maintained by recent models of the labour market, such as Burdett and Mortensen (1998) and Postel-Vinay and Robin (2002). There is no lack of models of how workers interact within firms (see e.g. Boyd et al. 1988 and Sattinger 1993), but they are often considered as applying to single or homogeneous firms or to a narrow segment of the labor market with special data available for it. Matched data sets create a new opportunity to consider questions that apply to the full spectrum of workplaces.

The joint technology we develop is a recursive task assignment model of production within firms based on Rosen (1982). Worker talent is allocated across tasks to equalize intermediate output and its demand from the top-level task. Workplaces with different workers have a different assignment problem and end up allocating talent differently. The empirical specification is chosen to include as a special case the separable technology where worker skill is consistent with a Mincer wage equation. In the linear case a worker's observed wage equals their value marginal product (VMP). But when heterogeneous workforces engage in joint production this condition is neither necessary nor sufficient to describe equilibrium wages. First, a worker's internal VMP can only be determined by hypothetically removing the worker and re-allocating the remaining workers to tasks and computing total output. The workplace creates a surplus jointly that must then be allocated jointly. We use multi-lateral Nash (1950) bargaining to allocate the surplus. Even with joint production but without labor market frictions, in a long run equilibrium all workers would be assigned coworkers such that their VMP is no lower than in any other workplaces. We study a static equilibrium in which workforces are given and frictions exist. This creates a second wedge between wages and VMP, because a single worker does not have a well-defined external VMP. The joint technology determines the worker's contribution to their workplace and its distribution across other workplaces. In our model there is competitive pressure on the outside alternatives (threat points) in the Nash bargaining problem. In the static equilibrium outside alternatives equal the expected outcome from a hypothetical search and bargaining outcome for another workplace.[1]

Our data come from one year of a matched panel of Norwegian workplaces (described in Salvanes et al. 1999) that combines information from a number of administrative databases to provide a complete picture of employment, earnings, transfers and education for the Norwegian population. Parameters are estimated by fitting the model's predictions to a 20% sample of all workplaces with more than one employee. The main question we address with our estimation of a joint economy equilibrium is the relationship between wages and seniority. The usual starting point for these questions is a Mincer wage equation such as

$$\ln W = b_0 \text{Exper} + b_1 \text{Exper}^2 + b_2 \text{Sen} + b_3 \text{Sen}^2 + \dots \tag{1}$$

[1] Teullings (1995), Ferrall (1997), and Costrell and Loury (2004) are recent examples of task assignment models. In each case no workplace-specific surplus is created so that the wage-ability relationship is determined by competition.

where Exper is years of labor market experience, and Sen is years with the current employer (e.g. Mincer and Jovanovic 1981). In our context we have a measure of actual experiences as opposed to potential experience (Age - Years of Schooling - Age at School Start). Usually (1) is estimated separately for men and women. Since seniority is based on job choices in the past, which in turn depends on unobserved determinants of current wages, OLS estimates are likely to be biased. A line of research has worked to construct consistent estimates of (1) or some variant of it. Some work stresses use of panel data on individuals to correct for endogeneity of seniority. Two recent examples are Altonji and Williams (2005), who find modestly rising concave seniority profiles and Dustmann and Meghir (2005) who find rising profiles and some flat or even declining profiles in some occupations.

We use the equilibrium Nash bargaining model of a workplace's whole payroll to address two issues related to seniority. First, much empirical work glosses over the indeterminant nature of the wage-seniority relationship. With firm-specific capital comes a bilateral surplus held by the firm-worker pair which faces no outside competitive pressure. Nothing requires wage profiles to trace out the marginal (social) return to a worker's seniority (even in the linearly separable model). This contrasts with general skills for which wages reflect value marginal product. Rather, a range of seniority profiles are consistent with equilibrium.[2] We pin down the difference between productivity and wages by modeling coworkers who produce together and share their part of the surplus in equilibrium. Rather than emphasizing panel observations we emphasize matched observations that provide a snapshot of whole workplaces to tie down total product of the workforce acting as a team.

Second, models that rely on linear output have no room to consider the role of factors internal to the firm in wage settings. We allow the surplus sharing rule to depend on

[2] This is separate from explanations of seniority profiles due to information and incentives, as discussed in Hutchens (1989) or Prendergast (1999).

relative seniority. This makes operational insider-outsider wage effects such as Lindbeck and Snower (1998). It also allows gender differences to depend on something that approximates office politics as well as productivity differences. Again, these considerations are consistent with competitive forces determining the value of outside alternatives but leaving a specific surplus to distribute. Consider a workplace where hypothetically all workers arrived one year earlier (thus seniority rises holding experience constant). Workers interact and skills are reallocated within the firm. The surplus changes and the worker's share of that surplus may go up or down because their relative seniority can go down even as it rises absolutely. These ambiguities, built into our equilibrium estimates of model parameters, can help explain why estimates of the return to seniority are variable and imprecise when they ignore coworker interaction and the sharing of specific surplus.[3]

We estimate both the joint production technology and its special case of linearly separable production. The linear model is rejected, and most one-digit industries are estimated to have significant coworker interactions. Allowing seniority to enter both bargaining and technology provides only a slightly better fit than making seniority unproductive. This model of co-worker interactions and seniority that is unproductive but relevant for distributing joint surplus becomes our preferred specification. We explore further implications of relaxing linear separability across workers. In particular the joint production model provides different interpretations of firm-size and male-female wage differentials.

2. The Joint Production Economy

2.1 Workers and Workforces

In the model a workplace matches the usual definition of a plant or establishment

[3] Buhai et al. (2008) develop a bargaining model of turnover, seniority and tenure. Their model has linearly separable production across co-workers but sequential bargaining between the firm and arriving workers creates a relationship between wages and and relative seniority. They also construct relative seniority from matched data although they do not impose equilibrium restrictions from the model on the data.

in employer-employee matched data sets. Namely, a workplace is a single physical site which may comprise the whole firm or one location of a multi-site firm. Each workplace produces a quantity of a single final good. It has an exogenously determined workforce attached to it consisting of N workers. Worker n in the workforce has a $1 \times P$ vector of observed and exogenous characteristics, x_n. The $N \times P$ matrix X containing the row vectors x_n describes the workforce.[4]

Observed characteristics of workers shift their talent. A worker contributes their talent which interacts with the talents of coworkers through a technology that determines output. Talent has both internal and external components. External (general) talent transfers to other workplaces. Internal (specific) talent is left behind if the worker leaves the current workplace. Computing the optimal allocation of talent, which is how the mapping from talent to output is completed, is simpler when the distribution of talents is smooth. With a finite heterogeneous workforce, a smooth talent distribution can be created by assuming that each worker provides not a point-valued talent but a talent distribution.[5]

Assumption A1: Talent. A worker with characteristics x_n has a talent distribution in their current workplace denoted $G(a; x_n\gamma)$ with corresponding density $g(a; x_n\gamma)$. The vector γ contains exogenous coefficients.

A1a. The index $x\gamma$ is composed of internal and external components:

$$\underbrace{x\gamma}_{\text{total}} = \underbrace{xM\gamma}_{\text{external}} + \underbrace{x(I-M)\gamma}_{\text{internal}}, \tag{2}$$

where M is an exogenous idempotent $P \times P$ matrix.

A1b. A worker's talent follows the exponential distribution: $G(a; x\gamma) = 1 - e^{-ae^{-x\gamma}}$.

[4] In order to focus on the role of joint production, the relationships among workplaces of a multi-site firm are ignored.

[5] One can interpret this assumption by supposing that the workplace sets allocation rules to maximize expected revenue before worker talent is realized. Workers come in to work repeatedly, drawing an amount of talent a from their own talent distribution. Based on their draw they play an assigned role. In small workplaces where the daily distribution of talent may diverge greatly from the expected distribution, a buffer stock of the intermediate good would smooth output .

The matrix M strips off the workplace-specific characteristics of a worker leaving their external characteristics that apply in all other workplaces. In the empirical specification the columns of x_n include functions of seniority (tenure) at the current workplace. A worker who moves to another workplace has seniority reset to zero but keeps all other characteristics. For example, if column 5 contains seniority and column 6 seniority squared then M would be the order P identity matrix except the 5th and 6th elements of the diagonal would be zero.

Figure 1 illustrates the distribution of talents of two workers. Worker 2 is the more able of the two. The talent distribution in the workplace is a vertical average of the two densities. In general the distribution is the mixture over N densities:

$$g(a; X\gamma) = \frac{1}{N}\sum_{n=1}^{N} g(a; x_n\gamma). \tag{3}$$

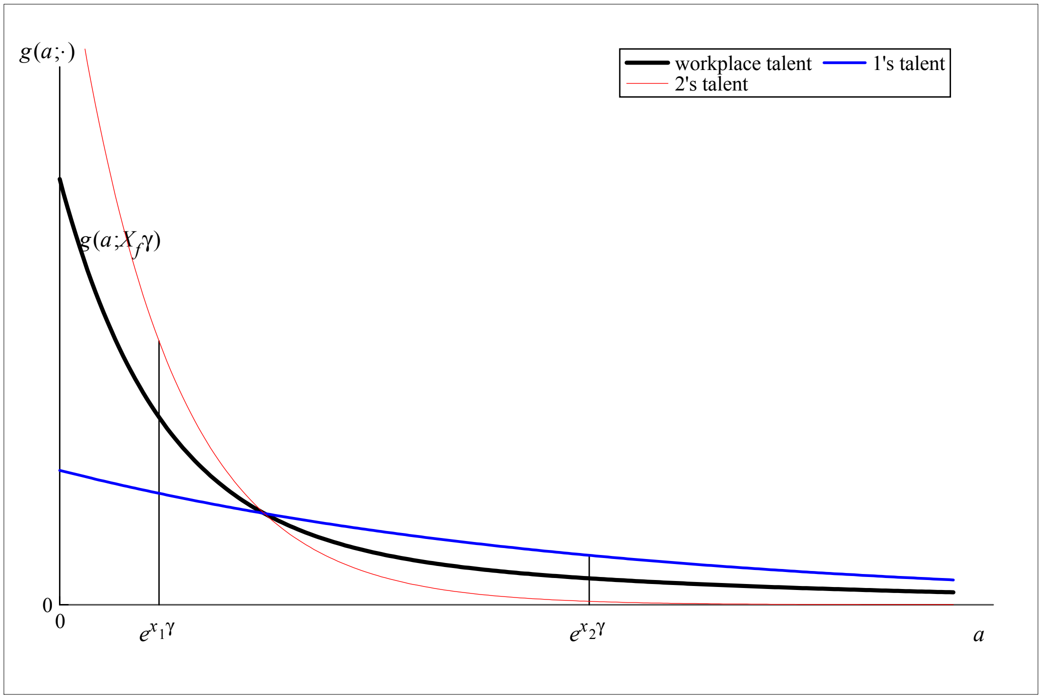
Assuming that a worker's talent follows the exponential distribution (A1b) is convenient for the empirical analysis. The scalar log of expected talent, $\ln E[a|x] = \ln(1/e^{-x\gamma}) = x\gamma$, plays the same role as "skill" in the typical human capital model with linearly separable technology. Since the exponential distribution is a one-parameter distribution a larger value of $x\gamma$ is first-order stochastic dominant over a smaller value (as seen in Figure 1). In equilibrium the value of outside alternatives is a monotonic function of a single index $xM\gamma$. We can associate $x\gamma$ with the usual Mincer-like earnings regression in special cases of the model.

2.2 Workplaces

The workforce uses a technology $Q\left(X\gamma; C\right)$, which expresses the value of per-worker output net of a fixed percentage of revenue taken by the owner of the technology. The technology can depend on workplace characteristics not embodied in the workforce (such as industry and location) and contained in the vector C. Total net revenue generated at a workplace equals $N \times \{\text{per-worker revenue}\} = NQ\left(X\gamma; C\right)$.[6] The technology requires that

[6] The interpretation of $Q()$ is discussed further in the section on bargaining.

Figure 1. Workplace Talent is a Mixture of Worker Talents



workers be assigned to one of two tasks. Task assignments are conditioned on a, so a worker spends some time in both tasks.[7]

Task 1 can be interpreted as primary production and task 2 as managing or secondary production. The tasks are ordered recursively as in Rosen (1982).[8] Primary production involves no interaction with other workers and relies only on the worker's talent. Managerial work requires the manager's talent and primary output.

Workers are indifferent to their task assignment. Task 1 output has no value outside the workplace and can be split and combined irrespective of worker identities. Task 2 uses as inputs both talent and task 1 output. Output from other workplaces cannot be used as input to task 2, and task 2 workers only care about the amount of task 1 output they get to

[7] A given worker is not dedicated to one task but switches between tasks according to the realized value of their talent. The proportion time spent managing varies with worker characteristics x_n as it shifts the distribution $g(a; x_n \gamma)$.

[8] It is possible to extend the model to more than two tasks (or levels). Information on assignment within workplace is not always available in administrative data, and without it the value added by allowing for more tasks is unclear.

use. They are indifferent to the talents and identities of the subordinates who produce the task 1 output. Together these assumptions lead to an internal market for primary (task 1) output in which only the overall distribution of talent matters not the number of workers or their individual talent distributions.

Assumption A2: Technology. Let $\phi(a) \in [0,1]$ denote a fraction of the talent a assigned to task 2 and $1 - \phi(a)$ the fraction assigned to task 1. Let $q_1(a)$ be the amount of task 1 output as an input to talent assigned to task 2. Let $I_\beta = 1$ if $\beta \geq 0$ and 0 otherwise. Then:

A2a. A worker with realized talent a in task 1 produces

$$Q_1(a;\beta) = \begin{cases} a & \text{if } \beta \geq 0 \\ a^{-\beta} & \text{if } \beta < 0 \end{cases} \quad = \quad a^{I_\beta(1+\beta)-\beta}.$$

A2b. A worker in task 2 produces

$$Q_2(a, q_1; \beta) = \begin{cases} a q_1^\beta & \text{if } \beta \geq 0 \\ \\ a^{1+\beta} q_1^{-\beta} & \text{if } \beta < 0 \end{cases} \quad = \quad a^{1+(1-I_\beta)\beta} q_1^{(I_\beta - 1)\beta}$$

For all values of β the task-specific technologies are Cobb-Douglas in inputs and additive across workers in the same task. The technology has distinct properties depending on whether β is above, below or exactly zero. For negative β the task 1 technology is concave in a and the task 2 output exhibits constant returns to scale. For positive β the task 1 technology is linear in a and task 2 output exhibits increasing returns to scale. For $\beta = 0$ the task 2 technology does not depend on input from lower levels and both technologies are linear in a.

Task 2 output is added up across workers and scaled by a coefficient $A > 0$ to generate revenue. The scaling coefficient includes the output price and the net contribution of fixed factors of production. Since workers are indifferent to their assignments, we can associate the overall technology with the revenue produced by revenue-maximizing task assignments:

Definition D1: Workplace Arrangements. An *optimal workplace arrangement* is a pair $\{\phi^\star(a), q_1^\star(a)\}$ of measurable functions that maximize the value of output subject to feasibility of task 1 output demand:

$$\{\phi^\star(a), q_1^\star(a)\} = \underset{\phi(a), q_1(a)}{\arg\max} \quad NA \int_0^\infty Q_2\left(a, q_1(a)\right)\phi(a) g(a; X\gamma) da \tag{4}$$

$$\text{subject to} \quad N\int_0^\infty Q_1(a)\left(1-\phi(a)\right) g(a; X\gamma)\, da = N\int_0^\infty q_1(a)\phi(a)\, g(a; X\gamma)\, da. \tag{5}$$

Let $Q(X\gamma; C)$ denote the value of output at the optimal arrangement.

The constraint (5) is simply a condition for the production process to be feasible in that the workplace must allocate talent internally so that the supply of task 1 output equals the demand coming from task 2. Less talented workers have a comparative advantage at task 1 because more talented workers assigned to task 2 are better able to combine their talent with output of others.

Implication I1: Task Assignment. Let λ denote the Lagrangian on (5). Then:

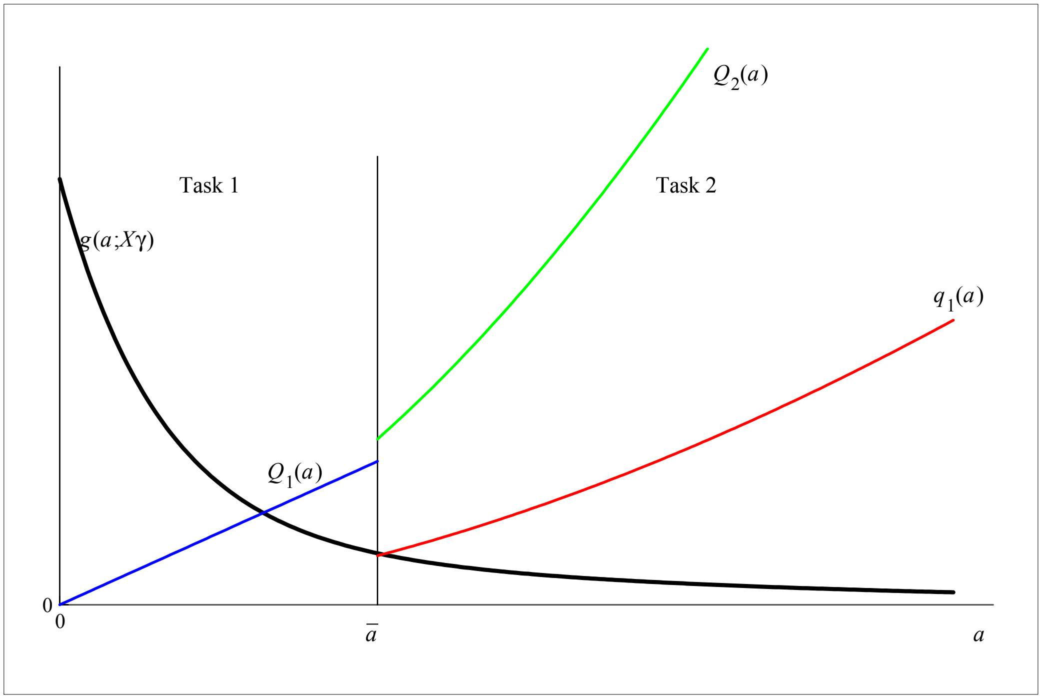
- **I1a.** Optimal task assignment is a cut-off rule. That is, all talent above a number $\bar{a}(\lambda; \beta, A, X)$ is assigned to task 2 and all talent below is assigned to task 1:

$$\phi^\star(a) = \begin{cases} 0 & \text{if } a \leq \bar{a}(\lambda; \beta, A, X) \\ 1 & \text{if } a > \bar{a}(\lambda; \beta, A, X). \end{cases} \tag{6}$$

- **I1b.** Given the technology the cut-off $\bar{a}$ is monotonic in λ and depends on X only through λ.
- **I1c.** For $\beta = 0$ the technology is linearly additive across workers. For $\beta \neq 0$ worker talents interact in determining workplace revenue.
- **I1d.** $Q\left(X_f\gamma; C_f\right)$ is a continuous function of β, A, and γ on their permitted ranges.

The form of $\bar{a}$ under the different cases of β are provided in the Appendix. Any talent devoted to primary output takes away from saleable output, but for $\beta \neq 0$ manager output is constrained by the internal supply of primary output. Holding constant the technology, a larger value of λ indicates that the workforce is more "top heavy," because two effects that move together: first, with greater λ the marginal task 2 assignment moves into task 1

Figure 2a. Optimally Assign Low Talent to Task 1 ($\beta > 0$)



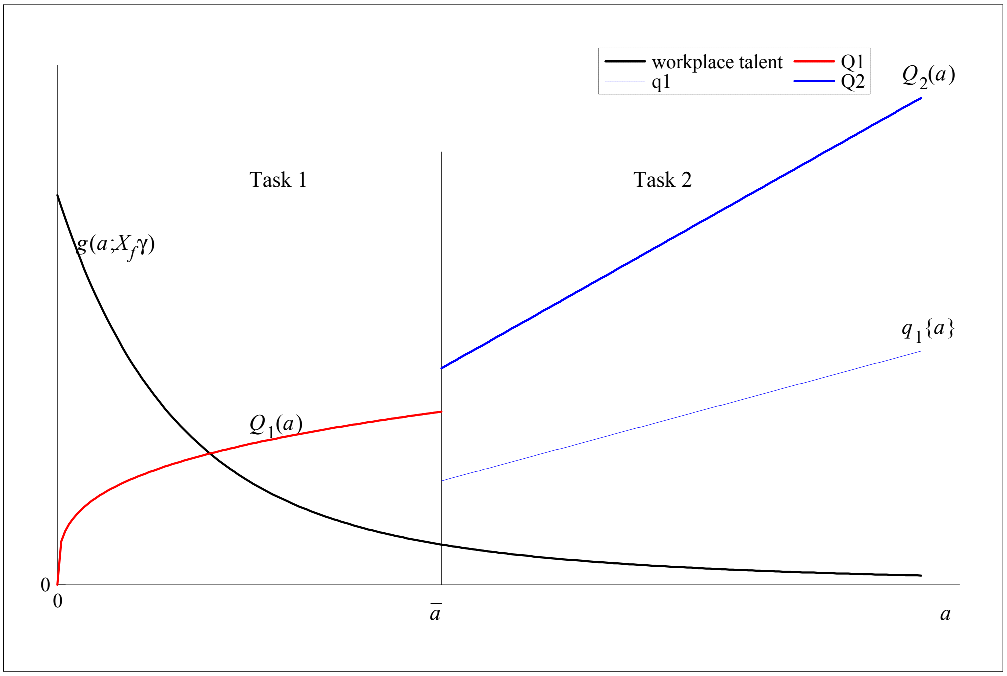
($\bar{a}$ increases); second, each worker working in task 2 faces a higher shadow price for input and gets less task 1 output to use ($q_1(a)$ decreases).

Figure 2a and Figure 2b illustrate optimal task assignment when $\beta > 0$ and $\beta < 0$, respectively. Given the distribution of talent $g(a; X_f\gamma)$, the values of $\bar{a}$ and λ are set to equate total $Q_1(a)$ to total $q_1(a)$. Per-worker output is the expected value of $Q_2(a)$. In the case $\beta = 0$ primary production becomes unnecessary. All talent is devoted to the managerial task. While exponential talent (A1) is maintained throughout the analysis, the first three properties of optimal task assignment require only the technology and a continuous unbounded support of talents. The assignments under each case of β are described fully in the Appendix.

2.3 Marginal Worker Contributions

Adding or subtracting a worker from a workplace changes N and shifts the talent distribution $G(a; X_f\gamma)$. There is a direct impact on primary supply and demand and final

Figure 2b. Optimally Assign Low Talent to Task 1 ($\beta < 0$)



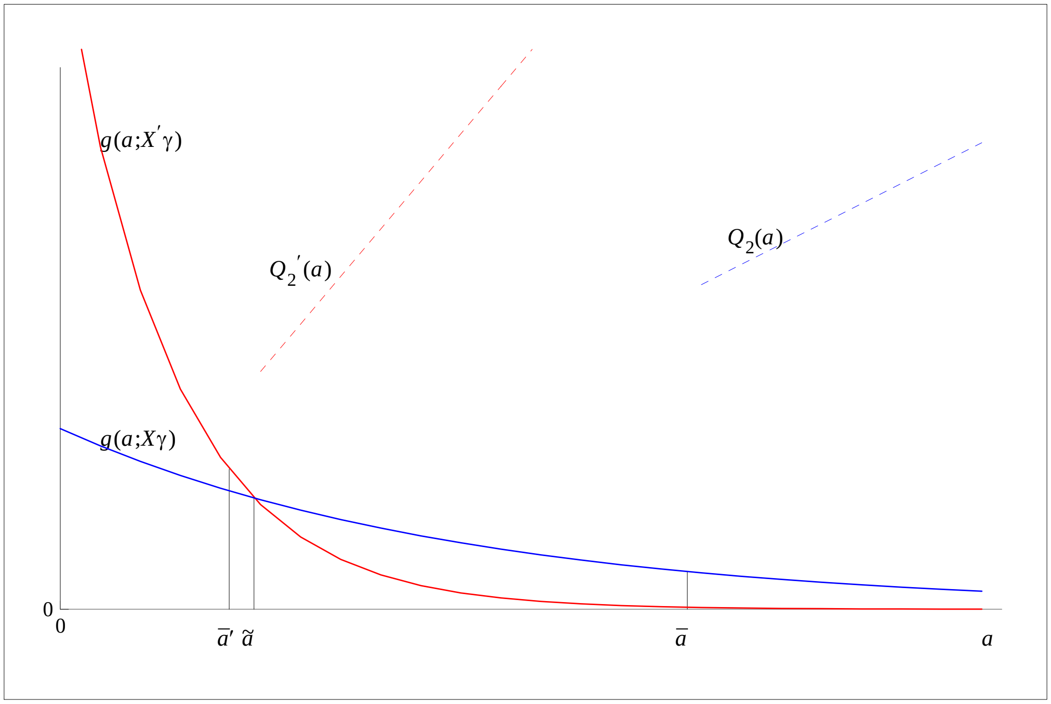
output, even if $\bar{a}_f$ and λ_f were held constant. The new distribution changes optimal task assignment. To describe marginal contribution we introduce some additional notation.

Definition D2: Talent and Value Added. Let $X_f^{\sim x}$ denote the addition (concatenation) of a worker with external characteristics x to workplace f. Thus $X_f^{\sim x}$ is a $(N_f + 1) \times P$ matrix with a last row equal to x. The value marginal product (revenue) of x in f is:

$$\text{VMP}_f(x) \equiv (N_f + 1)\, Q\left(X_f^{\sim x}\gamma; C_f\right) - N_f Q\left(X_f\gamma; C_f\right). \tag{7}$$

For a given workplace a more talented worker is always more productive on the margin than a less talented worker. By how much depends on the existing talent distribution. Although we treat as exogenous the composition of workforces, we briefly consider what the task assignment model says about matching of coworkers. First, when worker talents are complements in the technology (or more generally supermodularity as in Milgrom and Shannon 1994), it is well know that assortative matching occurs in the long run (Becker

Figure 3. Comparative Advantage and Workplace Talent



1973). The most talented workers will work with each other, the least with each other and so forth. However, it is also known that assignment models can break supermodularity (Kremer and Maskin 1996 and Legros and Newman 2002). In the model used here, primary output is complementary to work in the managerial task, but outputs within tasks are substitutes for each other. Since workers spend some time in each task there is a range of overall rates of substitution across coworkers. Workers who do mostly one task are substitutes for each other. They complement workers who spend more time in the other task. Thus, whatever the equilibrium assignment of coworkers is with such a technology it does not exhibit perfect segregation by skill.

Figure 3 illustrates this by comparing the talent distributions of two workforces, X and X'. The former is more talented than the later because its talent distribution stochastically dominates: $G(a;X\gamma) < G(a;X'\gamma)$ for all a. Because workforce X has more talent it has greater internal demand for Q_1. This results in a greater value of $\bar{a}$ (and greater shadow price λ) than for X'. The talented workers in X are forced to spend a large fraction of their time doing basic tasks because not enough co-workers are available to do these tasks. Meanwhile, workplace X' lacks high flyers to transform task 1 output. Now consider adding either a more talented worker with talent index $x\gamma$ or a less talented one, $x'\gamma < x\gamma$. Workplace X' can potentially out-bid X for worker x because hiring the better worker relaxes their constraint on leadership/management talent. Meanwhile, workforce X may prefer to add x' in order to produce task 1 output, freeing up time for their workers to engage in task 2 production. That is, the comparative and absolute advantages of two workers may differ: $\text{VMP}_X(x) > \text{VMP}_X(x')$ and $\text{VMP}_{X'}(x) > \text{VMP}_{X'}(x')$, but $\text{VMP}_X(x) - \text{VMP}_{X'}(x) < \text{VMP}_X(x') - \text{VMP}_{X'}(x')$. The more talented worker has an absolute advantage over the less talented worker regardless of the existing workforce, but endogenous task assignment can give the less talented worker the comparative advantage in a talented workforce. This complexity disappears when coworkers do not interact, as the next result establishes.

Implication I2: Separability. When the technology parameter β equals 0 then $\lambda = 0$. Also, VMP is separable across workers and log-linear in observables: $\ln \text{VMP}_f(x) = \ln A + x\gamma$.

A special case of the technology is the usual log-linear form for VMP that supports the ubiquitous log-linear human capital earnings equation of Mincer (1) . We would arrive at Mincer's equation with the assumption that workers are paid their VMP at their current firm (regardless of worker-workplace specific capital). The joint production model generalizes the linearly separable framework used for most empirical models of wages.

2.4 Wage Determination

Wage determination can be fairly straightforward when talents enter a linearly additive technology shared by all workplaces ($\beta = 0$ and A constant). For example, in a competitive equilibrium with free mobility and no specific human capital ($M = I$) workers are paid their value marginal product. Implication I2 shows that this model generates a Mincer wage equation under these assumptions.

Many considerations make wage determination less simple, including joint production and workplace-specific talents considered here. Now a worker's VMP defined in (7) depends on the talents of their potential coworkers and is computed by re-solving the task assignment problem of the existing workplace. In a joint production economy with exogenous workforces there is no single VMP to determine the wage. One way to determine wages is to consider them the outcome of a multilateral bargaining process for wages within a workplace. To simplify matters, we assume any worker's departure destroys the workplace. In this case, the two-person Nash solution extends to a multilateral situation (Lensberg 1988).[9]

Assumption A3: Bargaining. Wages are determined according to multi-lateral bargaining between the workforce and the employer.

A3a. The employer's bargaining power relative to the workforce is constant across all workplaces, and its outside alternative is zero (shutdown with no scrap value).

A3b. A worker with characteristics x has an outside alternative with value $V(xM\gamma)$, which appears in the vector $V(XM\gamma)$.

A3c. Workers bargain among themselves over their share of the overall surplus. Relative bargaining power depends on workplace-specific talent. That is, the $N \times 1$ vector of weights summing to 1 equals

$$\Pi\Big(X(I-M)\psi\Big) = [\pi_n] = \frac{e^{-X(I-M)\psi}}{\iota' e^{-X(I-M)\psi}}. \tag{8}$$

The $N \times 1$ vector ι contains 1's, and ψ is a vector of coefficients that relate seniority to relative bargaining power.

With employers having zero-valued threat points (A3a) and the total bargaining power of the workforce is constant (A3b), irrespective of the number and characteristics of the

[9] Multilateral bargaining is a complex situation when agents are heterogeneous and they can form sub-coalitions in order to escape agents who contribute less to the surplus (e.g. Krishna and Serrano 1996 and Stole and Zwiebel 1996).

workers, two simplifications occur that make the equilibrium model feasible to estimate on a large data set.

Implication I3: Separable Bargaining. Under (A3) the employer and the workforce each receive the same proportion of total revenue at any feasible workplace. Furthermore:

I3a. Without loss of generality, the technology $Q()$ can be interpreted as the workforce's share of revenue by re-defining parameter A to include factors that account for the workforce's bargaining power.

I3b. The bargaining allocation among coworkers can be separated from the bargaining outcome between the employer and its workforce.

Recall from Assumption A1 that $X\gamma$ is the vector of indices for total talents and $XM\gamma$ is the vector of general talents of a workforce. $X(I-M)$ is the matrix of workplace-specific shifters of talent. The ability to capture surplus is a linear combination of these factors, $X(I-M)\psi$. Given the form of Π the relative bargaining power of workers i and j is $\ln[\pi_i/\pi_j] = [(x_i - x_j)(I-M)\psi]$. Consider two special cases. When ψ is a zero vector then bargaining power is equal across workers; $\Pi = \iota/N$. Because outside alternatives do not depend on seniority a person's wage will be affected by seniority only through the technological contribution of the skills. But this contribution is shared with coworkers through the bargaining process. Now suppose ψ is zero except for the coefficient on seniority. As that coefficient increases it shifts bargaining power to more senior workers. In the limit the vector Π becomes an indicator vector for the worker with the most seniority. All the power accrues to the most senior worker who pays coworkers their outside alternatives and captures the surplus for themselves.

Definition D3: Nash Payroll. Denote the vector of outside values as $V(XM\gamma)$ and its average as $\bar{V}(XM\gamma) \equiv \frac{1}{N}\sum_{n=1}^{N} V(x_n M\gamma) = \iota' V(XM\gamma)/N$. Denote the *total* surplus generated by the workplace as $S(X\gamma, V(XM\gamma)) \equiv N\left[Q(X\gamma; C) - \bar{V}(XM\gamma)\right]$. A workplace is *feasible* if it produces a surplus: $S(X\gamma, V(XM\gamma)) \geq 0$. For a feasible workplace the multilateral *Nash* *payroll* is a vector $W^{\star} = [W_n^{\star}]$ that solves

$$W^{\star} = \underset{\{[W_n]\}}{\arg\max} \quad \prod_{n=1}^{N} \left[W_n - \bar{V}\left(x_n M\gamma\right)\right]^{\pi_n} \quad \text{s.t.} \quad \sum_n W_n \leq NQ(X\gamma; C). \tag{9}$$

Implication I4: The Nash Payroll in a Joint Production Workplace.

$$W^{\star}\Big(X\gamma\, , (I-M)X\psi,\, V\left(XM\gamma\right)\, ; C \Big) = V\Big(XM\gamma\Big) + S\Big(X\gamma\, ,\, XM\gamma\Big)\Pi\Big(X\left(I-M\right)\psi\Big). \tag{10}$$

That is, the Nash payroll is the vector version of the usual 'surplus sharing' result. Four channels generate wage variation as summarized in Figure 4. The Nash payroll depends on technology, talent, outside alternatives (as a function of general talents), and internal talents. In particular, the wages paid to any individual depends on both their own characteristics and the characteristics of their coworkers.

2.5 Outside Alternatives

Since the model is static and assignment of workers to workplaces is taken as given, we make auxiliary assumptions to pin down the value of outside alternatives. We derive an equilibrium value for $V()$ assuming a worker's alternative is to draw at random another existing workplace and join its workforce. Adding the worker shifts the optimal allocation of talents in that workplace. Since there is already an implicit "stage 1" bargain between the workforce and the employer the outside workers has a different status than current workers. This secondary market is a one-time alternative, so a moving worker's outside alternative is assumed to be a constant V_U. If the match with the randomly selected second workplace does not succeed, the worker expects to be unemployed at an exogenous value V_U.

Assumption A4: Hypothetical Alternatives. The economy consists of three stages:

S0. Each workforce f solves its assignment problem. Given $V(z)$ its feasibility is determined.

Figure 4. Wage Variation Within and Between Workplaces

Source	Index	Path to $W^\star$	Type
Internal	$X(I-M)\psi$	Power: $\Pi\,(X(I-M)\psi)$	Relative
External	$XM\gamma$	Alternatives: $V\,(XM\gamma)$	Individual
Total	$X\gamma$	Productivity: $Q(X\gamma; C)$	Joint
Sectoral	C	Technology: β, A	Exogenous

S1. Each worker is hypothetically matched to a randomly selected *feasible* workforce and bargains to join it ignoring other possible transitions. The existing workforce acts as a coalition bargaining with the hypothetical worker who has power $d \in [0, 1]$. The threat point for the workplace is to produce according to the outcome in stage 1. The threat point of the new worker is unemployment with an exogenous value V_U.

S2. Unemployed workers receive V_U. Feasible workplaces from the previous two steps produce. Workers are paid according to the Nash payroll.

This timing links wages with joint production in other workplaces. It uses a friction (only one randomly chosen workplace can be contacted) to avoid the problem of finding the optimal alternative for a given worker. It retains the assumption of bargaining among workers and assumes that arriving workers are evaluated according to their contribution to the workplace's surplus. With $d = 1$ the existing workforce extracts no surplus from a hypothetical worker who attempts to join the workplace, and with $d = 0$ the worker simply gets V_U from any hypothetical match. However we do not reconcile hypothetical bargaining with outsiders and the make up of existing workforces. That is, at some point in the past current coworkers arrived as outsiders and became insiders. That type of dynamic analysis is beyond the scope of this paper.

If all workplaces are feasible then all workers receive at least their outside alternative in their current workplaces. None would strictly prefer to follow through with the hypothetical search step S1.

Implication I5: Define an *equilibrium* as a function $V(z)$ such that all workplaces are feasible in stage S0 of (A4) and no worker strictly prefers a random match in S1 to their current workplace. Let

$$V(z) = V_U + \frac{d}{F}\sum_{f'=1}^{F}\left[VMP_{f'}(z) - V_U\right] I_{\{VMP_{f'}(z)\geq V_U\}}, \tag{11}$$

with $\text{VMP}_f(z)$ defined in (7) and d defined in S1 of (A4).

I5a. $V(z)$ is an equilibrium if all workplaces are feasible under it.

I5b. For V_U and d sufficiently close to 0 any workplace will be feasible under $V(z)$.

In other words, we consider the equilibrium such that the outside threat points of coworkers is their expected payoff from the bargaining to join a random outside workforce. This equilibrium is typically not unique because workplace-specific talent drives a wedge between the VMP in the current and outside workplaces. Coworkers can capture some of the gap without violating other equilibrium conditions. When $d = V_U = 0$ all workers prefer to work in their current firm than search. For a given technology and a given set of workforces finding an equilibrium can be assured by setting the values of V_u and d sufficiently close to 0. Higher values of either parameter change the distribution of wages by shifting wages from the surplus sharing component to the outside alternative component of the Nash payroll.

As with many equilibrium concepts this generalizes the textbook "wage = VMP" result. For example, consider a linear ($\beta = 0$) homogeneous technology, no value of unemployment ($V_U = 0$), full external surplus extraction ($h = 1$), and no internal talents ($M = I$). Then $V(z)$ equals the positive unique VMP of each talent level and the surplus generated by each workplace would be zero. Thus $W^{\star} = V(z) = VMP$ and all workplaces would be just feasible. This special case serves as a benchmark for the joint production technology. In our empirical analysis we allow the technology parameter A to differ across industries. So even with $\beta = 0$ there can be a surplus generated for workplaces in some industries.

2.6 Solution Method and Empirical Considerations

The goal of the empirical analysis is to compare predicted payrolls to observed payroll vectors, denoted W_f^o for $f = 1, \ldots, F$. In the model wages are deterministic. To match the data we introduce measurement error. The observed payroll is the equilibrium payroll $W^\star$ plus an iid normal error vector:[10] $W^o = W^\star + \epsilon$, and $\epsilon \sim N(0, \sigma^2 I)$.

The set of workforce and workplace characteristics is treated as observed and exogenous (as is the masking matrix M). Optimal task assignments and equilibrium outside alternatives are treated as unobserved.[11] We allow the production coefficient A and the exponent β to differ across industries. The vector of free parameters in the most flexible specification is written:

$$\theta \equiv (\, A(C) \quad \beta(C) \quad \gamma \quad \psi \quad h \quad V_U \quad \sigma \,). \tag{12}$$

For a parameter vector θ the optimal task assignment is computed within each workplace, which determines the total share of output available to the workforce, Q. To compute $V(z)$, each workplace in the sample has a randomly selected worker drawn from the whole sample attached to it. That worker's seniority was stripped from X. The result is $z = XM\gamma$. Optimal assignments and output were computed for each workplace, first for the actual workforce and then after inserting the hypothetical worker into the workforce. Letting Q and Q^h denote the actual and hypothetical outputs, the added workers share of the resulting match is computed following equation (11): $v(z) = V_u + d \max\{\mathrm{VMP}_f(z) - V_U, 0\}$.

[10] Typically wages are modeled with a log-linear specification. This creates a tendency to equate the average of log-wages with predicted log-wages. With wages skewed to the right the model would tend to under-predict average wages in levels. However, the model has a theoretical condition on average wage levels within workplaces, not average in logs. Imposing the condition that workplaces be feasible (average revenue exceeds average outside alternatives) may bias the estimate of outside alternatives in order to make up for the shortfall in average wages. So we employ the additive error. This means our linear technology is mostly closely related to a Mincer equation of the form $W^0 = Ae^{x\gamma} + \epsilon$ which would be estimated with non-linear least squares.

[11] This is true for the data used here. When the roles workers play inside their workplaces are available the model generates a probability that a worker with characteristics x is labeled a manager in their workplace equals the proportion of time they spend in task 2, $\exp\{-\bar{a}\exp\{-x\gamma\}\}$.

The result is F observations of z and $v(z)$. These were collected and $v(z)$ was regressed on powers of z:

$$V(z) = \hat{E}[v(z)|z] = b_0 + b_1 z + b_2 z^2 + b_3 z^3. \tag{13}$$

Since F is large, z is a one-dimensional index, and $V(z)$ is monotonic in z, this regression closely matches the conditional expectation of $v(z)$. It retains continuity of $W^\star$ in θ since $VMP_f(z)$ and $v(z)$ are continuous in z which in turn is continuous in θ.

With total output and outside alternatives computed, the log-likelihood for firm f comes from the normal density of the implied error terms: $l_f(\hat{\theta}; W_f^0, X_f) = -0.5/\sigma - \hat{\epsilon}'\hat{\epsilon}/\sigma^2$, and $\hat{\epsilon} = W^o - W^\star(X;\hat{\theta})$. In addition, the equilibrium requires that each workplace generate a positive surplus, $S\Big(X\gamma, XM\gamma\Big) > 0$. This is stringent, especially for small workforces where the characteristics of each worker has a big impact on the talent distribution. On the other hand, without the discipline of a positive surplus the model could predict workers are paid less than their outside alternative. To balance these concerns we penalize the likelihood for negative surpluses in workplaces with more than five workers. The overall objective is therefore:

$$\hat{\theta}^{ML} = \max_{\hat{\theta}} \sum_{f=1}^{F} \Big[l_f(\hat{\theta}; W_f^0, X_f) - D I_{\{S_f<0\}} I_{\{N_f>5\}} \Big]. \tag{14}$$

The penalty D was increased as estimation proceeded. Ultimately no penalty was incurred.

2.7 Identification

As discussed in the introduction, we use our estimated joint production equilibrium to reconsider how wages, seniority and other worker characteristics are related. A worker is affected by their own seniority-driven productivity and the net effect of their coworkers. The seniority coefficients estimated on individual data alone picks up the composite effect of all workplace seniority (even if individual turnover is exogenous). Further, under joint technology and multilateral allocation of the surplus, a worker only gets a share of their contribution to output. The equilibrium restriction isolates the effect of observables on productivity in the current firm from their effect on productivity in other firms that raises outside alternatives.

The share/power effect of seniority is disentangled from the productivity effect through the fact that total payrolls are observed in matched data. Firms with more senior workers should have greater overall payrolls, all else constant. As seniority affects productivity the effect is seen in the total payroll (the sum of the payroll vector $W^{\star}$). Since seniority is left at the workplace door it has no direct effect on the outside alternative $V(z)$, which is estimated through the relationship between payroll and general skill $z = MX\gamma$. The effect of seniority on surplus sharing is seen in the distribution of the payroll (the correlation between $W^{\star}$ and the elements of X related to seniority, $(I - M)X)$). A large representative sample of workplaces and their fully described workforces provides variation in total payroll (the productivity channel) and individual pay (the distributive channel) which separately identifies these effects in equilibrium.

This identification strategy is based on an explicit parameterized model of joint production that follows the literature on task assignment within organizations. The model also includes a nested model of linear, individualistic production which supports the standard wage equation (1) . The complex technology in (A2) ensures the special linear case is in the interior of the support of β. Inference about joint versus individualistic production becomes a standard test of the null hypothesis $\beta = 0$ in the interior of the parameter space. In contrast, working with data on individuals alone without their coworkers leaves total payroll unobserved. It also leaves relative seniority unobserved even with panel data. Thus the joint production parameter β is not identified from that kind of data, and in light of our model most models of wages maintain and leave untested the assumption $\beta = 0$. By the same token, this application of equilibrium task assignment model to matched data set maintains assumptions that can be relaxed with, say, individual-level panel data. So our analysis is a counter-balance to the historical focus on individualistic models of production when asking questions such as whether wages rise with seniority.

3. Empirical Analysis

3.1 Data Overview

3.1.1 Workforces

Beginning with the universe of Norwegian workplaces (both public and private) and people aged 16-75, attachment to an employer is based on the person's status as of November 30, 1997. Matching is based on a personal identification number assigned to all residents of Norway. Employers are identified with a unique number for the firm and a unique tax number for the plant or establishment. Because the current analysis is static, there is little concern here whether there are spurious workplaces created or destroyed in the administrative data.

3.1.2 Industry

The vector of workplace characteristics C is simply an indicator vector for the industry of the workplace, which is a variable for each worker. In some cases the industry code is missing, and in others workers in the same workplace are coded with different industries. The industry of the workplace is defined as the mode industry associated with its workers. If more than one mode exists, or if the industry code is missing completely the workplace is placed in a separate "no code" category. Some smaller industries are combined with larger ones to define 8 distinct categories listed in Table 1.

3.1.3 Earnings

Administrative data related to public pension credits record total annual 1997 earnings in all jobs and any unemployment insurance benefits. Denote this amount as EARN, expressed in thousands of 1997 Norwegian kroner, approximately US$150 in 1997. Let n^u and n^e denote months of unemployment and full-time education in 1997 (merged in from another administrative database). UI benefits received are approximately 0.6 times monthly earnings. And we assume that monthly earnings on any jobs held in 1997 either earlier or later than the November job are the same as earnings on that job. These assumptions and approximations imply EARN = $(12 - n^u - n^e)W^o + 0.6n^eW^o$. Then monthly earnings on the November job are

$$W^o_n \equiv \frac{\text{EARN}}{12 - n^e - 0.4n^u}.$$

For the vast majority of workers W^o is simply one-twelfth of their total 1997 earnings on their November job. The other assumptions come into play only if short jobs were held before or after the November job and only if the worker left school or spent some time unemployed during the year.

3.1.4 Hours, Experience and Seniority

Usual work hours per week are based on the data reported to the national insurance authorities mainly for sick-leaves and calculation of unemployment benefits. Work hours are described by three categories, and the worker's characteristic include an indicator for full-time.

Besides an indicator for females, the other elements of x consist of indicators for categories of years of schooling, merged from another administrative database (that actually contains detailed six-digit codes for both type and amount of education). The row of characteristics has $P = 16$ columns corresponding to the variables listed in Table 2. The external vector has 6 columns of zeros knocked out by M.

3.1.5 The Sample

In the original data, 1,719,983 people are associated with an employer. The sample is reduced by eliminating workers (and their workplace) with inconsistent job start dates, extreme earnings ($W^o < .1$ and $W^o > 1000$), and other missing variables. This eliminates 10,254 workplaces and 452,230 people. Next all workplaces with a single worker are eliminated, which eliminates 38,533 workplaces/observations. The result is 1,229,219 people working at 103,840 workplaces. From this a 20% sample of workplaces is drawn, based on 20,542 workplaces and 247,521 workers. The typical worker has 11 coworkers. A second 20% sample was drawn to be used for out-of-sample comparison of the equilibrium predictions of various model specifications.

Table 1 summarizes the data by workplace. The number of "no-code" workplaces is small except for very small workplaces for which conflicting industry coding is likely to occur. Small workplaces are the norm in all industries. Only in manufacturing and services is the percentage in 51-100 range even close to 10%. The last column shows that over half of all multi-worker workplaces in Norway have 2-5 workers and nearly 90% have 20 or fewer. For understanding the technology of joint production small workplaces are potentially very important.

Table 2 summarizes worker characteristics. Average monthly earnings are nearly NOK 19000 with a coefficient of variation of 65%. Women make up 45% of the workforce, and 76% of workers work full time. The average worker has 13 years of actual experience and 5.43 years of potential seniority. If full-time is a permanent status in a workplace then current full-time workers have acquired 4.434 / .76 = 5.84 years of seniority on average. Part-time workers have acquired only 4.15 calendar years of seniority (computed from the other numbers in the table). The modal education category is 10 or 11 years followed closely by 12 or 13 years.

Table 3 shows the joint distribution of industry and selected worker characteristic with common patterns. Services are dominated by women and have the lowest proportion of full-time workers. Construction is dominated by men and full-time workers. Agriculture/Mining/Elect. has the longest seniority, but transport has the most experience but the least seniority. FIRE and Services have the most educated workers.

Table 4 examines the variation in coworker characteristics within and across workplaces. For each worker the mean among coworkers was computed for selected variables. The variation in these coworker means was then decomposed between and within workplaces. In each case variation within workplaces is much lower than between workplaces. Not only do earnings vary more across workplaces than within, so do education, seniority, experience, full-time status and sex. The fact that between-workplace standard deviations are greater than the overall values reflects the positive correlation within workplaces, which is reported directly as the correlation between the worker value of the variables and their coworkers' mean. For example, the correlation of .54 for the female indicator shows that workplaces are partially segregated by sex. All the correlations are strongly positive. This partial assortative sorting suggests that models of earnings not based on matched data overstate the direct impact individual characteristics have on VMP.

3.2 Estimates

Table 5 reports estimates of two versions of the linear technology which sets the key parameter β to 0. They are the equivalent to a Mincer wage regression because they implement the usual assumption that workers are paid their VMP. This happens because they capture the whole share of their outside surplus ($h = 1 = 1 - V_U$) and the mask matrix M equals the identity matrix which means that all characteristics including seniority are related to general skills. Or, an equivalent interpretation is workers capture all the surplus from their seniority and employers capture none even though there specific nature of these skills means there is no competitive pressure on it. The difference between the columns is whether the external market is the whole economy (A constant) or the workplace's industry (A variable).

The estimates of γ follow the typical pattern of a wage regression. Women earn less than men; earnings are concave in experience and concave in seniority with a smaller range. Wages are slightly less sensitive in seniority for women and highly educated workers. The return to education implied by the coefficients on the categories is also typical. For example, using the mid-points of the 12-13 and 14-16 year categories yields a return to one year's schooling of 0.08. The model that restricts industries to have equal output coefficients is rejected with a χ^2_7 log-likelihood ratio of 7346=-2(694252-694579).

Table 6 reports estimates of our preferred specification of the joint technology model in which industry-specific βs and the bargaining power shifters ψ are estimated but internal talent (seniority) is assumed to have no affect on productivity. This model fits the data significantly better than the linear model (the log likelihood ratio test statistic is over 2000.) Overall many of the parameters follow patterns similar to the linear model. Many of the standard errors are small but of the same order of magnitude as those on regression coefficients on the same large data set.

3.2.1 Power or Productivity?

Table 7 summarizes five different specifications including the preferred specification in Table 6 and the unrestricted linear model in Table 5. Next to that model in Table 7 is a joint technology model in which bargaining power is constant. Each worker gets an equal share of the surplus. Allowing for interactions in the workplace but keeping power equal significantly improves the fit over the unrestricted linearly separable specification. The log-likelihood ratio statistic is 2473.5. The next column in Table 7 allows the linear seniority terms in the bargaining vector ψ to be non-zero. Freeing these four parameters improves the likelihood (test statistic is 1277.9) while not changing the seniority/productivity parameters greatly. Next comes the power-only specification already presented. This specification does not nest the linear power model since seniority has no productive aspect. It has four fewer free parameters yet results in an improved fit, which is statistically significant as the likelihood ratio is 515.7. This suggests that seniority's role in total output is very limited. Eliminating it altogether while allowing relative power to be non-linear is preferred.

Finally the last column presents the model that frees up all the parameters. This specification nests the other joint technology models and accordingly has the best likelihood value. The improvement over the preferred specification is very modest given the enormous sample size and the steady changes when freeing up other sets of parameters. The χ^2 statistics for the test of the preferred model is a mere 32, which is formally significant at the 1% level (the critical value with six degrees of freedom is 16.81). However, the parameter values for the nested model are somewhat problematic. For fulltime workers years with the workplace affects talent adversely throughout the career (the sum of the both the linear and quadratic components of γ are negative). It is difficult to argue that specific talents actually detract from total output throughout the career. A simpler explanation is that the data do not provide enough variation in total payroll and payroll distribution to separately identify the productive and distributional roles of seniority. Perhaps a more restrictive specification would provide coherent estimates of both effects, but the two cases presented in Table 7 suggest that seniority would still have a very small productivity effect. So we retain as our preferred specification that the seniority-related coefficients in γ are set to zero but ψ is estimated.

Further evidence against the nesting model in the final column of Table 7 is provided by applying estimates to a separate 20% sample of workplaces. The parameters and the equilibrium value of $V(z)$ from the estimation sample are used to compute the Nash payroll for each workplace in the holdout sample. The same likelihood comparisons are made for the holdout sample. The results are similar test statistics, but the linear technology and linear seniority specifications fit worse in the holdout sample. The equal power and preferred specification fit better (than in the estimation sample). In particular, the likelihood ratio test does not reject the preferred specification in the holdout sample. This suggests that the joint technology provides a real improvement over the linear model, but allowing seniority to be both productive and distributive leads to overfitting because the productive channel is extremely weak. Based on this evidence we now focus on the preferred specification with joint production and seniority-related bargaining power and how its explanation of the data compares to the rejected but more common linear specification.

3.2.2 Variation Within and Between Workplaces

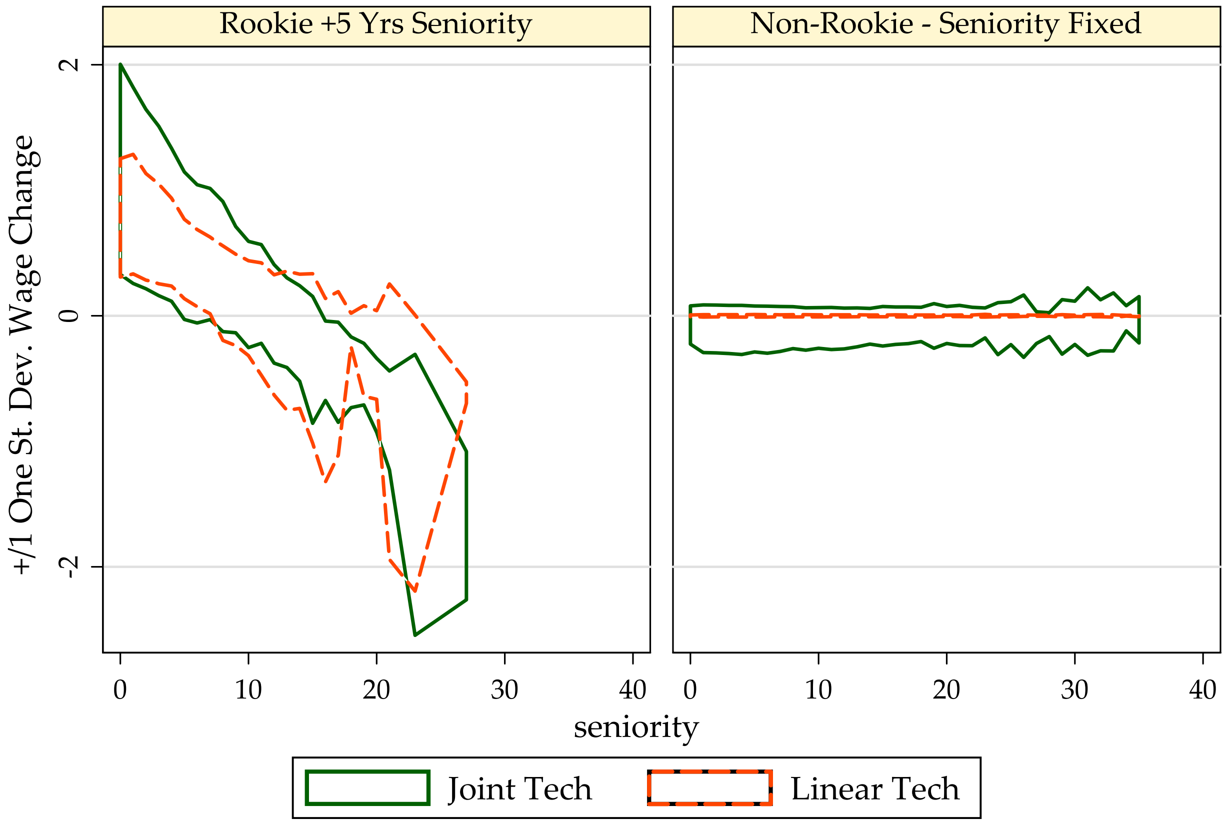
Table 8 compares the variance of wages in the data, the preferred joint estimates, and the linear technology estimates. As with all wage regressions a sizeable fraction of wage variation is unexplained (and is accounted for by measurement error in this analysis). Both models have predicted wage variation of about 58% of that in the data with the better fit of the joint model amounting to about .004% change. This is not surprising since the joint model introduces only a handful of new parameters to fit a quarter of a million observations. Despite ending up with similar overall variances the two models apportion it quite differently between and within workplaces. The joint model attributes a higher amount of variance between workplaces and less within. Within-workplace variation includes variation in the external returns to general skills ($V(z)$) and variation across workers in their share of the surplus ($S\pi$). The linear model creates a surplus only through inter-industry technology differences. So all within-workplace variation is due to $V(z)$. Each of the variances of $V(z)$ in the joint model are about 80-85% of the corresponding values in the linear model. The joint model attributes less variation to external factors, leaving surplus dividends to explain the rest.

3.3 The Wage Distribution and Technology

3.3.1 Seniority when Coworkers Matter

To illustrate the role of seniority we computed the response to increasing seniority by five years for one worker in each workplace. The worker chosen was the one with the lowest amount of seniority (the relative rookie). This change affects productivity and surplus in the linear technology and bargaining in the joint technology estimate. With joint production the rookie's wage change depends on their characteristics and their co-workers. There is a spillover effect for coworkers because their relative seniority is changed. The first panel of Figure 5 shows the distribution of impacts on the rookie whose seniority changed. It is represented by the one standard deviation ban around the mean change (over workplaces) as a function of the rookie's initial seniority. For

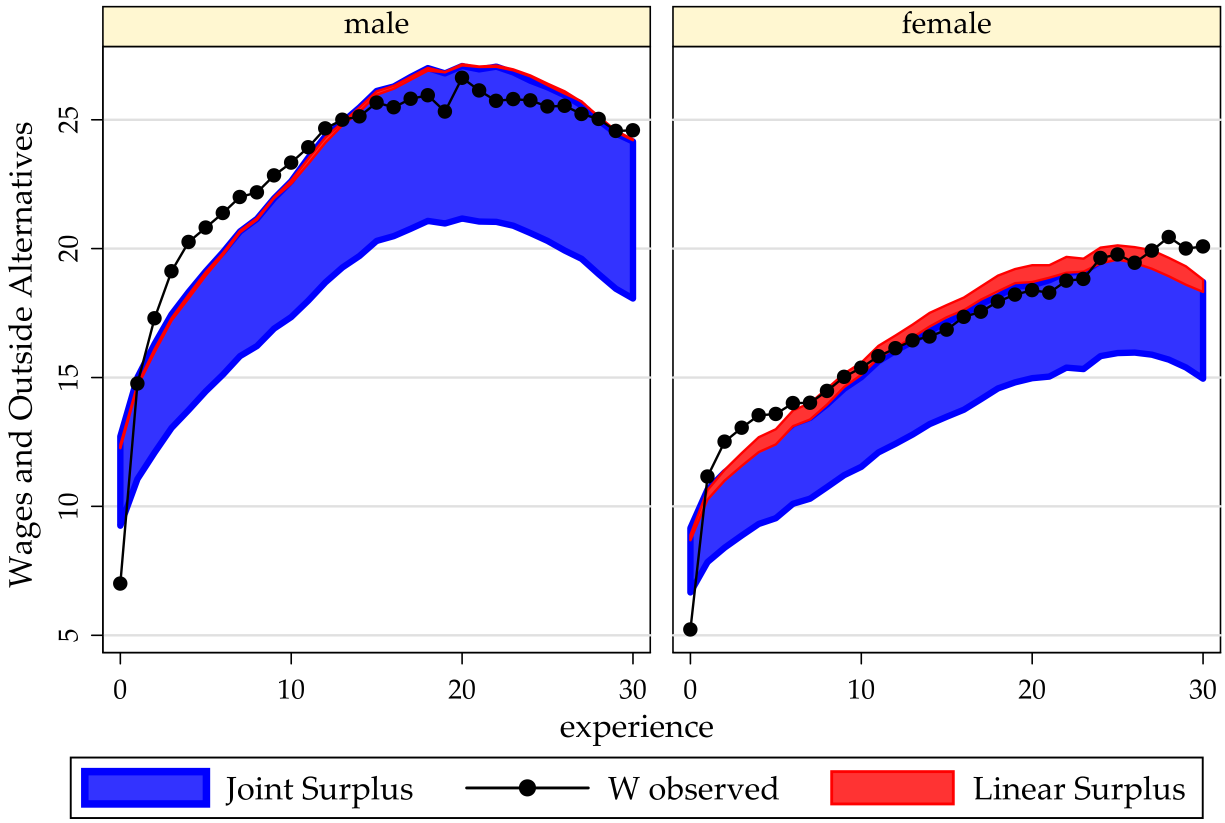
Figure 5. The Effect of Seniority and Technology Assumptions



Equilibrium response to adding 5 years seniority to the lowest ranked worker in each workplace. The curves show the ± one standard deviation bands around the mean differences in predicted wages.

these workers the typical changes are not that much different under the two joint and linear model. The variation is a little larger for new workers under the joint technology but smaller for old workplaces (in which the rookie already has substantial tenure). The downward trend reflects the concave seniority profiles under both technologies. The second panel of Figure 5 shows the spillover to other workers whose own seniority is fixed but their relative bargaining strength changes. This effect is much smaller that the direct impact, because it is spread over $N-1$ coworkers. But we see a much greater range under the joint technology then the linear one, in which only the industry-specific surplus is available for sharing.

Figure 6. Gender Wages Profiles and Technology



3.3.2 Gender Differentials

Now consider how the joint technology assumption changes the explanation for differences in average wages between men and women. The linear estimates in Table 5 attribute a 20% difference in productivity between men and women, all else constant.[12] Women also benefit from seniority less than men (a significant but not large difference). The coefficients on the interaction between experience and gender is not significant (recall that we use a measure of actual experience based on pension points. The joint technology estimates are not that different in magnitude, except the seniority effect in ψ is larger, and the experience differential is now significant.

However, Figure 6 shows that the these differences feed through to wages through

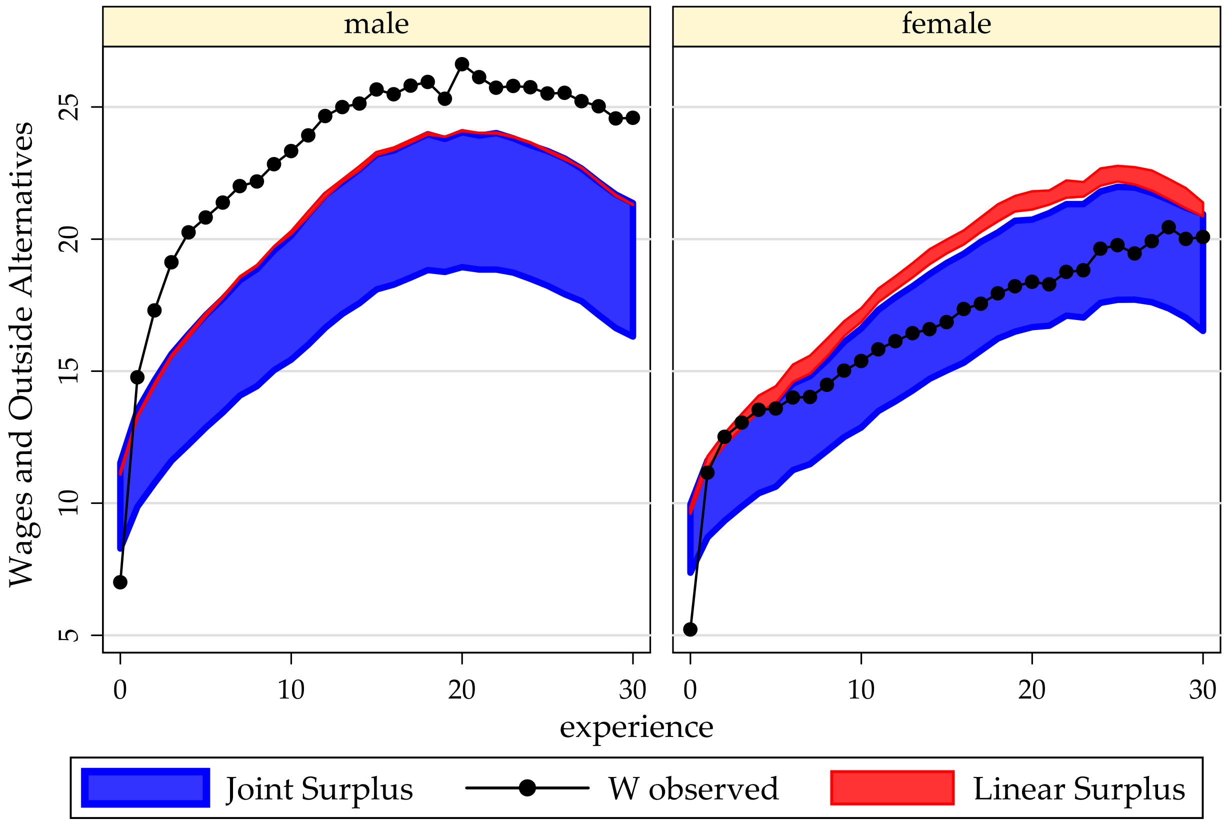
[12] Hægeland and Klette (1999) also study gender and experience wage differentials using measures of productivity available for a subset of the firms in our broader sample and a linearly separable interpretation of productivity.

different channels. The figure shows experience profiles for men and women (so both the direct effect of gender and differences in other characteristics are in play). The connected lines are observed average wages and the wage differential is apparent. The top of the blue shaded areas are the average predicted wages under the preferred joint technology estimates. For both men and women the model predicts slower growth when experience is low than in the data. The bottom of the blue shaded area is the average value of $V(z)$ for the people of that gender and experience level. We see that the profile of outside alternatives for men grows much quicker than for women and peaks in mid career. The size of the shaded area is therefore the average share of workplace-specific surplus. While outside alternatives account for most wage growth the share component of total wages grows for men over their careers. For women the share is nearly constant. This is due to the smaller effect of seniority for women. By contrast, the linear technology estimates have almost no surplus, and that which exists is solely due to industry differentials. That gap is shown in red in Figure 6 on top of the surplus from the joint technology surplus. The somewhat larger surplus for women is then due to industry differences and concentrations in Table 3. Thus, the joint technology estimates attribute some of the gender gap to workplace politics biased against women, the effect of which is most prominent in mid career.

Figure 7 shows the impact of gender differentials by considering a counterfactual that gives each worker a gender value of 0.5. That is, it makes the workforce neuter, raising talent for women and lowering it for men. Under a linear homogeneous technology overall mean log-productivity would be quite similar, except that women made up only 45% of the workforce. Under joint technology the whole workplace talent distribution shifts and the optimal task assignment changes. We compute the equilibrium response of $V(z)$ and display the same values as in Figure 6. The resulting experience profiles follow similar shapes and reduces but does not eliminate gender differences under both technologies.

The differences do not disappear since men and women differ in other characteristics

Figure 7. Equilibrium Effect of a Neuter Workforce



and these are left constant. For men the predicted values are nearly identical under the joint and linear technology. However, for women a visible gap between the linear and joint wages appears, with joint wage predictions below the outside alternative $V(z)$ under the linear technology. This is somewhat unexpected. At one level women should gain more than men from neutering the workforce. Women become more productive and tend to work with more women than men. Men become less productive and tend to work with more men than women. Thus without industrial differences in technology one would expect spillover affects to push female wages in the counterfactual above the linear wage profile. This does not happen because women tend to work in industries where there is less interaction between coworkers. Thus they gain less from the spillover effects than men lose (on average).

3.3.3 Workplace Size Differentials

Figure 8. Wages and Workplace Size

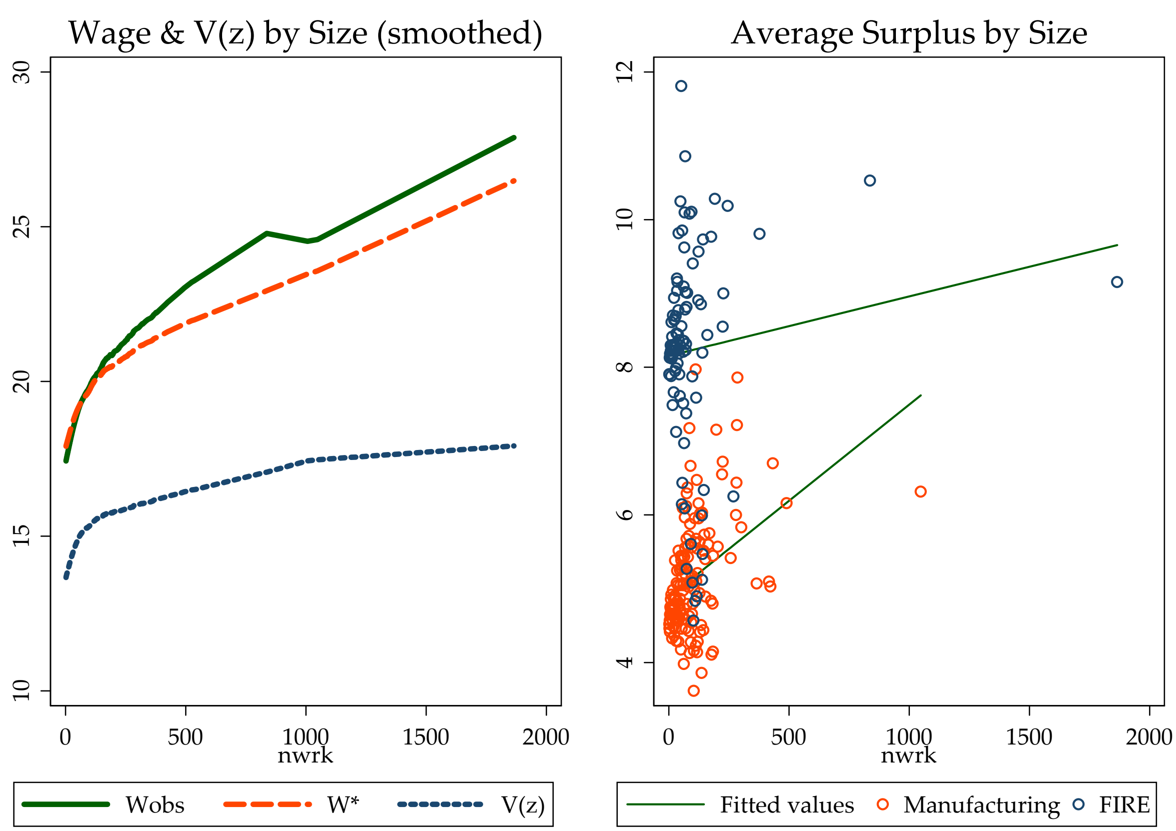


Figure 8 displays another aspect of the wage distribution. For the preferred specification, predicted wages ($W^\star$) and outside alternatives $V(XM\gamma)$ are averaged by workforce size, N_f. The resulting wage-size profile is displayed after smoothing. Recall that the model exhibits constant returns to scale in N_f. The predicted profile tracks the observed rising profile. This explanation of the firms size wage profile (Oi and Idson 1999) is partly based on differences in observable characteristics between large and small workforces (education and experience included). Changes in observed characteristics of individuals are captured by the slope of $V()$. Outside alternatives rise quickly for workers in small workplaces but flatten out more quickly than payrolls. But the displayed profile of outside alternatives shows the gap goes beyond this component. The gap between wages and alternatives, the average surplus, accounts for most of the profile beyond 100 workers. Part of this rise in the surplus is due technological differences across industry. Industries with larger workplaces tend to have larger values of A which is reflected in the surplus. As discussed earlier the value of A includes not just the contribution of other factors of production but any differences in the bargaining strength of workforces relative to employers. While our model does not exclusively model collective bargaining, but it does provide a surplus in equilibrium to be bargained over collectively.

Explaining the size profile by component of worker characteristics, technology and collective bargaining are not novel. But the estimated model provides one more component: better matching of talent to technology in larger firms. The second panel in Figure 8 shows the surplus ($W^{\star} - V$) by firm size for Manufacturing and FIRE. The fitted values from a regression interacting industry with the intercept and size are also shown. (The null hypothesis that the industry-specific size profiles are zero is rejected, $F_{8,747} = 11.93$.) The larger intercept for FIRE reflects a greater value of A, but the size profiles and their differences are due to both technology and coworker interaction. The profile is flatter in FIRE than in Manufacturing. Recalling that FIRE has an estimate of β near 0 there is much less scope for synergies between coworkers to increase surplus. The estimated β in Manufacturing is larger and allows for synergy. The fact that the trend is upward sloping means that the mixture of talents in larger workplaces tends to be better suited to the technology than in smaller workplaces.

3.3.4 A Possible Extension: Workplace Dynamics

Our model could be extended to form the basis of a dynamic analysis of matched panel data. We treat the current workers as given, but clearly a workforce at a point in time is a lagged endogenous outcome. The model generates reasons why workers would move from one workforce to another: as their skills and their coworker skills evolve they may match up better with a different workforce. The increased productivity due to that better match is spread among all new coworkers but if it outweighs the loss from the current firm the worker can be enticed to move. This worker loses seniority and hence bargaining power, but the estimated model allows for a bargaining parameter (d) specific to outside workers negotiating to join a workplace. In the model estimated here external workers capture almost 70% of the surplus from joining an outside firm. In this static analysis this parameter is not estimated on actual moves but rather is used to tune hypothetical outside alternatives $V(z)$. A dynamic model would have to complete the transition between arriving and existing workers.

Workers of any talent may find better matches, but workers with different talents will not always agree on their preferred destinations. As discussed earlier, high talent workers may be attracted to low talent workforces and vice versa. In a dynamic context, mobility slows down with seniority in the preferred model not to avoid loss of specific talent but loss in specific bargaining power. The joint production technology suggests that some of the observed wage-size profile is due to a correlation between workforce efficiency and firm size. A dynamic model might amplify this effect since small firms that happen to attract a good mix of workers are more attractive to outsiders than firms that have a bad match between technology and talent. Size grows not for scale reasons but to exploit early advantages which can persist.

4. Conclusions

This paper considers an alternative to equating a worker's wage as their VMP defined independently of their coworkers. To gain traction on this goal some of the lessons from research based on that assumption have been ignored. For example, we conduct a cross-sectional analysis treating the current characteristics of all workers as given, including their experience and seniority. Our model of coworker interaction is based on the task assignment model of production. Unlike most previous applications of task assignment models, our approach generates a firm-specific surplus that must be allocated among all coworkers. Equilibrium wages must also account for variation in an individual's value marginal product across outside workplaces. A multilateral Nash bargaining solution provides a generalization to the standard linearly separable model. Our specification makes it feasible to impose restrictions of the model on over 20,000 individual workplaces at each point in the estimation procedure. We parameterize the technology so that the linear technology is a special case, which is rejected with a conventional likelihood ratio test in favor a joint technology. The linear case is also outperformed in an out-of-sample validation.

Thus, accounting for joint production provides a viable alternative explanation of the data. The model adds only eight free parameters to the linear model, so the difference in fit is highly significant but not strikingly different. However, joint production provides a quite different explanation. Workplaces with more overall seniority are not more productive, in the sense that they support greater total payrolls. But within workplaces more senior workers get a larger share of the surplus than their external talent justify. The model attributes this to more bargaining power due to relative seniority. Relative seniority is not observed in data on individuals. Our estimates exploit the feature of matched data sets that total payroll is observed. Our analysis explains mixed results from previous research on seniority wage profiles through a weak (non-existent) productivity effect and a relative seniority effect. We have also shown that taking coworkers seriously can affect the interpretation of gender and firm-size wage differentials.

5. Appendix

Proof of Implication I1

I1a. In the absence of a cut-off rule, the workplace can allocate a fraction of the density of talent at level a to each task. Let $\phi(a) \in [0,1]$ denote the fraction assigned to the higher level 2. Then the Lagrangian can be written:

$$\begin{aligned}\mathcal{L} = \max_{\phi(a),q_1(a)} & A\int_0^\infty \phi(a)Q_2(a,q_1(a);\beta)g(a;X_f\gamma)da \\ & + \lambda\left[\int_0^\infty (1-\phi(a))Q_1(a;\beta)g(a;X_f\gamma)da - \int_0^\infty \phi(a)q_1(a)g(a;X_f\gamma)da\right] \\ & + \mu_0(a)\phi(a) + \mu_1(a)(1-\phi(a)).\end{aligned}$$

Given that $g(a;X_f\gamma) > 0$ for $a > 0$, the first order conditions for $q_1(a)$ can be re-arranged as

$$\lambda\phi(a) = \begin{cases} A|\beta|\phi(a)a^{1-|\beta|}[q_1(a)]^{|\beta|-1} & \beta < 0 \\ A\beta\phi(a)a[q_1(a)]^{|\beta|-1} & \beta > 0. \end{cases}$$

In either case the equation can be satisfied with $\phi(a) = 0$, or with $\phi(a) > 0$ and $q_1^\star(a)$, given as

$$q_1^\star(a;\lambda) = \begin{cases} \left(\frac{A|\beta|}{\lambda}\right)^{1/(1-|\beta|)} a & \beta < 0, \\ \left(\frac{A|\beta|}{\lambda}\right)^{1/(1-|\beta|)} a^{1/(1-|\beta|)} & \beta > 0 \end{cases}$$

The first order conditions for $\phi(a)$ can be written

$$\mu_1 - \mu_0 = \begin{cases} [Ka - \lambda a^{|\beta|}]g(a;X_f\gamma) & \beta < 0 \\ [Ka^{1/(1-\beta)} - \lambda a]g(a;X_f\gamma) & \beta > 0 \end{cases}$$

where

$$K = A\left(\frac{A|\beta|}{\lambda}\right)^{|\beta|/(1-|\beta|)} - \lambda\left(\frac{A|\beta|}{\lambda}\right)^{1/(1-|\beta|)}.$$

For interior solutions such that $0 < \phi(a) < 1$ the right hand sides are zero since $\mu_0 = \mu_1 = 0$. From A1 $g(a;X_f\gamma) > 0$, so the left hand side is zero only when the difference is zero. In both cases the difference is between a straight line through the origin and a positive power of a. Thus it is only zero for at most one point $\bar{a} > 0$. In both cases the difference begins at 0 for $a = 0$, goes negative and reaches 0 again at $a = \bar{a}$, then becoming positive. Thus, for $a < \bar{a}$ it must be that $\mu_0 > 0$ and $\mu_1 = 0$ and hence $\phi(a) = 0$. For $a > \bar{a}$, $\mu_0 = 0$ and $\mu_1 > 0$, and hence $\phi(a) = 1$. This proves that the optimal assignment of talent is (6). Solving for $a = \bar{a}$,

$$\bar{a} = \begin{cases} (\lambda/K)^{-1/(1+\beta)} & \beta < 0, \\ (\lambda/K)^{(1-\beta)/\beta} & \beta > 0. \end{cases}$$

I1b. From the expression for $\bar{a}$ we see that X_f only enters through λ. Monotonicity can be seen by rewriting the internal demand=supply constraint as

$$\int_0^{\bar{a}} Q_1(a;\beta)g(a;X_f\gamma)da = \int_{\bar{a}}^{\infty} q_1^{\star}(a;\lambda)g(a;X_f\gamma)da.$$

Holding $\bar{a}$ constant while increasing λ, the left hand side is constant. The right hand side is increasing (since $q_1^{\star}(a;\lambda)$ increases with λ for all a). Hence an increase in λ must be offset by an increase in $\bar{a}$, proving monotonicity between the two elements of the optimal workplace allocation.

I1c. When $\beta = 0$, then $Q_2(a, q_1; 0) = a$. Total production, $NQ\left(X\gamma; C\right) = A(\iota' E[a|X])$, is invariant to q_1 so it is optimal to allocate no talent to task 1 and workplace output is linearly separable across workers.

I1d. To solve the assignment problem one integral enters a single non-linear equation in λ. The root found using bi-section and Newton iteration. Continuity and monotonicity imply that the solution to λ is also continuous in the parameters as are workplace revenue and other aspects of the workplace.

I1e. For $\beta \neq 0$ continuity of the model's predictions is straightforward. The technology in A2 is smooth in β and all other parameters. Then $\bar{a}$ is continuous because it is the unique solution to a non-linear equation that varies continuously in the parameters and has a non-zero Jacobian everywhere. Given $\bar{a}$ the solution for λ is similarly continuous. Continuity extends through all integrals because the bounds and the integrands are continuous in the parameters, λ, and $\bar{a}$.

At $\beta = 0$, continuity is slightly complicated because the technology in A2 is continuous (but not differentiable) in β. Approaching 0 from either direction the technology is continuous and bounded. From below we see $\lim_{\beta\uparrow 0}\lambda \to 0$ and $\lim_{\beta\uparrow 0}\bar{a} = 0$. The contribution of task 1 to integrals approaches $\int_0^0 0^0 f(0)$. The density (3) is bounded at 0 under (A1). Thus the limit is 0. In other words, task 1 output goes to 0 even though output at exactly $a = 0$ is unbounded. For $\beta > 0$, it is difficult to prove what the limit of λ is as $\beta \downarrow 0$, or even if it has a limit. However, predictions only depend on λ and $\bar{a}$ through total output. Note that total supply of task 1 output is bounded by $K = (\iota' E[a|X])/A$ total output at $\beta = 0$ divided by A. Therefore total output must be below the value of giving each person assigned to task 2 K units of task 1 output. Substituting this into total output results in an upper bound for output of $AK^{\beta} \int_{\bar{a}}^{\infty} ag(a)da \leq A^{1-\beta} \left[\iota^{E}[a|X]\right]^{\beta+1}$. Optimal output must be below this upper bound which converges to separable output, $A(\iota' E[a|X]$, as $\beta \to 0$. This is feasible because it can be achieved by letting $q_1(a) \to 0$ and $\bar{a} \to 0$ as well. Thus, $NQ\,(X\gamma; C)$ is continuous at $\beta = 0$ and therefore continuous for $\beta \in (-1, 1)$.

Proof of Implication I3. Let R equal total revenue including the employer's share; let η denote the employer's bargaining power, and let $\pi_n^\star$ denote the power of worker n. The sum of the workforce parameters is $1 - \eta = \sum_{n=1}^{N} \pi_n^\star$. Let P denote the employers profit and W_n the worker's salary. Then the canonical Nash bargaining problem among the $N + 1$ agents can be written:

$$\max \quad P^{\eta} \prod_{n=1}^{N} (W_n - V_n^\star)^{\pi_n^\star} \quad \text{subject to} \quad P + \sum_{n=1}^{N} W_n = R.$$

The solution for the employer's share in a feasible workplace is $P = \eta(R - \sum V_n^\star)$ and the workforce as the whole receives $\sum W_n = (1 - \eta)(R - \sum V_n^\star)$. We can then consider the sub-problem of allocating this across workers. The problem can be written

$$\max \quad \prod (W_n - V_n^\star)^{\pi_n^\star/(1-\eta)} \quad \text{subject to} \quad \sum W_n = (1 - \eta)(R - \sum V_n^\star).$$

Then define $V_n = V_n^\star/(1 - \eta)$, $Q = R/(1 - \eta)$, and $\pi_n = \pi_n^\star/(1 - \eta)$. We then arrive at a problem equivalent to (D3).

Proof of Implication I4: See Lensberg (1988)).

Proof of Implication I5.

I5a. Under $V(z)$ no worker in a feasible workplace will prefer to leave and take a random match to another workplace.

I5b. For $V_U = 0$ and $h = 0$ $V(z) = 0$ for all z. All workplaces are feasible and all workers receive a positive share of the revenue, which also equals the surplus. Thus $V(z)$ is an equilibrium. $V(z)$ is continuous in V_U and h, so for some range of values above 0 all workplaces stay feasible and the corresponding $V(z)$ remains an equilibrium.

6. References

Altonji, Joseph and Nicholas Williams 2005. "Do Wages Rise with Job Seniority? A Reassessment," *Industrial and Labor Relations Review* 58, 3, 370-397.

Buhai, Sebastian I., Miguel Portela, Coen Teulings, and Aico Van Vuuren 2008. "Returns to Tenure or Seniority?" Tinbergen Institute Discussion Paper No. 08-010/3.

Becker, Gary S. 1973. "A Theory of Marriage: Part I," *Journal of Political Economy* 81, 4, 813-846.

Boyd, John H., Edward C. Prescott and Bruce D. Smith 1988. "Organizations in Economic Analysis," *Canadian Journal of Economics* 21, 3, 477-491.

Burdett, Kenneth and Dale T. Mortensen 1998. "Wage Differentials, Employer Size, and Unemployment," *International Economic Review* 39, 2, 257-273.

Costrell, Robert M. and Glenn C. Loury 2004. "Distribution of Ability and Earnings in a Hierarchical Job Assignment Model," *Journal of Political Economy* 112, 6, 1322-1363.

Dustmann, Christiann and Costas Meghir 2005. "Wages, Experience and Seniority," *Review of Economic Studies* 72, 1, 77-108.

Ferrall 1997. "Empirical Analysis of Occupational Hierarchies," *Journal of Human Resources* 32, 1 (Winter), 1-34.

Galindo-Rueda, Fernando and Jonathan Haskel 2005. "Skills, Workforce Characteristics and Firm-Level Productivity: Evidence from the Matched ABI/Employer Skills Survey," IZA Discussion Paper # 1542, March.

Haltiwanger, John, Julia Lane, James Spletzer, Jules Theeves, and Kenneth Troske (eds) 1999. *The Creation of Employer-Employee Matched Data*, (Amsterdam: North-Holland) 1999.

Hutchens, Robert M. 1989. "Seniority, Wages and Productivity: A Turbulent Decade," *Journal of Economic Perspectives* 3, 4 (Autumn), 49-64.

Hægeland, Torbjørn and Tor Jakob Klette 1999. "Do Higher Wages Reflect Higher Productivity? Education, Gender and Experience Premiums in a Matched Plant-Worker Data Set," in Haltiwanger et al. (op cit).

Kremer, Michael and Eric Maskin 1996. "Wage Inequality and Segregation by Skill," NBER Working Paper # 5718, August.

Krishna, Vijay and Roberto Serrano 1996. "Multilateral Bargaining," *Review of Economic Studies* 63, 1, 61-80.

Legros, Patrick and Andrew Newman 2002. "Monotone Matching in Perfect and Imperfect Worlds," *Review of Economic Studies* 69, 4 (October), 925-942.

Lensberg, Terje 1988. "Stability and the Nash Solution," *Journal of Economic Theory 45*, 330-341.

Lindbeck, Assar and Dennis J. Snower 1998. *The Insider Outsider Theory of Employment and Unemployment*, Cambridge: MIT Press.

Milgrom, Paul and Chris Shannon 1994. "Monotone Comparative Statics," *Econometrica* 62, 1, 157-180.

Mincer, Jacob and Boyan Jovanovic 1981. "Labor Mobility and Wages," in *Studies in Labor Markets*, Sherwin Rosen (ed.), Chicago: University of Chicago Press.

Nash, John 1950. "The Bargaining Problem, *Econometrica* 18, 155-62.

Oi, Walter and Todd L. Idson 1999. "Firm Size and Wages," in in O. Ashenfelter and D. Card, eds. *Handbook of Labor Economics*, Vol IIIB, Amsterdam: Elsevier, 2165-2214.

Postel-Vinay, F. and J.-M. Robin 2002. “Wage Dispersion with Worker and Employer Heterogeneity, *Econometrica* 70, 6, 2295-350.

Prendergast, Canice 1999. “The Provision of Incentives in Firms," *Journal of Economic Literature* 37, 1, 7-63.

Rosen, Sherwin 1982. “Authority, Control, and the Distribution of Earnings," *Bell Journal of Economics 13*, 2 (Autumn), 311-323.

Salvanes, Kjell Gunnar, Simon Burgess and Julia Lane 1999. “Sources of Income Dispersion and Firm Differences: Using Matched Employer-Employee data from Norway," in Haltiwanger et al. (op cit).

Sattinger, Michael 1993. “Assignment Models of the Distribution of Earnings," *Journal of Economic Literature* 31, 2, 831-880.

Stole, Lars A. and Jeffrey Zwiebel 1996. “Intra-Firm Bargaining under Non-Binding Contracts," *Review of Economic Studies* 63, 3, 375-410.

Teulings, Coen N. 1995. “The Wage Distribution in a Model of the Assignment of Skills to Jobs," *Journal of Political Economy* 103, 2, pp. 280-315

Table 1. Norwegian Workplaces, 1997 (20% sample, F=20,542)

Industry	Workplace Size Category (N_f) 2-5	6-20	21-50	51-100	>100	Total	Pct in 2-5
No code	774	226	32	5	2	1,039	74%
Elect.	522	177	52	18	10	766	68%
Manufacturing	750	699	263	91	76	1,811	41%
Construction	997	631	110	25	13	1,805	55%
retail trade	3610	2256	301	48	17	6,295	57%
Comm.	935	444	108	32	26	1,559	60%
FIRE	1256	601	144	46	28	2,085	60%
Services	2470	1760	686	218	83	5,284	47%
Total	11,314	6,794	1,696	483	255	20,542	55%
Dist. (%)	55%	88%	96%	99%	100%		

Table 2. Workers in Norway, 1997 (20% sample; N =247,521)

Variable	Name	diag(M)	Mean	St. Dev.
Earnings[a]	W^o	-	19.074	12.403
Female	FEM	1	0.452	
Experience[b]	EX	1	13.248	9.608
Experience Squared/100	EX^2	1	2.678	2.920
Experience X Female	EXxFEM	0	4.683	7.427
Seniority[c]	SN	0	5.432	5.976
Seniority^2 / 100	SN^2	0	0.652	1.258
Seniorty X Fulltime	SNxFT	0	4.434	5.975
(Seniority^2 / 100) X Fulltime	SN^2xFT	0	0.554	1.215
Seniorty X Female	SNxFEM	0	2.262	4.402
Seniorty X (E5+E6)	SNxED	0	0.978	3.155
Fulltime Worker	FT	1	0.760	
Education[d]	EDUC			
<= 9yrs	E1	1	0.130	
10 or 11 yrs	E2	1	0.332	
12 or 13 yrs	E3	1	0.313	
14-16 yrs	E4	1	0.151	
>= 17 yrs	E5	1	0.044	

[a] Monthly kroner / 1000 (approx. US$150).

[b] Full-time equivalent years since 1968, from public pension records.

[c] Years since joining workplace, censored at 20 years.

[d] Default category is no education level recorded (3% of the sample).

Table 3. Selected Worker Characteristics by Industry

	Worker Characteristics						
					Education (yrs)		
Industry	Female	Exper.	Sen.	Fulltm.	10-11	12-13	14-16
No code	0.41	11.59	1.09	0.73	0.29	0.35	0.15
Agr. Min. Elect.	0.21	15.94	7.73	0.83	0.33	0.38	0.07
Manufacturing	0.27	14.57	6.42	0.90	0.37	0.33	0.06
Construction	0.09	14.05	5.04	0.94	0.35	0.42	0.04
Trade	0.50	10.73	4.90	0.66	0.40	0.33	0.07
Comm.	0.28	15.64	4.77	0.83	0.40	0.34	0.06
FIRE	0.43	13.86	4.95	0.86	0.25	0.38	0.19
Services	0.67	13.02	5.67	0.66	0.29	0.23	0.29

Mean values of elements of x within industry.

Table 4. Coworker Characteristics

	Standard Deviations			Correlation
Mean of Co-Worker Values	overall	between workplaces	within workplaces	w/ worker value
Earnings	7.71	8.90	2.29	0.45
Education <= 9yrs	0.15	0.20	0.06	0.19
Education 14-16 yrs	0.20	0.19	0.05	0.43
Experience	5.32	6.18	1.48	0.40
Seniority	3.74	3.78	0.69	0.54
Fulltime Worker	0.27	0.32	0.07	0.49
Female	0.32	0.36	0.07	0.54

Variation across all workers of the mean of co-worker values.

Table 5. Linear Technology Parameter Estimates (Mincer-like Wage Regressions)

Par	Variable	Estimate	Std.Err	Estimate	Std.Err
d	External Barg. Power	1.000	-	1.000	-
Vu	Value of Unemployment	0.000	-	0.000	-
	Measurement Error SD	10.186 *	0.001	10.036 *	0.001
A	No code	10.450 *	0.102	10.676 *	0.115
A	Agric., Mining, Elect.	10.450	-	12.236 *	0.120
A	Manufact.	10.450	-	10.690 *	0.101
A	Construction	10.450	-	10.352 *	0.105
A	Wholesale & retail trade	10.450	-	10.810 *	0.102
A	Trans., Storage & Comm.	10.450	-	11.259 *	0.109
A	FIRE	10.450	-	12.401 *	0.115
A	Services	10.450	-	9.492 *	0.090
	Female	-0.232 *	0.0080	-0.196 *	0.0076
	Experience	0.052 *	0.0005	0.052 *	0.0005
	Experience Squared/100	-0.117 *	0.0014	-0.116 *	0.0014
	Experience X Female	-0.0006	0.0005	-0.0009	0.0076
	Seniority	0.036 *	0.0025	0.035 *	0.0023
	Seniority^2 / 100	-0.150 *	0.0138	-0.146 *	0.0128
	Seniorty X Fulltime	-0.026 *	0.0026	-0.025 *	0.0024
	(Seniority^2 / 100) X Fulltime	0.125 *	0.0139	0.121 *	0.0129
	Seniorty X Female	-0.003 *	0.0007	-0.003 *	0.0007
	Seniorty X (E5+E6)	-0.009 *	0.0003	-0.0075 *	0.0002
	Fulltime Worker	0.472 *	0.0085	0.436 *	0.0080
	Education <= 9yrs	-0.333 *	0.0054	-0.336 *	0.0053
	10 or 11 yrs	-0.252 *	0.0039	-0.256 *	0.0039
	12 or 13 yrs	-0.111 *	0.0036	-0.117 *	0.0036
	14-16 yrs	0.069 *	0.0042	0.091 *	0.0041
	>= 17 yrs	0.277 *	0.0043	0.287 *	0.0042
	-ln likelihood	698,252		694,579	
	2*increment (			7345.7	

* significant at the 1% level.

Table 6. Joint Technology Parameter Estimates (Unproductive Seniority)

Par	Variable		Estimate		Std.Err
d	External Barg. Power		0.695	*	0.0054
V_U	Value of Unemployment		-1.519	*	0.3666
	Measurement Error SD		9.951	*	0.0010
	No code		0.067		0.0544
	Agric., Mining, Elect.		0.326	*	0.0609
	Manufact.		0.161	*	0.0534
	Construction		0.068		0.0551
	Wholesale & retail trade		0.162	*	0.0537
	Trans., Storage & Comm.		0.157	*	0.0552
	FIRE		0.00031		0.0458
	Services		-0.671	*	0.0937
	Female		-0.247	*	0.0140
	Experience		0.060	*	0.0027
	Experience^2/100		-0.134	*	0.0062
	Experience X Female		0.00147	*	0.0006
	Fulltime Worker		0.479	*	0.0219
	Education	<= 9yrs	-0.415	*	0.0198
		10 or 11 yrs	-0.318	*	0.0150
		12 or 13 yrs	-0.147	*	0.0078
		14-16 yrs	0.067	*	0.0055
		>= 17 yrs	0.273	*	0.0131
	Seniority		0.087	*	0.0077
	Seniority^2 / 100		-0.280	*	0.0422
	Seniorty X Fulltime		-0.002		0.0066
	(Seniority^2 / 100) X Fulltime		0.032		0.0402
	Seniorty X Female		-0.032	*	0.0021
	Seniorty X (E5+E6)		0.0020		0.0008
	-ln likelihood		692,446		

Estimates of industry-specific coefficients A not reported. Seniority coefficients in set to 0. * significant at the 1% level.

Table 7. Comparison of Seniority Estimates

Par	Variable	Productive Seniority: Linear Tech.[a]	Productive Seniority: Equal Power	Productive Seniority: Linear Power	Flexible Power: Unprod. Seniority[b]	Flexible Power: Productive Seniority
d	External Barg. Power	1.000	0.734 *	0.726 *	0.695 *	0.710 *
V_U	Value of Unemployment	0.000	0.000	-0.654	-1.519 *	-0.986 *
	Measurement Error SD	10.036 *	9.987 *	9.961 *	9.951 *	9.948 *
	Seniority	0.0348 *	-0.0116	-0.0142 *	0	-0.0131 *
	$Seniority^2$ / 100	-0.1457 *	0.0386	0.0314	0	0.0183
	Seniorty X Fulltime	-0.0250 *	0.0143 *	0.0202 *	0	0.0173 *
	($Seniority^2$ / 100) X Fulltime	0.1211 *	-0.0450	-0.0506	0	-0.0303
	Seniorty X Female	-0.0029 *	-0.0010	-0.0040 *	0	-0.0027 *
	Seniorty X (E5+E6)	-0.0075 *	-0.0032 *	-0.0028 *	0	-0.0037 *
	Seniority	-	0	0.022 *	0.087 *	0.0398 *
	$Seniority^2$ / 100	-	0	0 *	-0.280 *	-0.0153
	Seniorty X Fulltime	-	0	0.007 *	-0.002	0.0340 *
	($Seniority^2$ / 100) X Fulltime	-	0	0 *	0.032	-0.1777 *
	Seniorty X Female	-	0	-0.026 *	-0.032 *	-0.0330 *
	Seniorty X (E5+E6)	-	0	0.005 *	0.002	0.00108
Estim. Sample	- ln likelihood	694,579	693,342	692,703	692,446	692,429
	2*increment ([d]	2473.5	1277.9	515.7	32.7	
Holdout Sample[c]	- ln likelihood	694,646	693,122	692,555	692,189	692,182
	2*increment ([d]	3048.0	1132.9	733.5	12.5	

a. Repeats portions of Table 5.

b. The preferred specification. Repeats portions of Table 6.

c. Holdout sample is another 20% subsample of the population including workplaces not in the estimation sample.

d. Incremental likelihood ratio test compares the specification in the column to the nesting specification in the next column.

* significant at the 1% level

Table 8. Wage Variation and Assumed Technology

		Standard Deviations		
	source	**Data**	**Joint**	**Linear**
Payroll	**overall**	12.40	7.29	7.24
	between workplaces	8.90	5.92	5.25
	within	9.98	5.08	5.44
V(z)	**overall**		5.81	6.89
	between workplaces		4.04	4.88
	within		4.63	5.44