J Prod Anal (2009) 31:1–14
DOI 10.1007/s11123-008-0117-3

The effects of match uncertainty and bargaining on labor market outcomes: evidence from firm and worker specific estimates

Subal C. Kumbhakar · Christopher F. Parmeter

Published online: 31 October 2008

Abstract In this paper we examine wage dispersion in labor markets across currently employed workers. We argue that differences in the potential productivity of a match (typically assumed to be known in the previous literature) generates a surplus between the minimum wage the worker is willing to accept and the maximum wage the firm is willing to offer for the job. Existence of this surplus leads to wage dispersion due to negotiating over the amounts extracted by each agent. Our objective is to estimate the surplus extracted by each firm-worker pair and the effect of the net extracted surplus on the wage, for each firm-worker pair using the two-tier stochastic frontier model. An empirical application finds that, on average, firms paid workers less than their expected productivity. More specifically, at the mean, the net effect of productivity uncertainty leads to equilibrium wages which are 3.33% below the expected productivity of matches.

Keywords Expected productivity · Random matching · Two-tier frontier

JEL Classifications C2 · J3 · J15 · J41

S. C. Kumbhakar (✉)
Department of Economics, Binghamton University, Binghamton, NY 13902-6000, USA
e-mail: kkar@binghamton.edu

C. F. Parmeter
Department of Agricultural and Applied Economics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0401, USA
e-mail: parms@vt.edu

> The matching process that brings together workers and employers fails to weed out all bad matches
> (Pries 2004, p. 194).

1 Introduction

Labor markets are designed to pair workers and firms efficiently to ensure a productive outcome. However, there are many obstacles that stand in the way of a perfectly efficient market sorting process. A major impediment to an efficient labor market is heterogeneity across both the worker and the firm concerning the productivity of the job which has implications for the wages they are willing to accept/pay. The effects of an inefficient market sorting process can be seen on the wage outcomes of worker-firm pairings. Some pairs will be characterized by identical workers with differing wages, while other pairs will be composed of identical firms, paying different wages for the same job. Thus, the impact of matching on labor market outcomes is key in understanding why observed wages are dispersed. The ability to pull out the effect that match quality and bargaining have on wage outcomes due to agent-specific heterogeneity is the central focus of this paper.

We argue that labor markets do not work perfectly. What this tells us is that all job formations are not ideal; ideal in the sense that the value of the job is the same to both worker and firm. We use a standard search/matching/bargaining model that accounts for wage dispersion generated by bargaining over an expected surplus given that the quality of any job match is only known imperfectly to each agent. While a litany of theoretical models have attempted to discern the nature and causes of wage dispersion, to our knowledge there does not exist a microeconometric model that is designed to pull out the effects of this manifest match uncertainty.

Our attempt to capture wage dispersion generated by unknown match quality comes directly from the standard framework of search/matching/bargaining models that are commonly accepted as a genuine precipitator of wage dispersion. However, these models generate an equilibrium wage distribution for the labor market as a whole, while we focus on wage dispersion once person specific (and firm specific) effects have been accounted for. We modify the intuition of these models slightly to account for observed individual heterogeneity and provide a new twist to the exact nature of wage dispersion. While intimately linked to bargaining, our idea goes beyond simple negotiation of wages to a more important issue, viz., one of worker productivity.

Many authors have assumed that once a match is made the productivity of the job at hand is perfectly revealed (e.g., Flinn and Heckman 1982b; Hosios 1990; Mortensen and Pissarides 1994). Realistically the case is, however, more likely that there is an *expected* productivity of the match, known to both parties based on observables, but the *actual* productive value of the match is unknown.[1] Here we consider a static setting and are only concerned with the effect that this unknown productive value plays on current wage formation and dispersion. Intuitively, firms must caution against initially overpaying workers whom they will lock into long term contracts and cannot terminate simply because they overestimated their productivity, i.e., not all matches are 'good'. This leads to firms attempting to pay workers less than their perceived productive value until the true value of the worker is revealed. Alternatively, workers attempt to receive a wage higher than their productive value as a means to guard against possible failure within the job and having to resort to a lower paying job in the future.

Unlike the standard models that formulate the wage being paid as the sum of the reservation wage and the part of the surplus extracted by the worker, it is more meaningful to consider the fact that the wage paid is related to the expected productivity of the match with fluctuations around the mean. This view is motivated by the insights of match uncertainty dating back to the work of Johnson (1978), Jovanovic (1979a, b), Viscusi (1979, 1980a, b, c, 1983), Wilde (1980), and Flinn (1986).[2] While the recent resurgence has been in the spirit of understanding macroeconomic unemployment fluctuations through dynamic modelling procedures, our focus is more on the static microeconomic implications of productivity uncertainty. Specifically, our model is an attempt to uncover the effects of match quality on wage dispersion across worker and firm types.[3]

The objective of this paper is twofold: (i) we provide an intuitive explanation about the creation of productive surplus in the labor market in a reduced form setting which provides the impetus for the use of the two-tier estimation procedure, and (ii) we show how to obtain firm- and worker-specific extracted surpluses. Both points differ from the seminal two-tier papers of Polachek and Yoon (1987, 1996). First, Polachek and Yoon (1987) formulated the two-tier model to capture ignorance while searching for a job, we take this further to incorporate bargaining between worker and firm. Second, Polachek and Yoon (1987, 1996) only estimated the expected impacts of ignorance on behalf of the firm and the worker, we use the intuition of Jondrow et al. (1982) to derive observation specific expected values of the impact of bargaining.

The remainder of the paper is organized as follows: Sect. 2 provides an explanation for the existence of productivity uncertainty stemming from match inefficiency that has arisen in the labor economics literature. It continues with an investigation of a traditional textbook bargaining model and tries to imbed productivity uncertainty within it, generating a two-tier frontier model. Section 3 briefly reviews the structure and intuition of the two-tier frontier estimation procedure, showcasing the appeal of this estimation technique for the problem at hand. Section 4 presents the derivations of the conditional distributions that allow us to estimate agent-specific extracted surplus. Section 5 presents an empirical application that demonstrates the use of our decomposition, while our final thoughts and further avenues for extensions are contained in Sect. 6.

2 An overview of wage dispersion and match productivity

2.1 Wage dispersion–a short history

Given the extent of the labor market search literature, both theoretical and empirical, it is neither necessary nor desirable to discuss here all the existing avenues that have attempted to explain wage dispersion. A short list includes surveys by Rothschild (1973) (for markets in general), Lippman and McCall (1976) (for labor markets explicitly) and recently Eckstein and Van den Berg (forthcoming).

[1] It might be revealed at a later date or not revealed at all.

[2] Recently ideas paralleling these works can be found in Pries (2004) and Nagypál (2004). A different, but relevant idea on match uncertainty, miss-matching, can be found in Marimon and Zilblotti (1999).

[3] Given that we have a supply side dataset we leave the effects of firm type on uncertainty for future research.

The penultimate paper that was the motivation behind endogenous wage dispersion was Diamond (1971) with major theoretical advances put forth in Butters (1977), Burdett and Judd (1983), Albrecht and Axell (1984), Burdett and Vishwanath (1988), Mortensen and Pissarides (1994), and Burdett and Mortensen (1998). Excellent empirical inspections are found in Eckstein and Wolpin (1990), Bowlus et al. (1995), Abowd et al. (1998), and Dey and Flinn (2005); as well as contributions with both a theoretical and empirical flavor, Flinn and Heckman (1982a, b), Bontemps et al. (1999a, b), Postel-Vinay and Robin (2002, 2003), and Flinn (2006).[4]

One of the main insights of equilibrium wage dispersion is that there is not one key element that generates it, whether it be differences in productivity across firms, allowing on-the-job search, heterogeneity in reservation wages, search frictions, or some as of yet undiscovered reason. Even today models capable of resolving the 'Diamond Paradox' are still being developed, see Gaumont et al. (2006) and Shapiro (2006). From an empirical standpoint there have been many studies that attempt to quantify the magnitude of dispersion, the effects of minimum wages on wage dispersion, structural estimation of the search theoretic models listed previously, as well as many applications of standard regression techniques to uncover the sources of equilibrium wage dispersion.

2.2 Uncertainty of match productivity

After Diamond (1971) laid out a theoretical model of price adjustment that did not generate equilibrium price dispersion, labor economists and search theorists alike tried to come up with strategies that allowed for the existence of wage dispersion in a cross section of homogeneous workers. It was Rothschild (1973) who noted that any model of search that cannot generate an endogenous non-degenerate equilibrium wage distribution was unsatisfactory.

Here we argue that when the match value is imperfectly known, perhaps a better way to model the wage setting mechanism is through the expected value of the match with fluctuations around the mean predicated upon workers and firms attempting to shield themselves from "bad" matches, i.e., the larger the surplus the more likely a match is bad for one party or another. Note that from the perspective of wages, matches can be considered bad for one party or the other, but not both. By relaxing the assumption of perfect knowledge of the match value we can, in a reduced form setting, attribute equilibrium wage dispersion to negotiation over the initial, unknown "true" value of the match and determine which types of workers and/or firms are benefitting from this uncertainty.

Firms may try to mitigate the occurrence of 'bad' matches by posting specific skill requirements *a priori* (Acemoglu 1999; Mortensen and Pissarides 1994). However, this will not eliminate all lesser skilled workers from applying. Alternatively, workers may try to only apply for jobs in a certain region or of a specific type, in the hope of aligning themselves in such a way to reduce the probability of engaging in untenable employment, but this does not guarantee that a perfect match is conceived. Thus, while both workers and firms can attempt to insulate themselves there is no reason to expect the matching process to work perfectly and consequently existing market inefficiencies can result in inappropriate firm-worker pairings.

Our motivation of match uncertainty is through the plausible assumption of firm and worker heterogeneity. We assume that for a given set of characteristics for the worker, there is a distribution of productive ability. Similarly, for a given set of firm characteristics, there is a distribution of productive outcomes. Both of these distributions can be seen as proxies for limit wages. More productive firms will have the ability to pay more while more productive workers will place a higher value on their time, i.e., have higher reservation wages. Additionally, if we assume that the matching process is not perfect (Pries 2004), then we have the foundation for productivity uncertainty.

To explain this further, assume that r represents the distribution of reservation wages associated with worker heterogeneity for a given set of characteristics and p distinguishes the distribution of maximum wage offers for a given set of characteristics based on firm heterogeneity. A match function takes a draw from each distribution and one of two outcomes occurs: either $r > p$ and no match is consummated, or $p \geq r$ creating a surplus which the agents negotiate over. By repeating this process for any combination of worker and firm attributes one can obtain upper and lower bounds on the potential wage outcome given the observed characteristics of the agents. The difference in the upper and lower bound of the potential wage is necessary for surplus generation.

A key difference with our definition of wage dispersion compared to the previous literature is that we have wage dispersion for any combination of characteristics, while many of the previous empirical papers arbitrarily divide the labor market into distinct segments or search for a seemingly homogeneous sample to work with. In other words we are looking at vertical wage dispersion (based on specific agent characteristics) versus horizontal wage dispersion (based on the assumption that everyone is identical).

[4] For a current synopsis of the state of the literature see the special issue of the *European Economic Review (2006)* in honor of Dale Mortensen.

2.3 Capturing productivity uncertainty in a reduced form setting

One particular model[5] of interest that exemplifies the analysis of the impact of surplus extraction on wage variation is a matching/bargaining model with a distribution of productivity across job matchings. This distribution of productivities implies that there is also a distribution of surpluses that arises from these, as of yet, unknown productivities. Which agent (worker or firm) extracts more of the surplus has been shown to depend upon their bargaining power and information (see Osbourne and Rubinstein 1990, Chap. 5). Following Pissarides (2000, Chap. 1) the optimal wage rate is,

$$wage = \underline{wage} + \eta\left(\overline{wage} - \underline{wage}\right) \tag{1}$$

where $\underline{wage}$ represents the worker's reservation wage, $\overline{wage}$ represents the firm's maximum wage offer ($\overline{wage} \geqslant \underline{wage}$), and η ($0 \leq \eta \leq 1$), is the bargaining power of workers.[6] In (1), $\eta\,(\overline{wage} - \underline{wage})$ represents the share of the surplus created from the formation of the job match that the worker receives when the job is filled.[7] The reservation wage of a worker is unobserved, and due to bargaining it is unlikely that the observed wage is equal to the reservation wage. From an econometric standpoint (1) is not operational because the reservation wage and the maximum wage offer are unobserved. Another shortcoming of (1) is that it only provides insight on the impact of bargaining from a worker's standpoint, and, given productivity uncertainty it does not help in unveiling what happens when negotiations take place over an unknown surplus. Thus, even with an estimate of $\eta(\overline{wage} - \underline{wage})$ nothing could be said about what is happening on either the firm's side of the market, or of the productive value of the match.

To make (1) operational we transform the model so that it not only captures the impact of worker's bargaining, but the impact of firm's bargaining as well. For this, we first denote the expected productivity of the match conditional on a vector of characteristics x by $\mu(x) = E(\theta \mid x)$, where θ is the actual, but unknown, productivity of the match. We condition on x as it is intuitive, both from a worker's as well as a firm's perspective that observable characteristics may influence productivity and are certainly used in hiring decisions by firms. By construction, $\underline{wage} \leqslant \mu(x) \leqslant \overline{wage}$ for those matches where a job is consummated. Consequently, $(\overline{wage} - \mu(x))$ is the firm's expected surplus from the match and $(\mu(x) - \underline{wage})$ is the worker's expected surplus from the match. We then use these notions of surplus to rewrite (1) as

$$\begin{aligned} wage &= \mu(x) - \mu(x) + \underline{wage} \\ &\quad + \eta(\overline{wage} - \underline{wage} + \mu(x) - \mu(x)) \\ &= \mu(x) + (\underline{wage} - \mu(x)) + \eta(\overline{wage} - \mu(x)) \\ &\quad - \eta(\underline{wage} - \mu(x)) \\ &= \mu(x) + \eta(\overline{wage} - \mu(x)) - (1-\eta)(\mu(x) - \underline{wage}). \end{aligned} \tag{2}$$

In this framework the worker can raise his/her wage by extracting a share of the firm's surplus, $\eta(\overline{wage} - \mu(x)) \geqq 0$, while the firm can lower the wage paid by extracting a share of the worker's surplus, $(1-\eta)(\mu(x) - \underline{wage}) \geqq 0$. The size of the extracted surplus by the worker depends upon the bargaining power of the worker, η, and the firm's expected surplus, $(\overline{wage} - \mu(x))$. Similarly, the level of the surplus extracted by the firm depends upon the firm's bargaining power, $(1-\eta)$, and the worker's expected surplus, $(\mu(x) - \underline{wage})$.[8]

A more intuitive way to understand Eq. 2 is to note that at the time of the match neither the worker nor the firm knows the productive value of the match. What each knows is the expected productivity of the match given observable characteristics. Thus firms have an incentive to offer lower wages to protect themselves against bad hiring policies,[9] while workers have an incentive to negotiate for higher wages to avoid entering into inefficient contracts. Thus workers try to extract some (all) of the surplus the firm is obtaining by hiring the worker, while the firm is trying to extract some (all) of the surplus the worker is acquiring by accepting the job. So, heuristically, our model is drawing upon previous studies that also look at the productive value of the match but they treat it as known and/or arising from a distribution (see Flinn and Heckman (1982b), Acemoglu and Shimer (2000), Postel-Vinay and Robin (2002, 2003), (Flinn 2006) for more on bargaining and productivity of matches).

Simply put, the idea of unknown productivity is very similar to the market for new Ph.D.s. When universities make hiring decisions, the productive value of the match is not known until a later date (usually when the candidate goes up for tenure) so the negotiated salary depends upon the job candidates skills and the characteristics of the university

[5] We thank an anonymous referee for suggesting the following framework to us.

[6] Pissarides (2000) used p instead of $\overline{wage}$ and rU instead of $\underline{wage}$. Also, we used η to represent relative bargaining power of workers, instead of β as in Pissarides. Furthermore, in our modeling framework η can be observation specific.

[7] The actual wage also represents a weighted average of the maximum offer and the reservation wage.

[8] Using these notions we can define the expected productivity, $\mu(x)$ formally as the conditional expectation of $wage$ given x when either there is no surplus to extract or surplus extracted by workers and firms are equal.

[9] See Shapiro (2006) for a similar idea along these lines.

hiring the candidate, both of which influence the productivity, but are not a perfect indicator of it. Given that the wage contract signed is for a predetermined number of years, it is in the interest of the candidate to get as high an amount as possible while the university should offer a lower wage not knowing if the candidate will be a good researcher. Thus, the university can extract the candidate's surplus until a later date when the contract is renegotiated based on a clearer picture of the productivity of the match. However, the candidate can extract surplus from the university in a similar manner to guard against being exploited until a later date when the contract is renegotiated. Our framework takes into account the influence of each party on the extent of wage dispersion around the expected productivity.

The wage equation in (2) has three distinct components. The first term, $\mu(x)$, represents the expected (productive value) wage of the worker given his/her characteristics x and is labelled as the benchmark wage (market value of the match). The second component shows the surplus extracted by the worker, while the third term (without the minus sign) is the surplus extracted by the firm. Since both workers and firms bargain and the effect of workers bargaining (surplus extraction) is to increase wages while the opposite happens due to firms bargaining (surplus extraction), what is relevant from the practical point of view is the net surplus, $NS = \eta(\overline{wage} - \mu(x)) - (1-\eta)(\mu(x) - \underline{wage})$, which indicates the overall effect of bargaining on wage. Thus the observed wage can be more or less than the benchmark wage, $\mu(x)$ depending upon the sign of NS. Individually, neither of these components has a meaningful interpretation unless the other component is zero.

One model that may be seen as an overarching model for those described above is that of Postel-Vinay and Robin (2002, 2003). They setup a model that allows the equilibrium wage to be a random variable composed of a worker effect, a firm effect, and a market friction effect. Here we could say that the worker effect is positive, the firm effect is negative, and the market friction effect is two-sided. Thus, while our idea of a three component effect falls in line with their research, our model does not fit in with the structural nature of Postel-Vinay and Robin. What is interesting though is the fact that we can control for wage dispersion due to differences in worker characteristics while at the same time allowing for wage dispersion propagated by the three random factors. In their model only the three components lead to wage dispersion as there was no human capital accumulation thus accounting for age, education, tenure, and experience were not relevant to their discussion.[10]

One may even go as far as describing our method as a reduced form cousin to the modelling framework proposed by Flinn (1986), except that we explicitly allow productivity to depend on observable characteristics with fluctuations propagated by negotiations over the unknown surplus that exists due to the inherent uncertainty of the job outcome. Many of the same issues with identification that arose in his paper apply here as well. We have a reduced form equation that cannot nonparametrically identify the dispersion parameters of interest without distributional assumptions, a common theme in stochastic frontier models as well. And, while our model is incapable of providing estimates of the structural parameters associated with the Pissarides model from which it is derived, it does encapsulate information about an important aspect of the wage formation process, viz., uncertainty. This model is also one of the first to investigate the affect of productivity uncertainty in a microeconometric setting on equilibrium wage dispersion.[11] Are these dispersion effects large or small, do they represent a large share of the variation in wages or a small share relative to unobserved worker heterogeneity? While these issues have been around for quite some time we feel that our analysis is important in that it provides estimates of the impact of match quality and uncertainty on equilibrium wage dispersion.

Given that we are working in a reduced form setting a pertinent question becomes "Is our framework consistent with optimizing behavior of economic agents?" If we treat match productivity as uncertain and predicate wages upon productivity, making them uncertain as well, then both workers and firms attempt to optimize some expected criterion, rather than a deterministic one. Thus workers and firms bargain over the unknown surplus that is created from the match. Previous studies have used alternative rules, such as treating the surplus as known and splitting it in half (Flinn and Heckman 1982b), treating the surplus as known but bargaining over the relative amounts (Burdett and Mortensen 1998; Dey and Flinn 2005; and Flinn 2006), or treating surplus as known, but using other firms to enter in a Bertrand type game that will extract most or all of the surplus after the job has been accepted (Postel-Vinay and Robin 2002, 2003). Our approach is consistent with the last two of the three cases prevalent in the literature today, except that the surplus is *unknown*. To our knowledge no other study has treated the match surplus as unknown.

At this point it is worth mentioning that match uncertainty and negotiating power is dependent upon market structure. Some labor markets work better than others to ensure highly productive matches, for example the market for nuclear physicists is expected to work better than the

[10] Although they did mention that the next logical step in their modelling framework would be to introduce human capital accumulation.

[11] See Shi (2006) for a recent theoretical insight into the effects of productivity on wages.

market for high school teachers. There are fewer candidates for nuclear physicists and the costs of obtaining the training to become one is more tedious and costly than that of obtaining a degree that allows one to teach in a high school. Also, because teachers have summers off, there are more candidates interested in being a teacher than a physicist. Thus, an extension of the model developed here, would be to test whether match certainty and surplus extraction differs based on metrics of market structure. One interesting example would be the labor markets of professional sports.[12]

3 A two-tier stochastic frontier model

3.1 Linking the bargaining model to a two-tier stochastic frontier model

The salient feature of the model in the preceding section is that the outcome variable has a lower and an upper bound. Polachek and Yoon (1987, 1996) (PY hereafter) used this notion and developed the two-tier frontier model in which these bounds are taken into account when estimating the model. In this section we put the labor market match quality example, discussed in the preceding section, in general terms and write the regression equation for the ith observation ($i = 1,\ldots,n$) in the format of a two-tier stochastic frontier model, viz.,

$$y_i = x_i'\delta + \varepsilon_i, \tag{3}$$

where y_i is the outcome variable, x_i is a vector of covariates, δ is the corresponding parameter vector, and $\varepsilon_i = v_i - u_i + w_i$ represents the composite error term encapsulating the difference between the observed outcome variable and $\mu(x) = x'\delta$.[13] In a buyer and seller framework $\mu(x)$ is the market value of the good. The lower-boundary (frontier) of price (y) is the minimum that the seller is willing to accept and is given by $\mu(x) - u$, $u \geq 0$. Similarly, the upper boundary (frontier) indicates the maximum that the buyer is willing to pay and is given by $\mu(x) + w$, $w \geq 0$. Because of the natural upper and lower boundaries of the outcome variable the frontier terminology is aptly used by PY. Furthermore, the frontiers are also likely to be affected by the presence of the noise term, v, that can take both positive and negative values and hence capture effects of random shocks.

[12] We thank an anonymous referee for bringing this link with the model to our attention.

[13] Although in (3) we are assuming $\mu(x) = x'\delta$ thereby making the assumption that $\mu(x)$ is linear in parameters, the linearity assumption is not necessary for the frontier model to work. One can, in principle, assume any functional form on $\mu(x)$.

To link the bargaining model in (2) to the two-tier model in (3) we rewrite the regression counterpart of (2) as

$$wage = \mu(x) - u + w + v \tag{4}$$

where $u = (1-\eta)(\mu(x) - \underline{wage}) \geqq 0, w = \eta(\overline{wage} - \mu(x)) \geqq 0$, and v is the classical error term. As mentioned before the worker can raise his/her wage by extracting a share of the firm's surplus, denoted by w. Similarly, the firm can lower the wage paid by extracting a share of the worker's surplus, denoted by u. The size of these extracted surpluses depends on the bargaining power parameter, η, the firm's expected surplus, $(\overline{wage} - \mu(x))$, and the worker's expected surplus, $(\mu(x) - \underline{wage})$.

If one believes the argument put forth in Postel-Vinay and Robin (2002, 2003), then wages should represent the maximum possible productivity of a worker within a firm. On the other hand, if the scenario is closer to the textbook model of Pissarides (2000), then workers are paid their reservation wage plus something extra, representing the surplus extracted from the match. Both these models fit into the single-tier stochastic frontier models. So there are models which impose limits (at least implicitly) on how high or low wages can be. However, neither of these approaches is complete as each ignores one of the extremities relating to observed wages. Furthermore, stochastic frontier approaches which impose these limits in estimation are not explored in any of these models (see Kumbhakar and Lovell (2000) for a variety of such models).

A model closer to an actual frontier estimator is that of Flinn and Heckman (1982b). In their model they used the lowest wage in the sample (first order statistic) as an estimate of the (common) reservation wage for the population of interest. This paper was one of the first to consider how truncation at the reservation wage impacted the econometric analysis. While this paper is similar in spirit to our ideas, there are some dissimilarities, notably, the fact that we are focusing on each worker having a (possibly) different reservation wage, an upper bound on wages due to firms limiting their wage offers, and the introduction of worker/firm specific characteristics in each surplus term.

Thus the two-tier stochastic frontier technique seems an adequate avenue to go down when exploring bargaining within a matching framework. Although it is possible to estimate u and w (details are given in the following section), without further assumptions one cannot recover the relative bargaining parameter (η which can be worker-specific), worker's surplus $(\mu(x) - \underline{wage})$ and firm's surplus $(\overline{wage} - \mu(x))$ from the estimates of u and w. Thus our focus is not on the bargaining power *per se* but the surpluses extracted by the worker and firm for each observed match. In fact, estimates of these extracted surpluses are more useful than bargaining power because the end result of the bargaining process is to alter the wage in favor of a particular agent. In fact, the extent of productivity uncertainty on wages can be found from estimates on $w - u$.[14] If this turns out to be positive then workers hold an advantage due to productivity uncertainty (by increasing their wages),[15] while the opposite is true if this measure is negative.

3.2 Distributional assumptions and likelihood function

The δ parameters in (3) can be obtained using standard regression techniques. For example, the OLS procedure will give unbiased estimates of the slope coefficients. Since u and w are one-sided, $E(\varepsilon)$ may not be zero, even if $E(v) = 0$. Consequently, the OLS estimate of the intercept will be biased.[16] Thus, if the objective is to estimate the δ parameters then the OLS estimator of the slope coefficients will be unbiased (unless one thinks along the lines of tenure and wage dispersion being correlated, in which case an instrument will be needed) and consistent. However, we are interested in not only estimating the δ parameters but also the surplus extraction components, i.e., to disentangle the one-sided error terms from the composed error term ε. For this reason, we estimate the model using the maximum likelihood (ML) method based on the following distributional assumptions of the error components, viz., u, v, and w. We assume that: (i) $v_i \sim i.i.d.\ N(0,\ \sigma_v^2)$, (ii) $u_i \sim i.i.d.\ Exp(\sigma_u,\ \sigma_u^2)$,[17] (iii) $w_i \sim i.i.d.\ Exp(\sigma_w,\ \sigma_w^2)$, and (iv) the error components are distributed independently of each other and from the regressors, x. The use of an exponential distribution is commonplace in standard single-tier stochastic frontier studies when ML is used. It should be noted that the two-tier frontier distribution is nonparametrically underidentified. Thus these distributional assumptions are necessary to conduct an empirical analysis.

Based on the above distributional assumptions, it is straightforward (but tedious) to derive the probability density function (pdf) of ε_i, $f(\varepsilon_i)$ which is[18]

$$\begin{aligned} f(\varepsilon_i) &= \frac{\exp\{\alpha_i\}}{\sigma_u + \sigma_w}\Phi(\beta_i) + \frac{\exp\{a_i\}}{\sigma_u + \sigma_w}\int_{-b_i}^{\infty}\phi(z)dz \\ &= \frac{\exp\{\alpha_i\}}{\sigma_u + \sigma_w}\Phi(\beta_i) + \frac{\exp\{a_i\}}{\sigma_u + \sigma_w}\Phi(b_i) \end{aligned} \tag{5}$$

where $a_i = \frac{\sigma_v^2}{2\sigma_w^2} - \frac{\varepsilon_i}{\sigma_w}$; $b_i = \frac{\varepsilon_i}{\sigma_v} - \frac{\sigma_v}{\sigma_w}$; $\alpha_i = \frac{\varepsilon_i}{\sigma_u} + \frac{\sigma_v^2}{2\sigma_u^2}$; $\beta_i = -(\frac{\varepsilon_i}{\sigma_v} + \frac{\sigma_v}{\sigma_u})$.

The log likelihood function for a sample of n observations is

MLE

$$\ln L(x;\theta) = -n\ln(\sigma_u + \sigma_w) + \sum_{i=1}^{n}\ln[e^{\alpha_i}\Phi(\beta_i) + e^{a_i}\Phi(b_i)] \tag{6}$$

where $\theta = \{\delta,\ \sigma_v,\ \sigma_u,\ \sigma_w\ \}$. The ML estimates of all the parameters can be obtained by maximizing the above log likelihood function. It should be noted that identification of all three standard deviations is achieved due to the fact that σ_u and σ_w appear in the likelihood equation separately, i.e., σ_u appears in α_i and β_i while σ_w appears in a_i and b_i.

The reason for the assumption of the exponential distributions for surpluses extracted by the firm and the worker is that the likelihood function can be expressed in a closed form and identification of the variance parameters is trivial. Moreover, while matching is a random process, we assume that markets work well enough that the generated surplus for any match is low. Thus, high values of extracted surplus, while probable in our model, occur with low probability. However, this does not preclude the use of other distributions for the one-sided error terms, such as log normal, gamma, half normal, truncated normal, etc.[19]

4 Measuring observation-specific extracted surplus

The main objective of estimating a two-tier stochastic frontier function is to obtain observation-specific estimates of extracted surplus by the worker and the firm, i.e., u_i and w_i from the composed error term ε_i, an estimate of which is obtained from the residuals of the wage equation, $y_i - x_i'\hat{\delta}$. In the standard single-tier frontier model, decomposition of the residual into inefficiency and noise components was accomplished by Jondrow et al. (1982). Here, we extend their technique to obtain observation-specific estimates of u

[14] Recall that this is our measure of net surplus introduced prior.

[15] Possibly from having relatively more bargaining power than firms.

[16] Note that although $E(u)$ and $E(w)$ are non-zero, $E(w - u)$ might be zero. If this happens then the OLS estimator of the intercept will also be unbiased. This, however, does not mean that surplus does not exist in the market.

[17] Here $Exp(\sigma_z, \sigma_z^2)$ denotes a random variable z that is exponentially distributed with mean σ_z and variance σ_z^2.

[18] The full derivations of all results are contained in the Appendix.

[19] In fact, further research into the distributional assumptions of the two-tiered method, aside from making the technique more general, may also provide greater insight into wage variations once the error decomposition has taken place. See Tsionas (2008) for estimation of the two-tier model using Gamma distributions instead of exponentials. Also, the effect of distributional assumptions on the ranking of firms in efficiency studies has been found to have minor differences in the rankings of producers (see Kumbhakar and Lovell 2000, p. 90).

and w. For this, we need to derive the conditional distributions of u_i and w_i, viz., $f(u_i|\varepsilon_i)$ and $f(w_i|\varepsilon_i)$. These are

$$f(u_i|\varepsilon_i) = \frac{\lambda \exp\{-\lambda \mathrm{u_i}\}\Phi(\mathrm{u_i}/\sigma_\mathrm{v} + \mathrm{b_i})}{\chi_{1i}} \tag{7}$$

and

$$f(w_i|\varepsilon_i) = \frac{\lambda \exp\{-\lambda \mathrm{w_i}\}\Phi(\mathrm{w_i}/\sigma_\mathrm{v} + \beta_\mathrm{i})}{\chi_{2i}} \tag{8}$$

where $\lambda = \frac{1}{\sigma_u} + \frac{1}{\sigma_w}$, $\chi_{1i} = \Phi(b_i) + \exp\{\alpha_i - a_i\}\Phi(\beta_i)$, and $\chi_{2i} = \Phi(\beta_i) + \exp\{a_i - \alpha_i\}\Phi(b_i) = \exp\{a_i - \alpha_i\}\chi_{1i}$.

With these conditional distributions we derive the conditional expectation of u_i as

$$\boxed{E(u_i|\varepsilon_i)} = \frac{1}{\lambda} + \frac{\exp\{\alpha_\mathrm{i} - \mathrm{a_i}\}\sigma_\mathrm{v}[\phi(-\beta_\mathrm{i}) + \beta_\mathrm{i}\Phi(\beta_\mathrm{i})]}{\chi_{1i}} \tag{9}$$

and the conditional expectation of w_i as

$$\boxed{E(w_i|\varepsilon_i)} = \frac{1}{\lambda} + \frac{\sigma_v[\phi(-b_i) + b_i\Phi(b_i)]}{\chi_{1i}} \tag{10}$$

which can be used to obtain observation-specific estimates of u_i and w_i, respectively.

Since the dependent variable in many regressions is in logarithmic form, one could interpret $E(u)$ and $E(w)$—the point predictor of u and w—as the percentage reduction and increase in wage due to bargaining by the firm and worker, respectively, when u and w are small. To get an exact percentage measure of wage reduction due to a firm's ability to extract surplus, one could follow two alternative routes. First, use $100[e^z - 1]$, for $z = E(u|\varepsilon)$, $E(w|\varepsilon)$. However, $E(e^z) \neq e^{E(z)}$. Thus, one could use $E(\exp(-z))$ for $z = u, w$ for computing the exact percentage decrease(increase) in wage due to firm's (worker's) bargaining.

To obtain the formula for computing observation-specific measures of $\exp(-u)$ and $\exp(-w)$, we need to derive the following conditional expectation, viz., $E(e^{-u_i}|\varepsilon_i)$ and $E(e^{w_i}|\varepsilon_i)$, which are:

$$\boxed{E(e^{-u_i}|\varepsilon_i)} = \frac{\lambda}{1+\lambda}\frac{1}{\chi_{2i}}\Big[\Phi(b_i) + \exp\{\alpha_i - a_i\} \times \exp\{\sigma_v^2/2 - \sigma_v\beta_i\}\Phi(\beta_i - \sigma_v)\Big] \tag{11}$$

pct

and

$$\boxed{E(e^{-w_i}|\varepsilon_i)} = \frac{\lambda}{1+\lambda}\frac{1}{\chi_{1i}}\Big[\Phi(\beta_i) + \exp\{a_i - \alpha_i\} \times \exp\{\sigma_v^2/2 - \sigma_v b_i\}\Phi(b_i - \sigma_v)\Big]. \tag{12}$$

These conditional expectations can be used as the point estimators of $\exp(-u)$ and $\exp(-w)$. The decomposition of ε into u and w suggests that the analyst does not need to make *a priori* assumptions about the bargaining power that a worker or firm has. The decomposition gives us a way to assess the impact of bargaining on the overall wage, once the negotiations have taken place.

5 An empirical application

We operationalize the methods discussed in the preceding sections by estimating a wage function using the data from Blackburn and Neumark (1992). The outcome variable, following Blackburn and Neumark, is log wage and the x variables in the log wage regression are: education, work experience, tenure, squares of education, experience and tenure, age, a proxy variable for unmeasured ability (IQ), and dummy variables for working in an urban area, being married, and working in the south. To further control for unobserved, inherent correlates to wage variations, we also use the number of siblings the worker has, the birth order of the worker, and the years of mother's and father's education. Given that there are missing observations for mother's education, father's education, and the number of siblings, our dataset is reduced from 936 to 663 observations. Thus, we use a subsample of the original data used in Blackburn and Neumark (1992).

Table 1 presents results from the standard OLS wage regression that ignores the effect of bargaining on observed wages, except by incorporating a dummy variable for black workers. In this set-up the estimated coefficient of the black dummy suggests that, on average, black workers earn

Table 1 Estimates of log wage regression function (OLS)

Variable	Estimate	Variable	Estimate
Constant	3.429	IQ	0.004
	0.000		0.003
Education	3.039	Education^2	−1.220
	0.015		0.045
Experience	0.172	Experience^2	−0.022
	0.308		0.823
Tenure	0.161	Tenure^2	−0.068
	0.016		0.109
Age	0.508		
	0.011		
Married	0.198	South	−0.043
	0.000		0.169
Urban	0.199	Black	−0.109
	0.000		0.051
Number of siblings	0.009	Birth order	−0.017
	0.253		0.151
Mother's education	0.010	Father's education	0.005
	0.123		0.319

The natural logarithm of the monthly wage is used as the dependent variable in the regression and there are 663 observations. Asymptotic p values are reported beneath each estimate. The R^2 for this regression is 0.285

Table 2 Estimates of log wage regression (Two-tier frontier)

Variable	Estimate	Variable	Estimate
Constant	3.858	IQ	0.004
	0.000		0.000
Education	2.037	Education^2	−0.756
	0.071		0.171
Experience	0.282	Experience^2	−0.101
	0.071		0.276
Tenure	0.154	Tenure^2	−0.057
	0.016		0.136
Age	0.495		
	0.008		
Married	0.206	South	−0.035
	0.000		0.241
Urban	0.221	Black	−0.102
	0.000		0.052
Number of siblings	0.009	Birth order	−0.015
	0.223		0.147
Mother's education	0.008	Father's education	0.008
	0.147		0.131
σ_v	0.190	σ_u	0.221
	0.000		0.000
		σ_w	0.189
			0.000

The natural logarithm of the monthly wage is used as the dependent variable in the regression and there are 663 observations. Asymptotic *p* values are reported beneath each estimate

about 11% less than white workers, *ceteris paribus*. By construction this coefficient is group specific. Thus, if it represents the effects of bargaining, it is an average for the whole group of black workers. As a result, nothing can be said about the effect of bargaining on an individual worker's wage whether black or white. This drawback is eliminated in the two-tier frontier model where one can estimate the impact of bargaining on wage for each worker-firm pair. These observation-specific results can then be used, if desired, to examine whether a particular group (defined in any manner) has more (less) influence on wages, *ceteris paribus*, for the group as a whole.

Estimated parameters from the two-tier frontier function are presented in Table 2. We are not too concerned with the fact that our data does not come from a population of recently hired workers. The reason being that the impact of negotiations over wages can have effects throughout the tenure of a worker and as such we can still determine if wages are higher or lower that they should be due to negotiations at the time the job was offered/accepted.[20]

Also, our estimates are of surplus extracted due to unknown bounds on productivity on both sides of the match. As long as those bounds remain after workers accept a match, i.e. there is still some productive uncertainty after the worker has been at the firm *t* years after the initial match, then surplus extraction will take place and cause variations in wages which we can then attempt to estimate.

The deterministic part of the frontier model is the same as the OLS model. The parameter estimates from the OLS and frontier models are quite similar. This suggests that one can use either of these models if the objective is to determine the marginal effect of covariates. However, if the interest is to obtain the impact of bargaining on wages, it is necessary to use the two-tier frontier approach which provides deeper insights on the effect of bargaining on wages for each employee-employer pair.

From the estimates of σ_v, σ_u and σ_w, we find that the unexplained variation in log wage $(\sigma_v^2 + \sigma_u^2 + \sigma_w^2)$ is 0.121. Of this unexplained variation, 70.4% is due to bargaining.[21] From the estimate of $E(w - u) = \sigma_w - \sigma_u$, one can say whether, on average, bargaining affects wages or not, and if so, in what direction. On the other hand, if $\sigma_w - \sigma_u = 0$ then one would predict that, at the mean, wages are not affected by bargaining. This, however, does not mean absence of bargaining because a zero mean does not imply that the quartiles, for example, will be zero. To get an answer to this we need to resort to the worker-firm pair estimates from the two-tier frontier approach.

Details on surplus extraction results (the percentage change in wages), based on observation specific estimates of $E(u|\varepsilon)$ and $E(w|\varepsilon)$, are reported in Tables 3–5. These tables display percentage changes measured relative to the benchmark log wage estimated from $\widehat{\log \text{wage}} = x_i{}'\hat{\delta}$. Table 3 shows that at the mean, surplus extracted by firms decreased wages by 25.2%, while, surplus extracted by workers increased wages by 21.03%. These opposite effects led to a decrease in wages (estimated from $E((w - u)|\varepsilon)$ by 3.33% relative to benchmark wages, *ceteris paribus*.[22] The first quartile value of net surplus is −13.20% which means wages are at least 13.20% below the expected productive value of the match for 25% of the sample). The top (meaning a positive surplus extraction on behalf of the worker) 25% of the surplus extractions are at least 9.68% relative to the benchmark wage (meaning that wages for 25% of the sample, are increased by at least 9.68% above the expected productive value of the match). Thus, wages

[20] Calculations with a different data set, not reported here, suggest that there is an additional impact from being a new worker that lowers wages. The results are available upon request.

[21] In their 1987 (1996) paper, Polachek and Yoon found that 79.8% (98.5%) of the unexplained wage variation was due to incomplete information.

[22] If the goal is to obtain an estimate of the mean of the net effect, one can use the estimated value of $E(w - u) = \sigma_w - \sigma_u$ which is −0.0339. This does not require use of the observation-specific estimates of *w* and *u*.

Table 3 Surplus extracted by firms and workers

	Mean (%)	Q1 (%)	Q2 (%)	Q3 (%)
Workers: $\hat{E}(w \mid \varepsilon)$	18.9	11.4	14.4	21.2
Firms: $\hat{E}(u \mid \varepsilon)$	22.1	12.2	16.4	25.0
Net surplus: $\hat{E}((w-u) \mid \varepsilon)$[a]	−3.2	−13.7	−2.0	9.0
Workers: $\hat{E}(1-e^{-w} \vert \varepsilon)$	15.9	10.3	12.8	18.1
Firms: $\hat{E}(1-e^{-u} \vert \varepsilon)$	18.1	10.9	14.4	21.1
Net surplus: $\hat{E}((e^{-u}-e^{-w}) \vert \varepsilon)$[a]	−2.2	−10.8	−1.6	7.2

Since the dependent variable is in logarithms we convert the estimates in the first panel in percentage form using $100[e^z-1]$, where $z=\hat{E}(. \mid \varepsilon)$. The second panel estimates are multiplied by 100 to express them in percentage form

[a] The mean and quartiles of net surplus were constructed after calculating $\hat{E}((w-u) \mid \varepsilon)$ and $\hat{E}((e^{-u}-e^{-w})|\varepsilon)$

Table 4 Surplus extracted by firms and workers across race

	Mean (%)	Q1 (%)	Q2 (%)	Q3 (%)
White workers[a]				
Workers: $\hat{E}(w \mid \varepsilon)$	18.9	11.4	14.3	21.2
Firms: $\hat{E}(u \mid \varepsilon)$	22.1	12.2	16.5	24.9
Net surplus: $\hat{E}((w-u) \mid \varepsilon)$	−3.2	−13.5	−2.2	9.0
Black workers[b]				
Workers: $\hat{E}(w \mid \varepsilon)$	18.4	11.1	15.0	20.6
Firms: $\hat{E}(u \mid \varepsilon)$	21.3	12.3	15.6	27.4
Net surplus: $\hat{E}((w-u) \mid \varepsilon)$	−2.9	−16.3	−0.6	8.3
White workers[a]				
Workers: $\hat{E}(1-e^{-w} \vert \varepsilon)$	15.9	10.3	12.7	18.2
Firms: $\hat{E}(1-e^{-u} \vert \varepsilon)$	18.1	10.9	14.4	21.0
Net surplus: $\hat{E}((e^{-u}-e^{-w}) \vert \varepsilon)$	−2.2	−10.7	−1.7	7.3
Black workers[b]				
Workers: $\hat{E}(1-e^{-w} \vert \varepsilon)$	15.7	10.0	12.7	18.2
Firms: $\hat{E}(1-e^{-u} \vert \varepsilon)$	17.8	11.1	13.8	22.8
Net surplus: $\hat{E}((e^{-u}-e^{-w}) \vert \varepsilon)$	−2.1	−12.8	−0.4	6.7

Since the dependent variable is in logarithms we convert the estimates in the first panel in percentage form using $100[e^z-1]$, where $z=\hat{E}(. \mid \varepsilon)$. The second panel estimates are multiplied by 100 to express them in percentage form

The mean and quartiles of net surplus were constructed after calculating $\hat{E}((w-u) \mid \varepsilon)$ and $\hat{E}((e^{-u}-e^{-w})|\varepsilon)$

[a] There are 609 observations for white workers

[b] There are 54 observations for black workers

are not reduced for all workers, and in fact, some workers managed to negotiate for a substantial increase in their wages over their expected productive outcomes. The lower panel of Table 4 gives the same information. The only difference is in the calculation of the percentage figures. Note that the estimates are for each worker-firm pair. We provide a summary of these in Table 3.

In addition to determining the impact of bargaining on observed wages for all worker-firm pairs, one can analyze the impact of bargaining on wages across different groups. Here we focus on black versus white workers, and married versus single workers. These results are reported in Tables 4 and 5, respectively.

Table 5 Surplus extracted by firms and workers across marital status

	Mean (%)	Q1 (%)	Q2 (%)	Q3 (%)
Married workers[a]				
Workers: $\hat{E}(w \mid \varepsilon)$	18.8	11.4	14.3	21.1
Firms: $\hat{E}(u \mid \varepsilon)$	22.0	12.2	16.5	24.7
Net surplus: $\hat{E}((w-u) \mid \varepsilon)$	−3.2	−13.3	−2.2	9.0
Single workers[b]				
Workers: $\hat{E}(w \mid \varepsilon)$	19.6	11.2	14.3	21.1
Firms: $\hat{E}(u \mid \varepsilon)$	22.4	11.9	15.7	26.4
Net surplus: $\hat{E}((w-u) \mid \varepsilon)$	−2.9	−15.2	−0.7	10.5
Married workers[a]				
Workers: $\hat{E}(1-e^{-w} \vert \varepsilon)$	15.8	10.3	12.7	18.1
Firms: $\hat{E}(1-e^{-u} \vert \varepsilon)$	18.0	20.8	14.4	10.9
Net surplus: $\hat{E}((e^{-u}-e^{-w}) \vert \varepsilon)$	−2.2	−10.5	−1.7	7.2
Single workers[b]				
Workers: $\hat{E}(1-e^{-w} \vert \varepsilon)$	16.4	10.1	13.2	19.1
Firms: $\hat{E}(1-e^{-u} \vert \varepsilon)$	18.5	10.7	13.8	22.1
Net surplus: $\hat{E}((e^{-u}-e^{-w}) \vert \varepsilon)$	−2.1	−12.0	−0.6	8.4

Since the dependent variable is in logarithms we convert the estimates in the first panel in percentage form using $100[e^z-1]$, where $z=\hat{E}(. \mid \varepsilon)$. The second panel estimates are multiplied by 100 to express them in percentage form

The mean and quartiles of net surplus were constructed after calculating $\hat{E}((w-u) \mid \varepsilon)$ and $\hat{E}((e^{-u}-e^{-w})|\varepsilon)$

[a] There are 597 observations for married workers

[b] There are 66 observations for non-married workers

From the estimates of the percentage change in wage (net surplus extracted) it is clear that, on average, both whites and blacks are receiving a wage reduction, relative to the benchmark, after controlling for being black on expected productivity. Thus, at the mean, there is not a significant difference between the surplus extraction of black and white workers across both measures (−3.2% vs. −2.9% and −2.2% vs. −2.1%, respectively). What is clear however, is that looking at the tails of the surplus extraction distributions across measures, the lower quartile suggests that black workers have 2–3% more extracted from the benchmark, while at the upper quartile, white workers are able to extract about 1% more than black workers relative to the benchmark.

These results are not surprising. First, controlling for being black in the expected productivity regression, linked with the similar distributions of surplus extraction, suggests that while blacks may be paid lower due to expected productivity, additional surplus extraction, compared with whites is negligible. Second, assuming uncorrelatedness of model covariates with the error terms means that we should not expect there to be significant differences between the surplus extraction distributions.

Switching to surplus extraction based on marital status, we see a similar picture. Both measures, across marital status, evaluated at the mean suggest there is no difference in surplus extraction based on marital status. What's more, the distributions of surplus extraction are quite similar as well. Again, this points to the fact that we are controlling for marital status in the expected productivity benchmark and the individual level surplus extraction measures are treated as uncorrelated with the covariates of the model. If one wanted to explicitly allow for the level of surplus extraction to depend on specific covariates, then the parameters of the two one-sided distributions could be modelled as such. Alternatively, if one believed that expected productivity did not depend on race or marital status, then a similar type of analysis would be better suited to reveal if surplus extraction depended on these worker features.

6 Conclusions

Firms and workers valuation of a job are inherently different, which leads room for negotiations over how much should be paid for the task at hand. A worker (firm) wants to extract as much of the surplus of the firm (worker) as possible. The surplus extracted by the firm reduces the wage while the surplus extracted by worker increases the wage. The net effect on the observed wage depends on the sum of these two opposing effects. In this paper we used a model that can identify workers' and firms' surplus extractions from the other party. The proposed technique allows us to estimate not only agent specific surplus extracted, but the net surplus extracted for each transaction as well. Once this net impact has been constructed comparisons across different strata of workers and/or firms may lead to a characterization of which qualities lead to better outcomes in a particular market.

We used the two-tier stochastic frontier technique to estimate the parameters of the model and to obtain observation-specific measures of extracted surplus by both the worker and the firm. This measure allows a secondary analysis of potential sources of bargaining power in the market. This secondary analysis could also be used to discern if particular groups of workers/firms are consistently being exploited in the market, in terms of extracting a smaller share of the extant surplus created from the match.

We provided an empirical application to examine the effect of worker and firm bargaining on wages, after controlling for worker characteristics. We found that, not only does a significant surplus exist, but the impact of bargaining over this surplus has an asymmetric effect on wages at the mean. Indeed, at the mean, the net effect of surplus extraction (bargaining) by workers and firms decreased wages by 3.33% from the benchmark/market wage, while at the median, wages were reduced by 2.06% relative to the benchmark wage. Our ability to measure the effect of bargaining on wages for each worker-firm pair allowed us to correlate wage fluctuations to a worker's race. In our application, we found that, across the distribution of surplus extraction, *ceteris paribus*, there are no significant differences between either white and black workers and married/unmarried workers.

Although in this paper we discuss a labor market model that captures the idea of bargaining over the surplus generated due to worker and firm heterogeneity and match inefficiency, we believe that the modeling strategy is general enough to include many other markets. Some other examples where this technique may be of use are auctions, used car markets, and hedonic price models such as the residential housing market. Thus, although the two-tier frontier technique was originally conceived as a method to learn about the impact of incomplete information in the labor market, we trust the applicability of the model goes beyond its original intention.

Acknowledgements The authors thank two anonymous referees, Suqin Ge, Christopher Hanes, Daniel Henderson, Nicolai Kuminoff, Xiang Lie, Knox Lovell, and Solomon Polachek. Comments and suggestions from seminar participants at Syracuse University and SMU as well as Sandra Ahearn's help in proofreading the Appendix are gratefully acknowledged. The usual disclaimer applies.

Appendix

Derivations of selected equations

Derivation[23] of Eq. 5:

Beginning with the definition of the composed error term $\varepsilon_1 = v - u$, the marginal distribution of this is, following Kumbhakar and Lovell (2000),

$$f(\varepsilon_1) = (1/\sigma_u)\left(\Phi(-\varepsilon_1/\sigma_v - \sigma_v/\sigma_u)\exp\left\{\varepsilon_1/\sigma_u + \sigma_v^2/2\sigma_u^2\right\}\right). \tag{A.1}$$

The three component error may then be written as $\varepsilon = \varepsilon_1 + w$, which implies that $\varepsilon_1 = \varepsilon - w$, yielding the following joint distribution, $g(\varepsilon, w) = g(\varepsilon_1, w) \cdot |d\varepsilon_1/d\varepsilon| = g(\varepsilon_1, w) = f(\varepsilon_1) \cdot f(w)$. Upon integrating out w one obtains the marginal distribution of ε. This is done below.

[23] To avoid notational clutter we dropped the i subscript in all the derivations.

$$f(\varepsilon)=\int_0^{\infty}\frac{1}{\sigma_u}\left(\Phi\left(-\frac{\varepsilon_1}{\sigma_v}-\frac{\sigma_v}{\sigma_u}\right)\exp\left\{\frac{\varepsilon_1}{\sigma_u}+\frac{\sigma_v^2}{2\sigma_u^2}\right\}\right)\times\frac{1}{\sigma_w}\exp\left\{-\frac{w}{\sigma_w}\right\}dw$$

$$=\frac{1}{\sigma_u\sigma_w}\left[\exp\left\{\frac{\varepsilon}{\sigma_u}+\frac{\sigma_v^2}{2\sigma_u^2}\right\}\int_0^{\infty}\Phi\left(\frac{w}{\sigma_v}-\left(\frac{\varepsilon}{\sigma_v}+\frac{\sigma_v}{\sigma_u}\right)\right)\times\exp\left\{-w\left(\frac{1}{\sigma_w}+\frac{1}{\sigma_u}\right)\right\}dw\right]$$

$$=\frac{-1}{\sigma_u+\sigma_w}\left[\exp\left\{\frac{\varepsilon}{\sigma_u}+\frac{\sigma_v^2}{2\sigma_u^2}\right\}\int_0^{\infty}\Phi\left(\frac{w}{\sigma_v}-\left(\frac{\varepsilon}{\sigma_v}+\frac{\sigma_v}{\sigma_u}\right)\right)\times d\left(\exp\left\{-w\left(\frac{1}{\sigma_w}+\frac{1}{\sigma_u}\right)\right\}\right)\right]$$

$$=\frac{-\exp\{\alpha\}}{\sigma_u+\sigma_w}\left[\Phi(w/\sigma_v+\beta)\exp\{-w\lambda\}|_0^{\infty}-\int_0^{\infty}\phi(w/\sigma_v+\beta)\exp\{-w\lambda\}dw\right]$$

$$=\frac{\exp\{\alpha\}}{\sigma_u+\sigma_w}\left[\Phi(\beta)+\exp\{-\alpha\}\exp\left\{\frac{\sigma_v^2}{2\sigma_w^2}-\frac{\varepsilon}{\sigma_w}\right\}\times\int_0^{\infty}\frac{1}{\sigma_v}\phi\left(\frac{w}{\sigma_v}-\left(\frac{\varepsilon}{\sigma_v}-\frac{\sigma_v}{\sigma_w}\right)\right)dw\right]$$

$$=\frac{\exp\{\alpha\}}{\sigma_u+\sigma_w}\Phi(\beta)+\frac{\exp\{\mathrm{a}\}}{\sigma_u+\sigma_w}\int_{-b}^{\infty}\phi(z)dz$$

$$=\frac{\exp\{\alpha\}}{\sigma_u+\sigma_w}\Phi(\beta)+\frac{\exp\{\mathrm{a}\}}{\sigma_u+\sigma_w}\Phi(b)$$

where $\alpha=\frac{\varepsilon}{\sigma_u}+\frac{\sigma_v^2}{2\sigma_u^2};\quad \beta=-\left(\frac{\varepsilon}{\sigma_v}+\frac{\sigma_v}{\sigma_u}\right);$

$$\lambda=\frac{1}{\sigma_u}+\frac{1}{\sigma_w}\quad a=\frac{\sigma_v^2}{2\sigma_w^2}-\frac{\varepsilon}{\sigma_w};\quad b=\frac{\varepsilon}{\sigma_v}-\frac{\sigma_v}{\sigma_w}.\tag{A.2}$$

Derivation of Eqs. 7 and 8:

$$f(u|\varepsilon)=\frac{f(u,\ \varepsilon)}{f(\varepsilon)}=\frac{(\exp\{\mathrm{a}\}/\sigma_{\mathrm{u}}\sigma_{\mathrm{w}})\exp\{-\lambda\mathrm{u}\}\Phi(\mathrm{u}/\sigma_{\mathrm{v}}+\mathrm{b})}{(1/(\sigma_u+\sigma_w))[\exp\{a\}\Phi(b)+\exp\{\alpha\}\Phi(\beta)]}=\frac{\lambda\exp\{\mathrm{a}\}\exp\{-\lambda\mathrm{u}\}\Phi(\mathrm{u}/\sigma_{\mathrm{v}}+\mathrm{b})}{[\exp\{a\}\Phi(b)+\exp\{\alpha\}\Phi(\beta)]}=\frac{\lambda\exp\{-\lambda\mathrm{u}\}\Phi(\mathrm{u}/\sigma_{\mathrm{v}}+\mathrm{b})}{\chi_1}\tag{A.3}$$

where $\chi_1=\Phi\ (b)+\exp\{\alpha-a\}\Phi\ (\beta)$. Similarly,

$$f(w|\varepsilon)=\frac{f(w,\ \varepsilon)}{f(\varepsilon)}=\frac{(\exp\{\alpha\}/\sigma_{\mathrm{u}}\sigma_{\mathrm{w}})\exp\{-\lambda\mathrm{w}\}\Phi(\mathrm{w}/\sigma_{\mathrm{v}}+\beta)}{(1/(\sigma_u+\sigma_w))[\exp\{a\}\Phi(b)+\exp\{\alpha\}\Phi(\beta)]}=\frac{\lambda\exp\{\alpha\}\exp\{-\lambda\mathrm{w}\}\Phi(\mathrm{w}/\sigma_{\mathrm{v}}+\beta)}{[\exp\{a\}\Phi(b)+\exp\{\alpha\}\Phi(\beta)]}=\frac{\lambda\exp\{-\lambda\mathrm{w}\}\Phi(\mathrm{w}/\sigma_{\mathrm{v}}+\beta)}{\chi_2}\tag{A.4}$$

where $\chi_2=\Phi(\beta)+\exp\ \{a-\alpha\}\Phi\ (b)=\exp\ \{a-\alpha\}\chi_1$.

Derivation of Eqs. 9 and 10:

$$E(u|\varepsilon)=\int_0^{\infty}u\frac{\lambda\exp\{-\lambda\mathrm{u}\}\Phi(\mathrm{u}/\sigma_{\mathrm{v}}+\mathrm{b})}{\chi_1}du$$

$$=\frac{-1}{\chi_1\lambda}\left[\int_0^{\infty}\Phi(u/\sigma_v+b)\,d(\exp\{-\lambda u\})+\lambda\int_0^{\infty}\Phi(u/\sigma_v+b)\,d(u\exp\{-\lambda u\})\right]$$

$$=\frac{1}{\chi_1\lambda}\left[\Phi(b)+\int_0^{\infty}\exp\{-\lambda u\}\phi(u/\sigma_v+b)\,du/\sigma_v+\lambda\int_0^{\infty}u\exp\{-\lambda u\}\phi(u/\sigma_v+b)\,du/\sigma_v\right]$$

$$=\frac{1}{\chi_1\lambda}\left[\Phi(b)+\frac{\exp\{\alpha\}}{\exp\{a\}}\left[\int_0^{\infty}\phi(u/\sigma_v+b+\sigma_v\lambda)du/\sigma_v+\lambda\int_0^{\infty}u\phi(u/\sigma_v+b+\sigma_v\lambda)\,du/\sigma_v\right]\right]$$

$$=\frac{1}{\chi_1\lambda}\left[\Phi(b)+\frac{\exp\{\alpha\}}{\exp\{a\}}\left[\int_{-\beta}^{\infty}\phi(z)\,dz+\sigma_v\lambda\times\int_{-\beta}^{\infty}z\phi(z)dz+\lambda\sigma_v\beta\int_{-\beta}^{\infty}\phi(z)\,dz\right]\right]$$

$$=\frac{1}{\chi_1\lambda}[\Phi(b)+\exp\{\alpha-a\}\Phi(b)+\sigma_v\lambda\exp\{\alpha-a\}[\phi(-\beta)+\beta\Phi(\beta)]]$$

$$=\frac{1}{\lambda}+\frac{\sigma_v[\phi(-\beta)+\beta\Phi(\beta)]}{\chi_2}.\tag{A.5}$$

The derivation for $E(w|\varepsilon)$ follows similarly as:

$$
\begin{aligned}
E(w|\varepsilon) &= \int_0^\infty w \frac{\lambda \exp\{-\lambda w\}\Phi(w/\sigma_v + \beta)}{\chi_2}\, dw \\
&= \frac{-1}{\chi_2\lambda}\left[\int_0^\infty \Phi(w/\sigma_v+\beta)\, d(\exp\{-\lambda w\}) \right. \\
&\quad \left. + \lambda \int_0^\infty \Phi(w/\sigma_v+\beta)\, d(w\exp\{-\lambda w\})\right] \\
&= \frac{1}{\chi_2\lambda}\left[\Phi(\beta) + \int_0^\infty \exp\{-\lambda w\}\phi(w/\sigma_v+\beta)\, dw/\sigma_v \right. \\
&\quad \left. + \lambda \int_0^\infty w\exp\{-\lambda w\}\phi(w/\sigma_v+\beta)\, dw/\sigma_v\right] \\
&= \frac{1}{\chi_2\lambda}\left[\Phi(b) + \exp\{a-\alpha\} \right. \\
&\quad \times \left[\int_0^\infty \phi(w/\sigma_v+\beta+\sigma_v\lambda)\, dw/\sigma_v \right. \\
&\quad \left.\left. + \lambda\int_0^\infty w\phi(w/\sigma_v+\beta+\sigma_v\lambda)\, dw/\sigma_v\right]\right] \\
&= \frac{1}{\chi_2\lambda}\left[\Phi(b) + \exp\{a-\alpha\}\left[\int_{-b}^\infty \phi(z)\, dz \right.\right. \\
&\quad \left.\left. + \lambda\sigma_v\int_{-b}^\infty z\phi(z)dz + \lambda\sigma_v b\int_{-b}^\infty \phi(z)\, dz\right]\right] \\
&= \frac{1}{\chi_2\lambda}[\Phi(\beta) + \exp\{a-\alpha\}\Phi(b) + \sigma_v\lambda\exp\{a-\alpha\} \\
&\quad \times [\phi(-b) + b\Phi(b)]].
\end{aligned} \tag{A.6}
$$

Thus,

$$
E(w|\varepsilon) = \frac{1}{\lambda} + \frac{\sigma_v[\phi(-b)+b\Phi(b)]}{\chi_1}. \tag{A.7}
$$

Derivation of Eqs. 11 and 12:

$$
\begin{aligned}
E(e^{-u}|\varepsilon) &= \int_0^\infty e^{-u}\frac{\lambda e^{-\lambda u}\Phi(u/\sigma_v+b)}{\chi_1}du \\
&= \frac{\lambda}{\chi_1}\int_0^\infty e^{-(1+\lambda)u}\Phi(u/\sigma_v+b)du \\
&= \left(\frac{-\lambda}{\chi_1(1+\lambda)}\right)\int_0^\infty \Phi(u/\sigma_v+b)d(e^{-(1+\lambda)u}).
\end{aligned} \tag{A.8}
$$

Using integration by parts, we get

$$
\begin{aligned}
E(e^{-u}|\varepsilon) &= \left(\frac{-\lambda}{\chi_1(1+\lambda)}\right)\left[\Phi(u/\sigma_v+b)e^{-(1+\lambda)u}\Big|_0^\infty \right. \\
&\quad \left. - \int_0^\infty e^{-(1+\lambda)u}\phi(u/\sigma_v+b)du/\sigma_v\right] \\
&= \left(\frac{\lambda}{\chi_1(1+\lambda)}\right)\left[\Phi(b) + e^{\alpha-a+.5\sigma_v^2-\sigma_v\beta} \right. \\
&\quad \left. \times \int_0^\infty \phi(u/\sigma_v+(b+\sigma_v(1+\lambda)))du/\sigma_v\right],
\end{aligned} \tag{A.9}
$$

and using the change of variable, $z = \frac{u}{\sigma_v} + (b+\sigma_v(1+\lambda)) \Rightarrow dz = du/\sigma_v$, we have

$$
E(e^{-u}|\varepsilon) = \left(\frac{\lambda}{\chi_1(1+\lambda)}\right)\left[\Phi(b) + e^{\alpha-a+.5\sigma_v^2-\sigma_v\beta}\Phi(\beta-\sigma_v)\right]. \tag{A.10}
$$

For the derivation of Eq. 12 we follow the same procedure as follows:

$$
\begin{aligned}
E(e^{-w}|\varepsilon) &= \int_0^\infty e^{-w}\frac{\lambda e^{-\lambda w}\Phi(w/\sigma_v+\beta)}{\chi_2}dw \\
&= (\lambda/\chi_2)\int_0^\infty e^{-(1+\lambda)w}\Phi(w/\sigma_v+\beta)dw \\
&= \left(\frac{-\lambda}{\chi_2(1+\lambda)}\right)\int_0^\infty \Phi(w/\sigma_v+\beta)\, de^{-(1+\lambda)w}.
\end{aligned} \tag{A.11}
$$

Using integration by parts

$$
\begin{aligned}
E(e^{-w}|\varepsilon) &= \left(\frac{-\lambda}{\chi_2(1+\lambda)}\right)\left[\Phi(w/\sigma_v+\beta)e^{-(1+\lambda)w}\Big|_0^\infty \right. \\
&\quad \left. - \int_0^\infty e^{-(1+\lambda)w}\phi(w/\sigma_v+\beta)\, dw/\sigma_v\right] \\
&= \left(\frac{\lambda}{\chi_2(1+\lambda)}\right)\left[\Phi(\beta) + e^{a-\alpha-b\sigma_v+.5\sigma_v^2} \right. \\
&\quad \left. \times \int_0^\infty \phi\left(\frac{w}{\sigma_v} + (\beta+\sigma_v(1+\lambda))\right)dw/\sigma_v\right].
\end{aligned} \tag{A.12}
$$

Finally, using the change of variable, $z = \frac{w}{\sigma_v} + (\beta + \sigma_v(1+\lambda)) \Rightarrow dz = dw/\sigma_v$, we have

$$
E(e^{-w}|\varepsilon) = \left(\frac{\lambda}{\chi_2(1+\lambda)}\right)\left[\Phi(\beta) + e^{a-\alpha-b\sigma_v+.5\sigma_v^2}\Phi(b-\sigma_v)\right]. \tag{A.13}
$$

References

Abowd J, Kramarz F, Margolis D (1998) High wage workers and high wage firms. Econometrica 67:251–333

Acemoglu D (1999) Changes in unemployment and wage inequality: an alternative theory and some evidence. Am Econ Rev 89:1259–1278

Acemoglu D, Shimer R (2000) Wage and technology dispersion. Rev Econ Stud 67:585–607

Albrecht J, Axell B (1984) An equilibrium model of search employment. J Polit Econ 92:824–840

Blackburn M, Neumark D (1992) Unobserved ability, efficiency wages, and interindustry wage differentials. Q J Econ 107:1421–1436

Bontemps C, Robin J-M, Van den Berg GJ (1999a) An empirical equilibrium job search model with search on the job and heterogeneous workers and firms. Int Econ Rev 40:1039–1075

Bontemps C, Robin J-M, Van den Berg GJ (1999b) Equilibrium search with continuous productivity dispersion: theory and nonparametric estimation. Int Econ Rev 41:305–358

Bowlus A, Keifer N, Neumann G (1995) Estimation of equilibrium wage distributions with heterogeneity. J Appl Econ 10:S119–S131

Burdett K, Judd L (1983) Equilibrium price distributions. Econometrica 51:955–970

Burdett K, Mortensen D (1998) Wage differentials, employer size, and unemployment. Int Econ Rev 39:257–273

Burdett K, Vishwanath T (1988) Balanced matching and labor market equilibrium. J Polit Econ 96:1048–1065

Burdett K, Wright R (1998) Two-sided search and nontransferable utility. Rev Econ Dyn 1:220–245

Butters G (1977) Equilibrium distributions of sales and advertising prices. Rev Econ Stud 44:465–491

Cole H, Rogerson R (1999) Can the Mortensen-Pissarides matching model match the business-cycle facts? Int Econ Rev 40:933–959

Dey M, Flinn C (2005) An equilibrium model of health insurance provision and wage determination. Econometrica 73:571–627

Diamond P (1971) A model of price adjustment. J Econ Theory 3:156–168

Eckstein Z, Wolpin K (1990) Estimating a market equlibrium search model from panel data on individuals. Econometrica 58:783–808

Eckstein Z, Van den Berg G (forthcoming) Empirical labor search: a survey. J Econom

Flinn C (1986) Wage and job mobility of young workers. J Polit Econ S88–S110

Flinn C (2006) Minimum wage effects on labor market outcomes under search matching and endogenous contact rates. Econometrica 74:1013–1062

Flinn C, Heckman J (1982a) Models for the analysis of labor force dynamics. In: Rhodes G, Basmann R (eds) Advances in econometrics. JAI Press, London, CT

Flinn C, Heckman J (1982b) New methods for analyzing structural models of labor force dynamics. J Econom 18:115–168

Gaumont D, Schindler M, Wright R (2006) Alternative theories of wage dispersion. Eur Econ Rev 50:831–848

Hosios A (1990) On the efficiency of matching and related models of search and unemployment. Rev Econ Stud 57:279–298

Jondrow J, Lovell CAK, Materov IS, Schmidt P (1982) On the estimation of technical inefficiency in the stochastic frontier production function model. J Econom 19:233–38

Johnson W (1978) A theory of job shopping. Q J Econ 92:261–278

Jovanovic B (1979a) Firm specific capital and turnover. J Polit Econ 87:1246–1260

Jovanovic B (1979b) Job matching and the theory of turnover. J Polit Econ 87:972–990

Kumbhakar SC, Lovell CAK (2000) Stochastic frontier analysis. Cambridge University Press, New York.

Lippman S, McCall J (1976) The economics of job search: a survey, Part I. Econ Inq 14:155–189

Marimon R, Zilibotti F (1999) Unemployment vs. mismatch of talent: reconsidering unemployment benefits. Econ J 109:266–291

Meeusen W, van den Broeck J (1977) Efficiency estimation from Cobb-Douglas production functions with composed error. Int Econ Rev 18:435–444

Mortensen D, Pissarides C (1994) Job creation and job destruction in the theory of unemployment. Rev Econ Stud 61:397–415

Nagypál É (2004) Learning-by-doing versus learning about match quality: can we tell them apart? Northwestern University, Mimeo

Nelson P (1970) Information and consumer behavior. J Polit Econ 78(2):311–329

Osbourne MJ, Rubinstein A (1990) Bargaining and markets. Academic Press, San Diego

Pissarides CA (2000) Equilibrium unemployment theory. 2nd edn, MIT Press, Cambridge

Polachek S, Yoon BJ (1987) A two-tiered earnings frontier estimation of employer and employee information in the labor market. Rev Econ Stat 69:296–302

Polachek S, Yoon BJ (1996) Panel estimates of a two-tiered earnings frontier. J Appl Econ 11:169–178

Postel-Vinay F, Robin J-M (2002) Equilibrium wage dispersion with heterogeneous workers and firms. Econometrica 70:2295–2350

Postel-Vinay F, Robin J-M (2003) The distribution of earnings in an equilibrium search model with state-dependent offers and counter-offers. Int Econ Rev 43:989–1016

Pries M (2004) Persistence of employment fluctuations: a model of recurring job loss. Rev Econ Stud 71:193–215

Rothschild M (1973) Models of market organization with imperfect information: a survey. J Polit Econ 81(6):1283–1308

Shapiro J (2006) Wage and effort dispersion. Econ Lett 92:163–169

Shi S (2006) Wage differentials, discrimination and efficiency. Eur Econ Rev 50:849–875

Tsionas EG (2008) Maximum likelihood estimation of non standard stochastic frontier models using the Fourier transform. Working paper, Athens University of Business and Economics

Viscusi WK (1979) Job hazards and worker quit rates: an analysis of adaptive worker behavior. Int Econ Rev 20(1):29–58

Viscusi WK (1980a) A theory of job shopping: a Bayesian perspective. Q J Econ 94(3):609–614

Viscusi WK (1980b) Sex differences in worker quitting. Rev Econ Stat 62(3):388–398

Viscusi WK (1980c) Self-selection, learning-induced quits, and the optimal wage structure. Int Econ Rev 21(3):529–546

Viscusi WK (1983) Employment relationships with joint employer and worker experimentation. Int Econ Rev 24(2):313–322

Wilde L (1980) An information-theoretic approach to job quits. In: Lippman S, McCall J (eds) Studies in the economics of search. Amsterdam, North Holland