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. Sorensen
Department of Economics 
VU University 
Amsterdam
sameos@gmail.com

February 5, 2009

Abstract:
To measure the importance of interactions at the workplace, this paper develops an equilibrium model of joint production and estimates it using administrative data from Norway. The result is a novel explanation for the relationship between earnings and seniority (tenure) and a re-interpretation of other results based on linearly separable technology. Coworkers interact through a task-assignment model of production, and the nested separable model is rejected in favor of the joint production model. Wages are determined through multi-lateral bargaining over the surplus that accrues to the workforce as a whole. A worker’s outside alternative is determined in equilibrium by expected wages when bargaining to join other workplaces. Seniority enters wages through skills and through relative bargaining power. The separate identification of these channels comes from the model’s implication that skills affect total wages and the workplace but power affects only the distribution. We find bargaining power is significantly related to seniority but we do not reject the hypothesis that seniority is unrelated to productivity.

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.
1. Introduction

This paper develops an equilibrium model of joint production designed to explain wages in matched data sets. It estimates the model using cross-sectional data for Norway in 1997. 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 for two channels through which seniority affects wages. First is the usual productivity channel. The second is through bargaining power over the surplus created by the workplace. The combination of rich data and an equilibrium model allows us to separate these channels. An individual’s seniority acts through the productivity channel by raising total workplace revenue. It works through the bargaining channel through relative seniority. Using full payroll data on a large sub-sample of all workplaces the absolute and relative effects of seniority can be disentangled. Maximum likelihood estimation of the joint production economy determines the strength of each channel in determining wages.

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 little is known about an individual’s employer the effects of coworkers are unobserved, which leads to models where output is linearly separable across its workers. The assumption of linearly separable workers has been maintained by recent models of the labour market, such as Burdett and Mortensen (1988) and Postel 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 interested in questions involving single or homogeneous firms or 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 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 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 analyze a static problem in which assignments to workers 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, 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

¹ Teulings (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.
such as

\[ \ln W = b_0 \text{Exper} + b_1 \text{Exper}^2 + b_2 \text{Sen} + b_3 \text{Sen}^2 + \ldots \] (1)

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 - 5). 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 do not attempt to correct corrections to OLS. Instead, we use the equilibrium Nash bargaining model of a workplace’s whole payroll to address two neglected issues related to seniority. First, much empirical work glosses over the indeterminancy of the competitive wage-seniority relationship. With firm-specific capital comes a bilateral surplus between the firm and worker match 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.

\(^2\) This is separate from explanations of seniority profiles due to information and incentives, as discussed in Hutchens (1989) or Prendergast (1999).
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 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 is only slightly better than making seniority unproductive. This becomes our preferred specification and with we explore some further implications of relaxing linear separability across workers. In particular joint production is an important additional element to the model’s interpretation of firm-size and male-female wage differentials.

2. The Joint Production Economy

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.
2.1 Workers and Workforces

In the model a workplace matches the usual definition of a plant or establishment 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 workers contribute 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 \( \gamma \) contains exogenous coefficients.

A1a. The index \( x_\gamma \) is composed of internal and external components:

\[
x_\gamma = x_{\text{total}} = x_{\text{external}} + x_{(I-M)\gamma} \quad \text{(2)}
\]

\(^4\) 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 of intermediate outputs would be needed to smooth output variations.
where $M$ is an exogenous and idempotent $P \times P$ matrix that masks out the internal components of total talent leaving only the external component.

**A1b.** A worker’s talent follows the exponential distribution:

$$G(a; x\gamma) = 1 - e^{-ae^{-x\gamma}}.$$

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:

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

Assuming that a worker’s talent follows the exponential distribution (A1b) is convenient for the empirical analysis. The log of mean talent $\ln E[a|x] = \ln(1/e^{-x\gamma}) = x\gamma$ is the scalar that 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 $Q(X; C)$, 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 $N \times \{\text{per-worker revenue}\} = NQ(X; C)$.\(^6\) The technology requires that workers be assigned to one of two tasks which generate intermediate outputs. Tasks assignments can be conditioned on the realized value of $a$, so a worker spends a fraction of any period 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 produc-

---

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

\(^7\) 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 $g(a; x_n)$. 

\(^8\) It is possible to extend the model to more than two tasks (or levels). Information on
tion 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 use. They are indifferent to the talents and identifies 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. The technology $Q(X;C)$ incorporates revenue-maximizing task and input assignments and the two task-specific technologies.

**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

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

**A2b.** A worker in task 2 produces

$$Q_2(a,q_1;\beta) = \begin{cases} aq_1^\beta & \text{if } \beta \geq 0 \\ a^{1+\beta}q_1^{-\beta} & \text{if } \beta < 0 \end{cases} = a^{1+(1-I_\beta)\beta}q_1^{(I_\beta-1)\beta}.$$

For all values of $\beta$ 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 $\beta$ is above, below or exactly zero. For negative $\beta$ the task 1 technology is concave in $a$ and the task 2 output exhibits constant returns to scale. For positive $\beta$ the task 1 technology is linear in $a$ and task 2 output exhibits increasing returns to scale. For $\beta = 0$ assignment within workplace is not always available in administrative data, and without it the value added by allowing for more tasks is unclear.
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 tasks assignments:

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

\[
\{\phi^*(a), q_1^*(a)\} = \arg \max_{\phi(a), q_1(a)} N A \int_0^\infty Q_2 \left(a, q_1(a)\right) \phi(a) g(a; X \gamma) da
\]

subject to \(N \int_0^\infty Q_1(a)(1 - \phi(a)) g(a; X \gamma) da = N \int_0^\infty q_1(a) \phi(a) g(a; X \gamma) da.\) (5)

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

The constraint (5) says the workplace must allocate talent internally so that the supply of task 1 output produced equals demand coming from task 2. Condition (5) is simply a condition for the production process to be feasible. Under the recursive technology 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 \(\lambda\) 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:

\[
\phi^*(a) = I_{a > \bar{a}(\lambda; \beta, A, X)}.
\] (6)

Another way to summarize the optimal arrangement is a task assignment \(t^*(a) = 2\) if \(a > \bar{a}\) and 1 otherwise.
I1b. Given the technology the cut-off $\bar{a}$ is monotonic in $\lambda$ and depends on $X$ only through $\lambda$.

I1c. For $\beta = 0$ the technology is linearly additive across workers. For $\beta \neq 0$ worker talents interact in determining workplace revenue.

I1d. $Q(X, \gamma; C_f)$ is a continuous function of $\beta$, $A$, and $\gamma$ on their permitted ranges.

The form of $\bar{a}$ under the different cases of $\beta$ 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 $\lambda$ indicates that the workforce is more “top heavy,” because of two effects that move together. First, with a greater $\lambda$ talent marginally assigned to task 2 moves down into task 1 ($\bar{a}$ increases). Second, each worker working in task 2 faces a higher shadow price for input and so cuts down their use of task 1 output accordingly ($q_1(a)$ decreases).

Figure 2a and Figure 2b illustrate optimal task assignment when $\beta > 0$ and $\beta < 0$, respectively. Given the distribution of talent $g(a; X, \gamma)$ the value of $\bar{a}$ and $\lambda$ are set so that expected amount of $Q_1(a)$ equals the expected amount of $q_1(a)$. Per-worker output is the expected value of $Q_2(a)$. Recalling that the talent distribution is the mixture over individual talents it can seen that the three workers would spend very different amounts of time in task 2.

In the case $\beta = 0$ primary production becomes unnecessary. All talent is devoted to the managerial task which in the limit is simply talent added up across workers. The first three properties of the optimal task assignment require only the assumptions on technology and continuous unbounded support of talents. The assumption of exponential talent (A1) is maintained throughout the analysis, and the resulting overall specification of the technology for all cases of $\beta$ is given in the appendix. In each case all the firm specific terms have closed forms in $\lambda$. One proper integral enters a single non-linear equation in $\lambda$, which can be computed robustly using bi-section and/or Newton iteration.
Figure 2a. Optimally Assign Low Talent to Task 1 ($\beta > 0$)

Smoothness and monotonicity imply that the solution to $\lambda$ is smooth and once found workplace revenue and most other aspects of the workplace have smooth closed form solutions in the underlying parameters.

**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 output, even if $\bar{a}_f$ and $\lambda_f$ were held constant. The new distribution changes optimal task assignment. To describe how marginal contribution varies with the new worker’s characteristics and the existing workforce characteristics some additional notation is useful.

**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:

$$
VMP_f(x) \equiv (N_f + 1) Q (X_f^{x \gamma}; C_f) - N_f Q (X_f; C_f).
$$

(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 it is worth considering the implications of what the optimal task assignment 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 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 recursive 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. Similar workers are substitutes while those with different talents are complements. Thus, whatever the equilibrium assignment of coworkers is with such a technology it does not exhibit perfect segregation by skill.

This is illustrated in Figure 4 by comparing two workforces, \( X \) and \( X' \). The former workplace can be said to more talented than the former because its talent distribution stochastically dominates the latter: \( G(a; X\gamma) < G(a; X'\gamma) \) for all \( a \). Because workplace \( X \) has more talent it has greater internal demand for \( Q_1 \). This results in a greater value of \( \bar{a} \) (greater value of \( \lambda \)) than the other workplace. 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 who can transform task 1 output. Now consider adding either a more or less talented worker \( (x'\gamma < x\gamma) \) to either of
these firms. Workplace $X'$ can potentially out-bid $X$ for worker $x$ because hiring the better worker relaxes their constraint on leadership/management talent. Meanwhile, workplace $X$ may prefer to add $x'$ in order to produce task 1 output, freeing up time for their existing workers to engage in task 2 production. That is, the comparative and absolute advantages 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 12: Separability.** When the technology parameter $\beta$ equals 0 then $\lambda = 0$ and VMP is separable across workers and log-linear in observables: $\ln \text{VMP}_f(x) \rightarrow \ln A + x\gamma$.

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 (mincer). 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 can be traced back to 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 ($M = I$) a worker is 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.\footnote{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).} 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).

**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 $V(x,M),\gamma,$ which appears in the vector $V(X,M,\gamma).$

**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

$$\Pi(X(I-M)\psi) = [\pi_n] = \frac{e^{X(I-M)\psi}}{e^{X(I-M)\psi}}.$$  \hspace{1cm} (8)

The $N \times 1$ vector $\psi$ contains all 1’s and $\psi$ is a vector of coefficients that relate seniority to relative power.

When employers have zero-valued threat points and the total bargaining power of the workforce is constant (irrespective of the number and characteristics of the workers) then two simplifications occur:

**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 $A$ and the outside alternative $V()$ to include factors that account for the workforce’s bargaining power.

I3b. The bargaining outcome among workers can be separated from the bargaining outcome between employers and their workforces.

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. It is presumed that in equilibrium outside alternatives will depend only on general talents of the worker. $X(I - M)$ is the matrix of workplace-specific components of talent. The ability to capture surplus is a linear combination of these factors, $X(I - M)\psi$. Given the form of $\Pi$ 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 $\psi$ is a zero vector then bargaining power is equal across workers and independent of $X\gamma$; each element of $\Pi$ becomes $1/N$. 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 $\psi$ is zero except for the coefficient on seniority. As it increases power shifts to senior workers. In the limit the vector $\Pi$ becomes an indicator vector for the worker with the most seniority. All the power accrues to the worker with the most internal (workplace-specific) talent who pays coworkers their outside alternatives and captures the whole surplus for themselves.

**Definition D3: Nash Payroll.** Denote the vector of outside values as $V(XM\gamma)$ and its average as $\bar{V}(XM\gamma) = 1/N \sum_{n=1}^{N} V(x_nM\gamma) = \psi V(XM\gamma) / N$. Denote the total surplus generated by the workplace as

$$S(X\gamma, V(XM\gamma)) = N \left[ Q(X\gamma; C) - \bar{V}(XM\gamma) \right].$$

(9)
Figure 4. Sources of Wage Variation within and Between Workplaces

<table>
<thead>
<tr>
<th>Source</th>
<th>Index</th>
<th>Path to $W^*$</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal</td>
<td>$X(\mathbf{I} - \mathbf{M})\psi$</td>
<td>Power: $\Pi(X(\mathbf{I} - \mathbf{M})\psi)$</td>
<td>Relative</td>
</tr>
<tr>
<td>External</td>
<td>$X M\gamma$</td>
<td>Alternatives: $V(X M\gamma)$</td>
<td>Individual</td>
</tr>
<tr>
<td>Total</td>
<td>$X\gamma$</td>
<td>Productivity: $Q(X\gamma; C)$</td>
<td>Joint</td>
</tr>
<tr>
<td>Sectoral</td>
<td>$C$</td>
<td>Technology: $\beta, A$</td>
<td>Exogenous</td>
</tr>
</tbody>
</table>

A workplace is feasible if it produces a surplus: $S(X\gamma, V(X M\gamma)) \geq 0$. For a feasible workplace the multilateral Nash payroll is a vector $W^* = \left[W^*_n\right]$ that solves

$$W^*\left(X\gamma; V(X M\gamma) ; C\right) = \arg\max_{\{W_n\}} \prod_{n=1}^{N} \left[W_n - \bar{V}(x_n M\gamma)\right]^\pi_n. \quad (10)$$

**Implication 14: The Nash Payroll in a Joint Production Workplace.**

$$W^*\left(X\gamma; V(X M\gamma) ; C\right) = V(X M\gamma) + S\left(X\gamma, X M\gamma\right) \Pi\left(X (\mathbf{I} - \mathbf{M})\gamma\right). \quad (11)$$

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 the assignment of workers to workplaces is exogenous, 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 and, given $V(z)$, determines its feasibility as a workplace.

**S1.** Each worker is hypothetically placed in a randomly selected feasible workforce, assuming no other workers change position. Workforces bargain with arriving outside workers one at a time ignoring the hypothetical movement of its own workers and the possible arrival of other workers. The existing workforce (or the employer) acts as a coalition bargaining with the hypothetical worker. The hypothetical worker has power $\eta \in [0, 1]$. 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 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 is a simple story to link wages with joint production in other workplaces. It retains the assumption of bargaining among workers and assumes that arriving workers are evaluated according to their contribution to the workplace’s surplus. However it does not attempt to reconcile that current workers arrived as outsiders in the past and somehow became insiders. 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.

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 15: 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

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

with $VMP_{f'}(z)$ defined in (7) and $d$ defined in S1 of (A4).

15a. $V(z)$ is an equilibrium if all workplaces are feasible under it.

15b. For $V_U$ and $d$ sufficiently close to 0 any workplace will be feasible under $V(z)$.

This equilibrium is typically not unique because workplace-specific talent drives a wedge between the VMP in the current and outside workplaces. Other workers can capture some of the gap without violating the equilibrium conditions. 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. If $d = V_U = 0$ then all workers would 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 one generalizes the simple “wage = VMP” in the textbook model. For example, consider a linear $(\beta = 0)$ homogeneous technology, no value to 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 $W^* = V(z) = VMP$ and all workplaces would be just feasible. This special case serves as a benchmark for the joint production technology.

2.6 Solution Method and Empirical Considerations

The goal of the empirical analysis is to compare predicted payrolls to observed payroll
vectors, denoted $W_f$ for $f = 1, \ldots, F$. In the theory wages are not stochastic. To match data we introduce measurement error. The observed payroll is the equilibrium payroll $W^*$ plus an iid normal error vector:

$$W^o = W^* + \epsilon, \quad \epsilon \sim N(0, \sigma^2 I). \tag{13}$$

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 wages within workplaces in levels, not 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.

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.\footnote{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, $\exp\{-\bar{a}\exp\{-x\gamma\}\}$.} We allow the production coefficient $A$ and the exponent $\beta$ to differ across industries. The vector of free parameters of the model is written:

$$\theta \equiv (A(C) \quad \beta(C) \quad \gamma \quad \psi \quad h \quad V_U \quad \sigma). \tag{14}$$

For a parameter vector $\theta$ 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 (\text{veq}): 

\[
v(z) = d\text{VMP}_f(z) > V_u (\text{VMP}_f(z) - V_u).
\]

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

\[
V(z) = \hat{E}[v(z)|z] = b_0 + b_1 z + b_2 z^2 + b_3 z^3. \tag{15}
\]

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 \( W^* \) in \( \theta \) since \( VMP_f(z) \) and \( v(z) \) are continuous in \( z \) which in turn is continuous in \( \theta \).

With total output and outside alternatives computed, the the log-likelihood for firm \( f \) is simply the normal density of the implied error terms:

\[
l_f(\hat{\theta}; W_f^0, X_f) = -0.5/\sigma - \hat{\epsilon}/\sigma^2
\]

and \( \hat{\epsilon} = W^o - W^*(X; \hat{\theta}) \). In addition, the equilibrium requires that each workplace generate a positive surplus, \( S(X\gamma, XM\gamma) > 0 \). This is a stringent requirement especially for small workplaces where the characteristics of each worker has a big impact on the implied talent distribution. On the other hand, without the discipline of a positive surplus the model would be allowed to predict workers get paid less than their outside alternative. To balance these concerns we penalize the likelihood for negative surpluses but only in workplaces with more than five workers. The overall objective is therefore:

\[
\hat{\theta}^{ML} = \max_\theta \sum_{f=1}^F l_f(\hat{\theta}; W_f^0, X_f) - DI_{S_f < 0} I_{N_f > 5}. \tag{16}
\]

The value of \( D \) was increased as estimation proceeded. Ultimately no penalty was incurred.

2.7 Identification

As discussed in the introduction, we use our estimation of a joint economy equilibrium to reconsider how wages, seniority and other characteristics are related. Given the assumption of exogenous \( X \) 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 productivity gain. 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. What identifies the share/power effect of seniority from the productivity effect of seniority other than the specific functional forms assumed above? The model generates the following key prediction: As seniority affects productivity the effect is seen in the total payroll (the sum of the payroll vector \( W^* \)). As seniority affects the distribution of the surplus the effect is seen in the distribution of the payroll (the correlation between \( W^* \) and the elements of \( X \) related to seniority, \( (I - M)X \)). Since seniority is left at the workplace door it has no direct effect on the outside alternative \( V(z) \), but that function must be estimated by using the relationship between the payroll and the general skills \( MX \) (which determine \( z \)). A large representative sample of workplaces and their fully described workforces provides variation in total payroll (productivity) and the distribution of payroll (sharing) that separately identifies these effects within an equilibrium context.

Of course this identification is parametric. It is based on an explicit parameterized model of joint production that follows the literature on task assignment within organizations. The model can also produce as a special (nested) case a linear, individualistic production function which supports the standard wage equation (1). The complex technology in (A2) ensures the special case lies in the interior of the support of \( \beta \). 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 unobserved total payroll. It also leaves relative seniority unobserved even with panel data. Thus the joint production parameter \( \beta \) is unidentified from that kind of data, and therefore the assumption \( \beta = 0 \) has been

22
maintained and untested. 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 this 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.\textsuperscript{11} 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 matched worked. In some cases the industry code is missing and in others workers in the same workplace report 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 com-

\textsuperscript{11} 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, along with detailed summaries of output, are available from the first author’s website.
bined with larger ones to define 8 distinct categories listed in Table 1. In the specification reported below only the coefficient \( \lambda \) is allowed to vary by industry.

3.1.3 Earnings

Earnings are derived from administrative data used to compute public pension credits. In the data earnings are actual annual 1997 earnings in all jobs in Norwegian kroner (NOK). With the worker’s start date available it is possible to adjust total earnings for attachments that began in 1997. However, earnings would include earnings from earlier jobs. Instead, months of unemployment in 1997 (merged in from another administrative database) are used to correct for jobless months and monthly earnings in November job and any earlier or later 1997 jobs are assumed to be equal. Wages are expressed in monthly thousands of kroner (NOK1000):

\[
W^* = \frac{1}{1000} \times \frac{\text{total 1997 earnings in NOK}}{12 - \# \text{ months unemployed in 1997}}.
\]

In 1997, NOK1000 equalled approximately US$150.

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. Experience is computed from pension points accumulated since 1968 when the Norwegian public pension system began. By inverting the schedule for how pension points are earned total work experience over the thirty years between 1968 and 1997 can be closely approximated. This measure of actual experience is converted to full-year equivalent. The main limitation to this measure of experience is that workers on leave continue to earn pension points, so such spells are not excluded. Seniority at the current workplace cannot be deduced from pension points, but an accurate job-start date for the current job is available. Seniority is censored at approximately 20 years because start dates only go back until April, 1978. The years-since-joining variable is interacted with current
full-time status, which assumes that current full-time status is a good proxy for past status within the workplace. It will be inaccurate for people who have recently changed their regular work hours. Furthermore, seniority should be interpreted as ‘potential seniority’ because it includes leaves of absence. As with potential experience the upward bias of the measure is worse for women due to maternity leaves.

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 zero.

### 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 ($W^a < .1$ and $W^a > 1000$), 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.

Table 1 summarizes the data by workplace. The number of “no-code” workplaces is small except for very small workplaces for which uni-modal 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 senior. FIRE (Financial and Retail Services) and other 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. So not only do earnings vary more across workplaces than within, so do all the major characteristics of 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 may overstate the direct VMP of individual characteristics if production takes place together with coworkers and surpluses are shared.
3.2 Estimates

Table 5 reports estimates of two versions of the linear technology which sets the key parameter \( \beta \) 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 \( \gamma \) 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. Allowing industries to have different productivities improves the fit significantly. (Adding seven parameters results in an improvement in the likelihood.)

Table 6 reports estimates of our preferred specification of the joint technology model in which industry-specific \( \beta \)s and the bargaining power shifters \( \psi \) 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 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 regression coefficients.

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 significantly improves the fit over the unrestricted linearly separable specification. The change in log-likelihood is 1236.7. The next column in Table 7 allows the linear seniority terms in the bargaining vector $\psi$ to be non-zero. Freeing these four parameters improves the likelihood by 638.9 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. 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 $\chi^2$ 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 $\gamma$ 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 $\gamma$ are set to zero but $\psi$ is estimated.

To further investigate how the technology assumption affects the role of seniority we computed the response to increasing each worker’s seniority by five years. In the linear technology this changes the wage due to their share of the changed firm-specific productivity. Because the seniority profile is not monotonic some workers gain and some lose by this change. With the preferred joint technology seniority has no affect on productivity or outside alternatives. Its effect comes solely through bargaining power. When all workers gain five years those with relatively less seniority gain power while others lose. In addition, workers in very large firms see very little response because relatively seniority has little effect when everyone gains. Figure 5 shows the distribution of the wage changes by actual seniority. The connected line is the mean change in wages, which in all cases is nearly zero, although more senior workers lose on average in both technologies. More importantly is the one standard deviation around the mean changes illustrated by the shaded regions. With joint technology the aggregate response to seniority is more varied than with the linear assumption, although in percentage terms the variation is still small relative to average wages.

3.3 The Wage Distribution and Technology

3.3.1 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
Figure 5. The Marginal Effect of Seniority is Small

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.

3.3.2 Gender Differentials

Now consider how the joint technology assumption changes the explanation for dif-
ferences 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. Women also benefit from seniority less than men (a significant but not large difference). Recalling that we use a measure of actual experience based on pension points, the difference in experience is insignificant. The joint technology estimates are not that different in magnitude, except the seniority effect in $\psi$ is larger, and the experience differential is now significant.

However, Figure 6 shows that the effect of these differences feeds into wages through different channels. The figure show 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. 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 mainly 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. 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.

3.3.3 Workplace Size Differentials

Figure 7 displays another aspect of the wage distribution. For the preferred specification, predicted wages \((W^*)\) and outside alternatives \(V(XM_γ)\) are averaged within workforces and then averaged over workforces of the same size \(N_f\). The resulting wage-size profile is displayed after smoothing. Observed wages \(W^o\) are also averaged and smoothed. Recall that the model exhibits constant returns to scale in \(N_f\) holding constant the distribution of talents. So any pattern across firm sizes arises out of the estimates, not an assumption that large firms are more productive per se. The predicted profile tracks the observed rising profile, although the model wages are too high for workplaces with fewer than 100 workers and too low for firms up to size 200. But the inflection points match closely. Beyond 200 workers the number of firms in Norway is small and the model matches captures the flattening profile.
This explanation of the size profile is partly based on differences in observable characteristics between large and small workforces (education and experience included). But the displayed profile of outside alternatives shows the gap goes beyond this component. As with the previous figure both predicted wages ($W^*$) and and outside alternatives ($V(XM_\gamma)$) are shown. 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. The gap between wages and alternatives, the average surplus, accounts for most of the profile beyond 100 workers. Seniority is unproductive and changes in individual talent are captured by $V()$. So this rising surplus arises from the model because large workforces have on average more efficient talent distributions than the average small workforce.
3.3.4 Extensions: Workplace Dynamics

These results suggest extending our analysis would lead to further insights. Clearly a workforce at a point in time is a lagged exogenous outcome. Our model can form the basis of a dynamic analysis of matched panel data. 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 will move. Workers of all levels of talent can find better matches but they will not always agree on their preferred destinations. 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 because of lost specific talents but a loss in bargaining power.

As shown above, the joint production technology suggests that some of the wage-size profile is due to workforce efficiency. A dynamic model might amplify this effect. 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. The efficient firm finds it easy to grow in size and build on their advantage.

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 of the technology makes it feasible to impose all the restrictions of the model on over 20,000 individual workplaces at each point in the estimation procedure. And we carefully parameterize the technology so that the linear case is a special case of the linearly separable model, which is rejected with a conventional likelihood ratio test.

Thus, accounting for joint production provides a viable alternative explanation of the data. The restrictive model adds only eight free parameters to the linear model, so the difference in fit is significant but not strikingly different. However, the joint production provides a quite different explanation for the distribution of wages. The model suggests that wages are related to seniority because it affects surplus sharing not productivity. 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. So our estimates explain mixed results concerning seniority wage profiles through a weak (non-existent) productivity effect and a heterogeneous relative seniority effect. Joint production explains the firm-size wage differential beyond differences in observable worker characteristics. Larger workforces produce more surplus through more efficient skills distributions. In addition, the productivity gap underlying the male-female wage differential is smaller than in the linear model. The counterfactual of eliminating sex differences shows a complex effect on wages.
5. Appendix

5.1 Proofs of Implications

Proof of Implication 11

P: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:

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

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

\[
A|\beta|\phi(a)a^{1-|\beta|}[q_1(a)]^{|\beta|-1} = \lambda \phi(a) \text{ for } \beta < 0,
\]

\[
A|\beta|\phi(a)a[q_1(a)]^{|\beta|-1} = \lambda \phi(a) \text{ for } \beta > 0.
\]

This equation can be satisfied with \( \phi(a) = 0 \), or with \( \phi(a) > 0 \) and \( q_1^*(a) \), given as

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

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

\[
[Ka - \lambda a|\beta|]g(a; X_f \gamma) = \mu_1 - \mu_0, \quad \beta < 0,
\]

\[
[Ka^{1/(1-\beta)} - \lambda a]g(a; X_f \gamma) = \mu_1 - \mu_0, \quad \beta > 0,
\]

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 (??) \( g(a; X_f \gamma) > 0 \). So the left hand side is zero only when the difference is zero. In both cases the difference is between a line and a positive power of \( a \). Thus it is only zero for at most one point \( \tilde{a} > 0 \). 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 \( \phi(a) = 1 \). Solving for \( a = \bar{a} \) we get

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

This proves that the optimal assignment of talent is a cut-off rule with \( t^*(a) = 1 \) below \( \bar{a} \) and 2 above.

**P.I1b.** From the expression for \( \bar{a} \) we see that \( X_f \) only enters through \( \lambda \). Monotonicity can be seen by rewriting the internal demand-supply constraint as

\[
\int_0^{\bar{a}} Q_1(a; \beta)g(a; X_f\gamma)da = \int_0^\infty q_1^*(a; \lambda)g(a; X_f\gamma)da.
\]

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

**P.I1c.** When \( \beta = 0 \), then \( Q_2(a, q_1; 0) = a \), and hence total production is invariant to \( q_1 \). No allocation of talent to intermediates is therefore optimal.

**P.I1d.** For \( \beta \neq 0 \) continuity of the model’s predictions is straightforward. The technology (??) is is smooth in \( \beta \) and all other parameters. Then \( \bar{a} \) is continuous because it 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 \( \lambda \) similarly continuous. Continuity extends through all integrals because the bounds and the integrands are continuous in the parameters, \( \lambda \), and \( \bar{a} \).

At \( \beta = 0 \), continuity is slightly complicated because the technology in A2 is continuous (but not differentiable) in \( \beta \). Approaching 0 from either direction the technology is continuous and bounded. From below we see that \( \lim_{\beta \to 0} \lambda \to 0 \) and \( \lim_{\beta \to 0} \bar{a} = 0 \). The contribution of task 1 to integrals approaches \( \int_0^0 0^0 f(0) \).
Since the density (??) is bounded at 0 the limit is 0. In other words, task 1 output goes to 0 even though output at exactly \( a = 0 \) is unbounded. Thus, all predictions of the model

**Proof of Implication 13.** Let \( R \) equal total revenue including the employer’s share; let \( \eta \) denote the employer’s bargaining power, and let \( \pi^* \)t\( n \) denote the power of worker \( n \). The sum of the workforce parameters is \( 1 - \eta = \sum_{n=1}^{N} \pi^*_n \). 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:

\[
\max \quad P^n \prod_{n=1}^{N} (W_n - V^*_n)^{\pi^*_n} \quad \text{subject to} \quad P + \sum_{n=1}^{N} W_n = R.
\]

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

\[
\max \quad \prod (W_n - V^*_n)^{\frac{\pi^*_n}{1-\eta}} \quad \text{subject to} \quad \sum W_n = (1 - \eta)(R - \sum V^*_n).
\]

We can then define \( V_n = V^*_n/(1 - \eta) \); \( Q = R/(1 - \eta) \); and \( \pi_n = \pi^*_n/(1 - \eta) \) and arrive at an problem equivalent to the one defined in (D3).

**Proof of Implication 14:** See (Lensesberg (1988)).

**Proof of Implication 15.**

**P:15a.** Under \( V(z) \) no worker in a feasible workplace will prefer to leave and take a random match to another workplace.

**P:15b.** For \( \nu_v = 0 \) 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 \( \nu_v \) and \( h \), so for some range of values above 0 all workplaces stay feasible and the corresponding \( V(z) \) remains an equilibrium.
6. References


Jobs,"  *Journal of Political Economy* 103, 2, pp. 280-315
Table 1. Norwegian Workplaces, 1997 (20% sample, F=20,542)

<table>
<thead>
<tr>
<th>Industry</th>
<th>Workplace Size Category (Ni)</th>
<th>Total</th>
<th>Pct in 2-5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2-5</td>
<td>6-20</td>
<td>21-50</td>
</tr>
<tr>
<td>No code</td>
<td>774</td>
<td>226</td>
<td>32</td>
</tr>
<tr>
<td>Agric., Mining, Elect.</td>
<td>522</td>
<td>177</td>
<td>52</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>750</td>
<td>699</td>
<td>263</td>
</tr>
<tr>
<td>Construction</td>
<td>997</td>
<td>631</td>
<td>110</td>
</tr>
<tr>
<td>Wholesale &amp; retail trade</td>
<td>3610</td>
<td>2256</td>
<td>301</td>
</tr>
<tr>
<td>Trans., Storage &amp; Comm.</td>
<td>935</td>
<td>444</td>
<td>108</td>
</tr>
<tr>
<td>FIRE</td>
<td>1256</td>
<td>601</td>
<td>144</td>
</tr>
<tr>
<td>Services</td>
<td>2470</td>
<td>1760</td>
<td>686</td>
</tr>
<tr>
<td>Total</td>
<td>11,314</td>
<td>6,794</td>
<td>1,696</td>
</tr>
<tr>
<td>Cumulative Dist. (%)</td>
<td>55%</td>
<td>88%</td>
<td>96%</td>
</tr>
</tbody>
</table>
Table 2. Workers in Norway, 1997 (20% sample; N = 247,521)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Name</th>
<th>diag(M)</th>
<th>Mean</th>
<th>St. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earnings$^a$</td>
<td>W$^o$</td>
<td>-</td>
<td>19.074</td>
<td>12.403</td>
</tr>
<tr>
<td>In Earn</td>
<td>InWo</td>
<td>-</td>
<td>2.760</td>
<td>0.705</td>
</tr>
<tr>
<td>Female</td>
<td>FEM</td>
<td>1</td>
<td>0.452</td>
<td></td>
</tr>
<tr>
<td>Experience$^b$</td>
<td>EX</td>
<td>1</td>
<td>13.248</td>
<td>9.608</td>
</tr>
<tr>
<td>Experience Squared/100</td>
<td>EX$^2$</td>
<td>1</td>
<td>2.678</td>
<td>2.920</td>
</tr>
<tr>
<td>Experience X Female</td>
<td>EXxFEM</td>
<td>0</td>
<td>4.683</td>
<td>7.427</td>
</tr>
<tr>
<td>Seniority$^c$</td>
<td>SN</td>
<td>0</td>
<td>5.432</td>
<td>5.976</td>
</tr>
<tr>
<td>Seniority$^2$ / 100</td>
<td>SN$^2$</td>
<td>0</td>
<td>0.652</td>
<td>1.258</td>
</tr>
<tr>
<td>Seniority $^2$ X Fulltime</td>
<td>SNxFFT</td>
<td>0</td>
<td>4.434</td>
<td>5.975</td>
</tr>
<tr>
<td>(Seniority$^2$ / 100) X Fulltime</td>
<td>SN$^2$xFT</td>
<td>0</td>
<td>0.554</td>
<td>1.215</td>
</tr>
<tr>
<td>Seniority X Female</td>
<td>SNxFEM</td>
<td>0</td>
<td>2.262</td>
<td>4.402</td>
</tr>
<tr>
<td>Seniority X (E5+E6)</td>
<td>SNxED</td>
<td>0</td>
<td>0.978</td>
<td>3.155</td>
</tr>
<tr>
<td>Fulltime Worker</td>
<td>FT</td>
<td>1</td>
<td>0.760</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>EDUC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>none recorded</td>
<td>E0</td>
<td>1</td>
<td>0.030</td>
<td></td>
</tr>
<tr>
<td>&lt;= 9 yrs</td>
<td>E1</td>
<td>1</td>
<td>0.130</td>
<td></td>
</tr>
<tr>
<td>10 or 11 yrs</td>
<td>E2</td>
<td>1</td>
<td>0.332</td>
<td></td>
</tr>
<tr>
<td>12 or 13 yrs</td>
<td>E3</td>
<td>1</td>
<td>0.313</td>
<td></td>
</tr>
<tr>
<td>14-16 yrs</td>
<td>E4</td>
<td>1</td>
<td>0.151</td>
<td></td>
</tr>
<tr>
<td>&gt;= 17 yrs</td>
<td>E5</td>
<td>1</td>
<td>0.044</td>
<td></td>
</tr>
</tbody>
</table>

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

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

c Monthly kroner / 1000 (approx. US$150).
Table 3. Selected Worker Characteristics by Industry

<table>
<thead>
<tr>
<th>Industry</th>
<th>Female</th>
<th>Exper.</th>
<th>Sen.</th>
<th>Fulltm.</th>
<th>Education (yrs)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10-11 12-13 14-16</td>
</tr>
<tr>
<td>No code</td>
<td>0.41</td>
<td>11.59</td>
<td>1.09</td>
<td>0.73</td>
<td>0.29 0.35 0.15</td>
</tr>
<tr>
<td>Agric., Mining, Elect.</td>
<td>0.21</td>
<td>15.94</td>
<td>7.73</td>
<td>0.83</td>
<td>0.33 0.38 0.07</td>
</tr>
<tr>
<td>Manufact.</td>
<td>0.27</td>
<td>14.57</td>
<td>6.42</td>
<td>0.90</td>
<td>0.37 0.33 0.06</td>
</tr>
<tr>
<td>Construction</td>
<td>0.09</td>
<td>14.05</td>
<td>5.04</td>
<td>0.94</td>
<td>0.35 0.42 0.04</td>
</tr>
<tr>
<td>Wholesale &amp; retail trade</td>
<td>0.50</td>
<td>10.73</td>
<td>4.90</td>
<td>0.66</td>
<td>0.40 0.33 0.07</td>
</tr>
<tr>
<td>Trans., Storage &amp; Comm.</td>
<td>0.28</td>
<td>15.64</td>
<td>4.77</td>
<td>0.83</td>
<td>0.40 0.34 0.06</td>
</tr>
<tr>
<td>FIRE</td>
<td>0.43</td>
<td>13.86</td>
<td>4.95</td>
<td>0.86</td>
<td>0.25 0.38 0.19</td>
</tr>
<tr>
<td>Services</td>
<td>0.67</td>
<td>13.02</td>
<td>5.67</td>
<td>0.66</td>
<td>0.29 0.23 0.29</td>
</tr>
</tbody>
</table>

Mean values of elements of x within industry.
Table 4. Coworker Characteristics

<table>
<thead>
<tr>
<th>Mean of Co-Worker Values</th>
<th>Standard Deviations</th>
<th>Correlation w/ worker value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>overall</td>
<td>between workplaces</td>
</tr>
<tr>
<td>Earnings</td>
<td>7.71</td>
<td>8.90</td>
</tr>
<tr>
<td>Education &lt;= 9yrs</td>
<td>0.15</td>
<td>0.20</td>
</tr>
<tr>
<td>Education 14-16 yrs</td>
<td>0.20</td>
<td>0.19</td>
</tr>
<tr>
<td>Experience</td>
<td>5.32</td>
<td>6.18</td>
</tr>
<tr>
<td>Seniority</td>
<td>3.74</td>
<td>3.78</td>
</tr>
<tr>
<td>Fulltime Worker</td>
<td>0.27</td>
<td>0.32</td>
</tr>
<tr>
<td>Female</td>
<td>0.32</td>
<td>0.36</td>
</tr>
</tbody>
</table>

Variation across all workers of the mean of co-worker values.
Table 5. Linear Technology Parameter Estimates (Mincer-like Wage Regressions)

<table>
<thead>
<tr>
<th>Par</th>
<th>Variable</th>
<th>Estimate</th>
<th>Std.Err</th>
<th>Estimate</th>
<th>Std.Err</th>
</tr>
</thead>
<tbody>
<tr>
<td>d</td>
<td>External Barg. Power</td>
<td>1.000</td>
<td>-</td>
<td>1.000</td>
<td>-</td>
</tr>
<tr>
<td>Vu</td>
<td>Value of Unemployment</td>
<td>0.000</td>
<td>-</td>
<td>0.000</td>
<td>-</td>
</tr>
<tr>
<td>σ</td>
<td>Measurement Error SD</td>
<td>10.186 *</td>
<td>0.001</td>
<td>10.036 *</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>No code</td>
<td>10.450 *</td>
<td>0.102</td>
<td>10.676 *</td>
<td>0.115</td>
</tr>
<tr>
<td></td>
<td>Agric., Mining, Elect.</td>
<td>10.450</td>
<td>-</td>
<td>12.236 *</td>
<td>0.120</td>
</tr>
<tr>
<td></td>
<td>Manufact.</td>
<td>10.450</td>
<td>-</td>
<td>10.690 *</td>
<td>0.101</td>
</tr>
<tr>
<td></td>
<td>Construction</td>
<td>10.450</td>
<td>-</td>
<td>10.352 *</td>
<td>0.105</td>
</tr>
<tr>
<td></td>
<td>Wholesale &amp; retail trade</td>
<td>10.450</td>
<td>-</td>
<td>10.810 *</td>
<td>0.102</td>
</tr>
<tr>
<td></td>
<td>Trans., Storage &amp; Comm.</td>
<td>10.450</td>
<td>-</td>
<td>11.259 *</td>
<td>0.109</td>
</tr>
<tr>
<td></td>
<td>FIRE</td>
<td>10.450</td>
<td>-</td>
<td>12.401 *</td>
<td>0.115</td>
</tr>
<tr>
<td></td>
<td>Services</td>
<td>10.450</td>
<td>-</td>
<td>9.492 *</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>-0.232 *</td>
<td>0.0080</td>
<td>-0.196 *</td>
<td>0.0076</td>
</tr>
<tr>
<td></td>
<td>Experience</td>
<td>0.052 *</td>
<td>0.0005</td>
<td>0.052 *</td>
<td>0.0005</td>
</tr>
<tr>
<td></td>
<td>Experience Squared/100</td>
<td>-0.117 *</td>
<td>0.0014</td>
<td>-0.116 *</td>
<td>0.0014</td>
</tr>
<tr>
<td></td>
<td>Experience X Female</td>
<td>-0.0006</td>
<td>0.0005</td>
<td>-0.0009</td>
<td>0.0076</td>
</tr>
<tr>
<td></td>
<td>Seniority</td>
<td>0.036 *</td>
<td>0.0025</td>
<td>0.035 *</td>
<td>0.0023</td>
</tr>
<tr>
<td></td>
<td>Seniority² / 100</td>
<td>-0.150 *</td>
<td>0.0138</td>
<td>-0.146 *</td>
<td>0.0128</td>
</tr>
<tr>
<td></td>
<td>Seniority X Fulltime</td>
<td>-0.026 *</td>
<td>0.0026</td>
<td>-0.025 *</td>
<td>0.0024</td>
</tr>
<tr>
<td></td>
<td>(Seniority² / 100) X Fulltime</td>
<td>0.125 *</td>
<td>0.0139</td>
<td>0.121 *</td>
<td>0.0129</td>
</tr>
<tr>
<td></td>
<td>Seniority X Female</td>
<td>-0.003 *</td>
<td>0.0007</td>
<td>-0.003 *</td>
<td>0.0007</td>
</tr>
<tr>
<td></td>
<td>Seniority X (E5+E6)</td>
<td>-0.009 *</td>
<td>0.0003</td>
<td>-0.0075 *</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>Fulltime Worker</td>
<td>0.472 *</td>
<td>0.0085</td>
<td>0.436 *</td>
<td>0.0080</td>
</tr>
<tr>
<td></td>
<td>Education &lt;= 9 yrs</td>
<td>-0.333 *</td>
<td>0.0054</td>
<td>-0.336 *</td>
<td>0.0053</td>
</tr>
<tr>
<td></td>
<td>10 or 11 yrs</td>
<td>-0.252 *</td>
<td>0.0039</td>
<td>-0.256 *</td>
<td>0.0039</td>
</tr>
<tr>
<td></td>
<td>12 or 13 yrs</td>
<td>-0.111 *</td>
<td>0.0036</td>
<td>-0.117 *</td>
<td>0.0036</td>
</tr>
<tr>
<td></td>
<td>14-16 yrs</td>
<td>0.069 *</td>
<td>0.0042</td>
<td>0.091 *</td>
<td>0.0041</td>
</tr>
<tr>
<td></td>
<td>&gt;= 17 yrs</td>
<td>0.277 *</td>
<td>0.0043</td>
<td>0.287 *</td>
<td>0.0042</td>
</tr>
</tbody>
</table>

-ln likelihood | 698,252 | 694,579

* significant at the 1% level.
Table 6. Joint Technology Parameter Estimates (Unproductive Seniority)

<table>
<thead>
<tr>
<th>Par</th>
<th>Variable</th>
<th>Estimate</th>
<th>Std.Err</th>
</tr>
</thead>
<tbody>
<tr>
<td>d</td>
<td>External Barg. Power</td>
<td>0.695</td>
<td>* 0.0054</td>
</tr>
<tr>
<td>V_U</td>
<td>Value of Unemployment</td>
<td>-1.519</td>
<td>* 0.3666</td>
</tr>
<tr>
<td>σ</td>
<td>Measurement Error SD</td>
<td>9.951</td>
<td>* 0.0010</td>
</tr>
<tr>
<td>β</td>
<td>No code</td>
<td>0.067</td>
<td>0.0544</td>
</tr>
<tr>
<td></td>
<td>Agric., Mining, Elect.</td>
<td>0.326</td>
<td>* 0.0609</td>
</tr>
<tr>
<td></td>
<td>Manufact.</td>
<td>0.161</td>
<td>* 0.0534</td>
</tr>
<tr>
<td></td>
<td>Construction</td>
<td>0.068</td>
<td>0.0551</td>
</tr>
<tr>
<td></td>
<td>Wholesale &amp; retail trade</td>
<td>0.162</td>
<td>* 0.0537</td>
</tr>
<tr>
<td></td>
<td>Trans., Storage &amp; Comm.</td>
<td>0.157</td>
<td>* 0.0552</td>
</tr>
<tr>
<td></td>
<td>FIRE</td>
<td>0.00031</td>
<td>0.0458</td>
</tr>
<tr>
<td></td>
<td>Services</td>
<td>-0.671</td>
<td>* 0.0937</td>
</tr>
<tr>
<td>γ</td>
<td>Female</td>
<td>-0.247</td>
<td>* 0.0140</td>
</tr>
<tr>
<td></td>
<td>Experience</td>
<td>0.060</td>
<td>0.0027</td>
</tr>
<tr>
<td></td>
<td>Experience²/100</td>
<td>-0.134</td>
<td>* 0.0062</td>
</tr>
<tr>
<td></td>
<td>Experience X Female</td>
<td>0.00147</td>
<td>* 0.0006</td>
</tr>
<tr>
<td></td>
<td>Fulltime Worker</td>
<td>0.479</td>
<td>* 0.0219</td>
</tr>
<tr>
<td></td>
<td>Education &lt;= 9yrs</td>
<td>-0.415</td>
<td>* 0.0198</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 or 11 yrs</td>
<td>-0.318 * 0.0150</td>
</tr>
<tr>
<td></td>
<td></td>
<td>12 or 13 yrs</td>
<td>-0.147 * 0.0078</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14-16 yrs</td>
<td>0.067 * 0.0055</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt;= 17 yrs</td>
<td>0.273 * 0.0131</td>
</tr>
<tr>
<td>Ψ</td>
<td>Seniority</td>
<td>0.087</td>
<td>* 0.0077</td>
</tr>
<tr>
<td></td>
<td>Seniority² / 100</td>
<td>-0.280</td>
<td>* 0.0422</td>
</tr>
<tr>
<td></td>
<td>Seniority X Fulltime</td>
<td>-0.002</td>
<td>0.0066</td>
</tr>
<tr>
<td></td>
<td>(Seniority² / 100) X Fulltime</td>
<td>0.032</td>
<td>0.0402</td>
</tr>
<tr>
<td></td>
<td>Seniority X Female</td>
<td>-0.032</td>
<td>* 0.0021</td>
</tr>
<tr>
<td></td>
<td>Seniority X (E5+E6)</td>
<td>0.0020</td>
<td>0.0008</td>
</tr>
<tr>
<td></td>
<td>-In likelihood</td>
<td>692,446</td>
<td></td>
</tr>
</tbody>
</table>

Estimates of industry-specific coefficients A not reported. Seniority coefficients in γ set to 0. * significant at the 1% level.
Table 7. Comparison of Seniority Estimates

<table>
<thead>
<tr>
<th>Par</th>
<th>Variable</th>
<th>Productive Seniority</th>
<th></th>
<th>Flexible Power</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Linear Power Tech.(^a)</td>
<td>Equal Power</td>
<td>Linear Power</td>
<td>Unprod. Seniority(^b)</td>
</tr>
<tr>
<td>d</td>
<td>External Barg. Power</td>
<td>1.000</td>
<td>0.734 (^*)</td>
<td>0.726 (^*)</td>
<td>0.695 (^*)</td>
</tr>
<tr>
<td>(\psi)</td>
<td>Value of Unemployment</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.654</td>
<td>-1.519 (^*)</td>
</tr>
<tr>
<td>(\sigma)</td>
<td>Measurement Error SD</td>
<td>10.036 (^*)</td>
<td>9.987 (^*)</td>
<td>9.961 (^*)</td>
<td>9.951</td>
</tr>
<tr>
<td>(\gamma)</td>
<td>Seniority</td>
<td>0.0348 (^*)</td>
<td>-0.0116</td>
<td>-0.0142 (^*)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Seniority(^2) / 100</td>
<td>-0.1457 (^*)</td>
<td>0.0386</td>
<td>0.0314</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Seniority X Fulltime</td>
<td>-0.0250 (^*)</td>
<td>0.0143</td>
<td>0.0202 (^*)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(Seniority(^2) / 100) X Fulltime</td>
<td>0.1211 (^*)</td>
<td>-0.0450</td>
<td>-0.0506</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Seniority X Female</td>
<td>-0.0029 (^*)</td>
<td>-0.0010</td>
<td>-0.0040 (^*)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Seniority X (E5+E6)</td>
<td>-0.0075 (^*)</td>
<td>-0.0032</td>
<td>-0.0028 (^*)</td>
<td>0</td>
</tr>
<tr>
<td>(\psi)</td>
<td>Seniority</td>
<td>-</td>
<td>0</td>
<td>0.022 (^*)</td>
<td>0.087 (^*)</td>
</tr>
<tr>
<td></td>
<td>Seniority(^2) / 100</td>
<td>-</td>
<td>0</td>
<td>0 (^*)</td>
<td>-0.280 (^*)</td>
</tr>
<tr>
<td></td>
<td>Seniority X Fulltime</td>
<td>-</td>
<td>0</td>
<td>0.007 (^*)</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(Seniority(^2) / 100) X Fulltime</td>
<td>-</td>
<td>0</td>
<td>0 (^*)</td>
<td>0.032</td>
</tr>
<tr>
<td></td>
<td>Seniority X Female</td>
<td>-</td>
<td>0</td>
<td>-0.026 (^*)</td>
<td>-0.032 (^*)</td>
</tr>
<tr>
<td></td>
<td>Seniority X (E5+E6)</td>
<td>-</td>
<td>0</td>
<td>0.005 (^*)</td>
<td>0.002</td>
</tr>
</tbody>
</table>

\(-\ln\) likelihood:

- 694,579  693,342  692,703  692,446  692,429

2*increment (\(\chi^2\)):

- -2,473.5  -1,277.9  -516  -32.7

\(^a\) Repeats portions of Table 5.
\(^b\) The preferred specification. Repeats portions of Table 6.

* significant at the 1% level
Table 8. Wage Variation and Assumed Technology

<table>
<thead>
<tr>
<th>source</th>
<th>Standard Deviations</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data</td>
<td>Joint</td>
<td>Linear</td>
</tr>
<tr>
<td>Payroll</td>
<td>overall</td>
<td>12.40</td>
<td>7.29</td>
</tr>
<tr>
<td></td>
<td>between workplaces</td>
<td>8.90</td>
<td>5.92</td>
</tr>
<tr>
<td></td>
<td>within</td>
<td>9.98</td>
<td>5.08</td>
</tr>
<tr>
<td>V(z)</td>
<td>overall</td>
<td>5.81</td>
<td>6.89</td>
</tr>
<tr>
<td></td>
<td>between workplaces</td>
<td>4.04</td>
<td>4.88</td>
</tr>
<tr>
<td></td>
<td>within</td>
<td>4.63</td>
<td>5.44</td>
</tr>
</tbody>
</table>