Unobserved Worker Quality and Inter-Industry Wage Differentials[†]

Suqin Ge and João Macieira

Department of Economics, Virginia Tech

December 29, 2013

Abstract

This paper provides quantitative assessment of two alternative explanations of inter-industry wage differentials: worker heterogeneity in the form of unobserved quality and firm heterogeneity in the form of firm's willingness to pay (WTP) for workers' productive attributes. We develop an empirical model of labor demand and apply a two-stage, non-parametric procedure to recover worker and firm heterogeneity. In the first stage we recover the unmeasured worker quality by estimating a nonparametric hedonic wage function. In the second stage we infer each firm's WTP parameters for worker attributes using first order conditions from the demand model. We apply our approach to quantify inter-industry wage differentials using individual data from NLSY79 and find that worker quality accounts for approximately two-third of the inter-industry wage differentials.

Keywords: wage determination, inter-industry wage differentials, labor quality, hedonic models.

JEL Codes: J31, J24, C51, M51

[†]We would like to thank Carlos Madeira, Ken Wolpin and the seminar and conference participants at Virginia Tech, the Hong Kong Economic Association conference in Tianjin and the Latin American Meetings of the Econometric Society (LAMES) in Santiago for their valuable comments and suggestions. Contact information: Ge, Department of Economics, Virginia Tech, Email: ges@vt.edu; Macieira, Department of Economics, Virginia Tech, Email: joaom07@vt.edu.

1 Introduction

Substantial evidence exists on large and persistent wage differentials among industries for workers with the same observed productivity characteristics such as education and experience. The (unexplained) inter-industry wage differentials have attracted the attention of economists for decades as they are used to examine the alternative theories of wage determination and the underlying forces of wage structural change.[1] Explanations for inter-industry wage differentials largely fall into two categories. The first one emphasizes the role of the worker-specific productive abilities that are not measured in the data (Murphy and Topel, 1987). The second one emphasizes the importance of firm-specific heterogeneity in the form of compensating wage differences (Rosen, 1986), efficiency wage (Katz, 1986; Krueger and Summers, 1988), and rent sharing (Katz and Summers, 1989; Nickell and Wadhwani 1990). Gibbons and Katz (1992) provide empirical assessment of the two explanations by following a sample of (approximately) exogenously displaced workers but remain agnostic that either explanation alone can fit the empirical evidence on inter-industry wage differentials.

There is continuing debate regarding how much the observed inter-industry wage differentials can be explained by the unobserved worker or firm characteristics. To disentangle simultaneous worker- and firm-level heterogeneity in wage determination, microdata matching characteristics of firms to characteristics of their workers are preferred. Abowd, Kramarz and Margolis (1999) are able to decompose inter-industry wage differences in France into a worker fixed effect and a firm fixed effect by using a large matched employer-employee panel data. However, such matched employer-employee panels are not frequently accessible to researchers.

In this paper, we develop an empirical model of labor demand and apply a two-stage, nonparametric procedure to recover unobserved worker and firm heterogeneity in a hedonic wage equation. First, we recover unobserved worker quality nonparametrically using an estimator based on results from Bajari and Benkard (2005) and Imbens and Newey (2009). This estimator exploits both the uniqueness of equilibrium wage function and its monotonicity in unobserved worker attributes to identify worker quality while allowing unobserved quality to be correlated with other observed worker characteristics such as education and experience. Second, we infer firm-specific willingness to pay (WTP) with respect to both observed and unobserved worker attributes nonparametrically by using model results relating WTP and first-order conditions for profit maximization. Once the unobserved worker and firm effects are identified, we can quantitatively assess their importance in explaining inter-industry wage

[1] Thaler (1989) presents a review on the debate whether the residual inter-industry wage differentials can emerge from a competitive equilibrium or simply reflect non-competitive forces such as efficiency wage. Katz and Auto (1999) provide a comprehensive survey on changes in the wage structure.

differentials using widely available individual data.

Since the pioneer work of Rosen (1974), hedonic models have been widely used in empirical literature. Our approach builds on the classic hedonic model and borrows insights from recent work on differentiated product demand estimation in industrial organization.[2] We model labor demand as a discrete choice of a bundle of worker attributes. Worker quality is modeled as a worker attribute unobserved by the econometrician but valued by employers. Recent advances in industrial organization have proposed nonparametric methods to identify product characteristics observed by the consumers but not by the researcher (e.g., Bajari and Benkard 2005). We apply these methods to recover worker quality. Similar to the hedonic literature, the marginal prices of worker characteristics are estimated as random coefficients in a hedonic wage function.

We use our estimates of worker and firm effects to analyze inter-industry wage differentials. Our labor demand model is estimated on individual data from NLSY79 to explore the importance of the two effects in wage determination. Our estimates show that worker effects are statistically more important than firm effects in explaining wages. This results is consistent with the finding in Abowd, Kramarz and Margolis (1999) using matched employer-employee panel data in France. We find unmeasured worker quality to account for about two-third of the inter-industry wage differential.

Observed worker characteristics that are supposed to account for productivity differences typically explain no more than 30 to 40 percent of the wage variations across workers. The existence of a large residual variance suggests differences in unmeasured worker ability: high-ability worker earn high wages. In our empirical analysis, we find that the percentage of explained wage differentials across workers nearly double when log wage regressions on observed worker attributes are augmented by the estimated unobserved worker quality. Large wage dispersion across employers is also consistent with inefficient matching, indicating that worker reallocation may improve labor market efficiency. Our empirical framework allows us to recover firm's WTP for worker productive attributes, which may be interpreted as the match value between employers and employees. Variations in WTP for worker attributes across firms of different industry affiliation imply that a reallocation of workers across firms may increase match efficiency.

[2] Most of the hedonic literature considers a market with a continuum of products and perfect competition and all product characteristics are assumed to be perfectly observed. Rosen's estimation strategy is criticized by Brown and Rosen (1982), Epple (1987), and Bartik (1987) that preference estimates are biased because consumers with a strong preference for a product characteristic would purchase more of that characteristic. Recent work by Bajari and Benkard (2005) relaxes these assumptions and proposes a hedonic model of demand for differentiated products which accounts for unobserved product characteristics and heterogeneous consumers. Ekeland, Heckman and Nesheim (2004) and Heckman, Matzkin and Nesheim (2010)also offer excellent discussion on the identification issues in the estimation of hedonic models.

This paper is organized as follows. In section 2 we present the hedonic labor demand model and discuss its properties. In section 3 we outline the estimation methods used to recover unobserved worker quality and employer preferences for worker attributes. In section 4 we describe the data used in our empirical analysis. Section 5 presents and discusses estimation results. Section 6 concludes and outlines possible extensions for future research. All derivations and auxiliary results can be found in the Appendixes.

2 A Model of Labor Demand

In this section we describe a labor demand model for heterogeneous workers. Consider an economy where labor markets are indexed by $t = 1, ..., T$. These markets are either a time series for a single labor market or a cross-section of markets. In each market there are $j = 1, ..., J_t$ workers and $i = 1, ..., V_t$ job vacancies. Each job vacancy is a single-worker firm, which decides whether to hire a worker to fill the vacancy.

Each worker is represented by a bundle of characteristics that potential employers value differently, and M of the characteristics can be observed by both the employer and the researcher. Let X_{jt} denote a $1 \times M$ vector of worker $j's$ observed characteristics. Examples of observed worker characteristics include education, work experience, and gender. We use a scalar ξ_{jt} to represent a characteristic of the worker that is observed only by the employer. The unobserved characteristic reflects the fact that there are some worker attributes, such as productive abilities, communication skills, and career ambition, that are valued by the employer but are often not observed by the researcher. For simplicity, we interpret the variable ξ_{jt} as representing unmeasured worker quality that is rewarded in labor markets.

The output produced by worker j at employer i in market t is given by the production function $F_i(E_{jt}, K_{it})$, where E_{jt} is the labor efficiency units of worker j and K_{it} is the composite non-labor input including all intermediate inputs and capital. The variable E_{jt} measures the different skill levels of labor in terms of different quantities of efficiency unit.[3] We denote the set of available labor efficiency units at time t by $\Xi_t \equiv \{E_{0t}, E_{1t}, ..., E_{J_t t}\}$, where E_{0t} represents no hiring.

Employers are profit maximizers that choose labor input E_{jt} and non-labor input K_{it}, given market wage rate w_{jt}, rental price r_{it} of non-labor input K_{it}, and output price p_{it}. Formally, employer i's problem is

$$\max_{(E_{jt}, K_{it}) \in \Xi_t \times \mathbb{R}_0^+} \pi_{it} = p_{it} F_i(E_{jt}, K_{it}) - w_{jt} - r_{it} K_{it}, \tag{1}$$

[3] Sattinger (1980, pp. 15–20) provides a review and discussion on the efficiency unit assumption.

where the production function $F_i(E_{jt}, K_{it})$ is assumed to be continuously differentiable and strictly increasing in K_{it}. The first order condition on K_{it} implicitly defines a unique employer-specific optimal choice of the composite non-labor input, given its rental price, an labor efficiency level, and the output price.

$$\frac{\partial \pi_{it}}{\partial K_{it}} = p_{it}\frac{\partial F_i}{\partial K_{it}} - r_{it} = 0 \Longrightarrow K^*_{it} = K^*_i(E_{jt}, p_{it}, r_{it}). \tag{2}$$

Replacing the optimal choice of non-labor input in (1), the employer's problem simplifies to choose an optimal labor input E_{jt}:

$$\max_{E_{jt}\in\Xi_t} \pi_{it}(E_{jt}) = R_{it}(E_{jt}) - w_{jt}, \tag{3}$$

where $R_{it}(E_{jt})$ is the employer-specific revenue per worker net of non-labor costs, i.e.,

$$R_{it}(E_{jt}) = p_{it}F_i(E_{jt}, K^*_i(E_{jt}, p_{it}, r_{it})) - r_{it}K^*_i(E_{jt}, p_{it}, r_{it}). \tag{4}$$

We model a worker's labor efficiency units as a function of her characteristics such that $E_{jt} = E(X_{jt}, \xi_{jt})$. Then the employer's decision becomes a discrete-choice problem of choosing at most one worker to maximize profit on the job vacancy:

$$\max_{j\in\{0,1,\ldots,J_t\}} \pi_{it}(X_{jt}, \xi_{jt}) = R_{it}(X_{jt}, \xi_{jt}) - w_{jt}. \tag{5}$$

If more than one worker generates the same profits for the employer, we assume the employer would randomly pick one to fill the vacancy. The option of not hiring is denoted by $j = 0$.

In the heterogeneous labor demand model outlined above, there is a unique equilibrium wage function $\mathbf{w}_t(X_{jt}, \xi_{jt})$ in each market t, mapping the set of worker characteristics onto the set of wages. The equilibrium wages have the following properties: (1) there is one wage for each bundle of worker characteristics; (2) the wage function is increasing in the unobserved worker quality. The following proposition establishes these results.

Proposition 1 *Suppose that $R_{it}(X_{jt}, \xi_{jt})$ is (i) Lipschitz continuous in (X_{jt}, ξ_{jt}) and (ii) strictly increasing in ξ_{jt}, for all employers $i = 1, ..., V_t$, then there exists a unique, Lipschitz-continuous equilibrium wage function $\mathbf{w}_t(X_{jt}, \xi_{jt})$ that is strictly increasing in ξ_{jt} for each market t.*

The proof is provided in Appendix A.[4]

[4] We follow a similar strategy taken by Bajari and Benkard (2005) in their differentiated product demand model.

Suppose that worker characteristic m, denoted by $x^c_{j,m,t}$, is a continuous variable and that worker j^* is profit maximizing for employer i. Then the following first-order conditions must hold:

$$
\begin{aligned}
\frac{\partial R_{it}(X_{j^*t}, \xi_{j^*t})}{\partial x^c_{j,m,t}} &= \frac{\partial \mathbf{w}_t(X_{j^*t}, \xi_{j^*t})}{\partial x^c_{j,m,t}}, &(6)\\
\frac{\partial R_{it}(X_{j^*t}, \xi_{j^*t})}{\partial \xi_{jt}} &= \frac{\partial \mathbf{w}_t(X_{j^*t}, \xi_{j^*t})}{\partial \xi_{jt}}. &(7)
\end{aligned}
$$

Thus a firm's optimal labor demand will be one at which the value the firm derives from the last unit of each worker characteristic is exactly equal to the implicit price it had to pay for that unit. If this were not so then the firm could increase their profits by choosing an alternative worker with different bundle of worker attributes.

Some restrictions on the revenue per worker function $R_{it}(X_{jt}, \xi_{jt})$ will be required for identification. We assume this function to be linear in (X_{jt}, ξ_{jt}) with firm-specific coefficients, that is,

$$R_{it}(X_{jt}, \xi_{jt}) \equiv \beta_{i,0} + X_{jt} \cdot \boldsymbol{\beta}_{i,X} + \beta_{i,\xi}\xi_{jt}. \tag{8}$$

Given this functional form assumption, the employer's problem in Equation (16) becomes

$$\max_{j \in \{0,1,\ldots,J_t\}} \beta_{i,0} + X_{jt} \cdot \boldsymbol{\beta}_{i,X} + \beta_{i,\xi}\xi_{jt} - \mathbf{w}_t(X_{jt}, \xi_{jt}). \tag{9}$$

The firm's first-order conditions in Equations (6) and (7) on any continuous characteristic $x^c_{j,m,t}$ and ξ_{jt} evaluated at the observed optimal choice j^* are

$$
\begin{aligned}
\beta_{i,x^c_{j,m,t}} &= \frac{\partial \mathbf{w}_t(X_{j^*t}, \xi_{j^*t})}{\partial x^c_{j,m,t}}, &(10)\\
\beta_{i,\xi} &= \frac{\partial \mathbf{w}_t(X_{j^*t}, \xi_{j^*t})}{\partial \xi_{jt}}. &(11)
\end{aligned}
$$

While seemingly arbitrary, the linearity assumption in Equation (8) can be derived under mild conditions on model primitives.[5] In what follows, we illustrate how the linear revenue function can be derived from common specifications of labor efficiency and the production

[5] The proposed functional form is not required for identification. Other parametric specifications could be considered, such as a linear function where continuous variables are in logarithms rather than in levels (e.g., Bajari and Benkard 2005, Bajari and Khan 2005). We tried this latter specification, but its performance on explaining inter-industry wage differentials was not significantly different than the linear-in-levels specification used in this paper. The linear-in-levels case has the advantage of clear interpretation of $\boldsymbol{\beta}_i$ as WTP vector for worker characteristics, so we focus our analysis on this specification.

function. We suppress the market subindex t in our notation for ease of exposition.

Consider the following specification for labor efficiency units of worker j with characteristics vector $\left(x_{j,1}, x_{j,2}, \cdots, x_{j,M}, \xi_j\right)$,

$$E_j = \rho_0 + \rho_1 x_{j,1} + \rho_2 x_{j,2} + \cdots + \rho_M x_{j,M} + \rho_\xi \xi_j, \quad \forall j = 1, ..., J. \tag{12}$$

In addition, consider a CES production function

$$F_i(E_j, K_i) = [\lambda_i E_j^{\sigma_i} + (1-\lambda_i) K_i^{\sigma_i}]^{1/\sigma_i}, \tag{13}$$

where λ_i governs the income shares between labor and non-labor inputs, and σ_i determines the elasticity of substitution between inputs.

The first-order condition of the employer's problem with respect to K_i implies that its optimal demands takes the form of $K_i^* = \delta_i E_j$, where

$$\delta_i = \left[\frac{\lambda_i}{\left(\frac{r_i}{p_i(1-\lambda_i)}\right)^{\sigma_i/(1-\sigma_i)} - (1-\lambda_i)}\right]^{1/\sigma_i}. \tag{14}$$

Profit from hiring worker j, given the optimal choice of non-labor input, becomes

$$\pi_{ij} = p_i F_i(E_j, \delta_i E_j) - w_j - r_i \delta_i E_j. \tag{15}$$

Therefore, under the CES assumption, the employer's problem simplifies to

$$\pi_i = \max_{j \in \{0,1,...,J)} \{\gamma_i E_j - w_j, 0\}, \tag{16}$$

where the profit of not hiring ($j = 0$) is equal to zero. Intuitively, γ_i represents the dollar value of the marginal productivity of labor efficiency units for employer i. Under the model primitives, this coefficient is given by

$$\gamma_i = p_i \left[\lambda_i + (1-\lambda_i)\delta_i^{\sigma_i}\right]^{1/\sigma_i} - r_i \delta_i. \tag{17}$$

Combining Equations (12) and (16) gives a parametric form of the revenue per worker function,

$$R(X_j, \xi_j; \boldsymbol{\beta}_i) = \gamma_i E_j = \beta_{i,0} + X_j \cdot \boldsymbol{\beta}_{i,X} + \beta_{i,\xi} \xi_j, \tag{18}$$

where the coefficient vector $\boldsymbol{\beta}_i$ is the product of the vector of efficiency unit coefficients in Equation (12) and γ_i. This employer-specific revenue function allows an intuitive interpretation for the parameter vector $\boldsymbol{\beta}_i$. The coefficients $\boldsymbol{\beta}_{i,X}$ and $\beta_{i,\xi}$ represent employer i's WTP for characteristic vector X_j and ξ_j, respectively, and we allow each firm to have a unique set of WTP parameters. When the optimal choice is not hiring, all the coefficients in the revenue function are equal to zero. Similar specifications are commonly used to estimate preference parameters in the literature on demand estimation in differentiated product markets (see Berry 1994, Berry, Levinsohn and Pakes 1995; Petrin 2002; Bajari and Benkard 2005; Bajari and Kahn 2005). This type of random coefficient models are considerably more flexible than standard logit or probit models.

3 Estimation of the Labor Demand Model

The equilibrium pricing function implied by most hedonic models is of the nonseparable form $Y = g(X, \varepsilon)$, where Y is the product price, X is a vector of observed characteristics, and ε is a variable representing unobserved attributes. Our equilibrium wage function also consists of a functional where X and ε are nonseparable.[6] A large body of literature examines the estimation and identification of both the function $g(.)$ and the unobserved term ε (e.g. Matzkin 2003; Chesher 2003; Chernozukov et al. 2007). While most estimators proposed in this literature allow for at most one variable in X to be correlated with ε (e.g. Bajari and Benkard 2005, Imbens and Newey 2009), our application contemplates two variables in X, both worker education and experience, to be correlated with the unobserved worker quality ε.

Our estimation strategy proceeds in two steps. In the first step, we recover the unobserved worker quality up to a normalization using nonparametric methods based on the identification results of Matzkin (2003).[7] To take into account the potential correlation between worker quality and other observed worker characteristics, we use an extended version of the estimators proposed by Bajari and Benkard (2005) and Imbens and Newey (2009). In the second step, we use the first-order conditions in Equations (10) and (11) to infer firm-specific parameters on their WTP for continuous worker characteristics.

3.1 Estimation of Unobserved Worker Quality

Since unobserved worker quality has no inherent units, we normalize ξ_{jt} to lie in the interval $[0, 1]$ by using a monotonic transformation $F_\xi(\xi_{jt})$, where $F_\xi(\xi_{jt})$ is the cumulative distrib-

[6] Chernozukov and Hansen (2005, p. 248) motivate their IV model of quantile treatment effects with a returns-to-training model whose variables have an interpretation similar to ours.

[7] Matzkin (2003) demonstrates that the unobserved component ε in a nonlinear function $Y = m(X, \varepsilon)$ is only identified up to a normalization.

ution function (cdf) of ξ_{jt}. For the case where observed characteristics X_{jt} are uncorrelated with ξ_{jt}, Bajari and Benkard (2005) show that $F_\xi(\xi_{jt}) = F_{w|x}(w_{jt}|X_{jt})$, where $F_{w|x}(.|.)$ denotes the cdf of wages conditional on worker characteristics. In the context of the labor demand model we consider, however, observable worker characteristics such as education and experience are likely correlated with the unobserved worker quality. To confront the endogeneity problem, we develop an estimator in the spirit of Bajari and Benkard (2005) and Imbens and Newey (2009).

A control variable V is a variable such that X and ε are independent conditional on V. Our first step of estimation builds on recent estimators that condition on control variables as an alternative to traditional IV estimators to deal with endogenous regressors (e.g. Blundell and Powell 2003, 2004, Imbens and Newey 2009, Bajari and Benkard 2005, Petrin and Train 2010, Farre, Klein and Vella 2010). Let X_0 and X_1 be sub-vectors of the observed characteristics vector such that $X = (X_0, X_1)$.[8] In addition, let $X_0 = (x_{01}, .., x_{0M_0})$ represent the variables in X that may be correlated with the unobserved quality ξ, where M_0 denotes the number of endogenous variables in X_0. Subvector X_1 represent the vector of exogenous variables. We assume that the researcher also observes a vector of instruments Z with dimension $G \geq M_0$.

Theorem 1 of Imbens and Newey (2009) shows that, when $M_0 = 1$, the variable $\eta_1 = F_{x_{01}|X_1,Z}(x_{01}|X_1, Z)$ is a control variable, such that X and ξ are independent conditional on η_1. We consider an extended setup for an arbitrary number of endogenous regressors. We specify reduced-form regressions

$$x_{0m} = h_m(X_1, Z, \eta_m), \quad m = 1, ..., M_0, \tag{19}$$

where η_m is an error term such that $(\xi, \eta_1, ..., \eta_{M_0})$ are jointly independent of (X_1, Z), and each $h_m(.)$ is an unknown function strictly increasing in η_m. The following proposition shows that $(\eta_1, ..., \eta_{M_0})$ are control variables that can be used in the estimation of the unobserved worker quality ξ after a normalization.

Proposition 2 *Let $F_{x_{0m}|X_1,Z}(.|.)$ denote the cdf of the endogenous characteristic x_{0m} conditional on the vector of exogenous characteristics X_1 and an instrument set Z. If each η_m is normalized to lie in the interval $[0,1]$ such that, for each $m = 1, ..., M_0$, $\eta_m = F_{x_{0m}|X_1,Z}(x_{0m}|X_1, Z)$, then X and ξ are independent conditional on $\eta = (\eta_1, ..., \eta_{M_0})$ and,*

$$\xi = \int_{\boldsymbol{\eta} \in [\mathbf{0},\mathbf{1}]^{M_0}} F_{w|x,\boldsymbol{\eta}}(w|X, \eta) d\eta. \tag{20}$$

[8] To simplify notation, we suppress both the individual subindex j and the market subindex t in what follows.

Our proof extends the Theorem 1 in Imbens and Newey (2009) and the Theorem 4 of Bajari and Benkard (2005) by allowing for multiple endogenous characteristics, and it is provided in Appendix A.

The unobserved worker quality can be recovered in three steps. First, for each endogenous variable indexed by $m = 1, ..., M_0$, we estimate the values of η_m using an empirical analog of $F_{x_{0m}|X_1,Z}(.|.)$. Second, we use the recovered series for η_m to nonparametrically estimate $F_{w|x,\boldsymbol{\eta}}(.|.)$, the integrand function in Equation (20). Third, the estimates of worker quality are obtained by integrating η out using Halton draws of the M_0-dimensional unit cube.[9] The same procedure is applied to all workers $j = 1, ..., J_t$ and markets $t = 1, ..., T$.

Over the past few decades, a number of nonparametric methods, such as kernel method and series estimators, have been proposed to estimate conditional cdf. Imbens and Newey (2009) document that series estimators are preferable in empirical frameworks similar to ours. Within the class of series estimators, the mixtures of normal distributions are a frequently used nonparametric estimator (e.g., Bajari, Fox and Ryan, 2007; Bajari et al. 2011) because of its desirable approximation and consistency properties (e.g., Norets 2010). For our application, we have adopted this type of estimator as it fits the data well and is computationally more tractable for the numeric integration in Equation (20) than other methods.

More specifically, our estimator for the conditional probability distribution function (pdf) $\hat{f}$ of a variable Y, given a $1 \times H$ vector of covariates U, is a weighted mixture of normal densities

$$\hat{f}(Y|U;\boldsymbol{\theta}) \equiv \sum_{r=1}^{R(N)} \alpha_r(U, \theta^{\alpha})\phi(Y|\mu_r, \sigma_r), \tag{21}$$

where $R(N)$ represents the (integer) number of normal densities as an (increasing) function of sample size N, $\boldsymbol{\theta}$ is the vector of parameters of the density function, and $\phi(.|\mu_r, \sigma_r)$ is a normal density with mean μ_r and standard deviation σ_r. The corresponding conditional cdf of Y is

$$\hat{F}(Y|U;\boldsymbol{\theta}) \equiv \sum_{r=1}^{R(N)} \alpha_r(U, \boldsymbol{\theta}^{\alpha})\Phi(Y|\mu_r, \sigma_r), \tag{22}$$

where $\Phi(.|\mu_r, \sigma_r)$ denotes the cdf of the same normal distribution. Each normal density in Equation (21) is weighted by a multinomial logit function $\alpha_r(U, \boldsymbol{\theta}^{\alpha})$ with a $(H + 1) \times 1$

[9] Halton draws consist of a sequence of numbers within the unit interval that uses a prime number as its base (Halton 1960). For example, the first 8 numbers in the sequence corresponding to base 3 are 1/3, 2/3, 1/9, 4/9, 7/9, 2/9, 5/9, 8/9. To span the domain of the M_0-dimensional unit cube it suffices to form Halton draws using different prime numbers for each dimension. Its advantages over random draws from $U[0, 1]$ in terms of lower variance and fewer number of draws have been documented in recent studies (Petrin and Train, 2010; Bhat 2001).

parameter vector $\boldsymbol{\theta}^{\alpha}$ defined as

$$\alpha_r(U, \boldsymbol{\theta}^{\alpha}) = \begin{cases} \dfrac{1}{1 + \sum_{l=2}^{R(N)} \exp\left(\theta_{0,l}^{\alpha} + U \cdot \boldsymbol{\theta}_{U,l}^{\alpha}\right)} & if\ r = 1, \\ \dfrac{\exp\left(\theta_{0,r}^{\alpha} + U \cdot \boldsymbol{\theta}_{U,r}^{\alpha}\right)}{1 + \sum_{l=2}^{R(N)} \exp\left(\theta_{0,l}^{\alpha} + U \cdot \boldsymbol{\theta}_{U,l}^{\alpha}\right)} & if\ r = 2, ..., R(N). \end{cases} \tag{23}$$

Norets (2010) demonstrates that this specification approximates arbitrarily well the true conditional pdf of Y given U.

In each market $t = 1, ..., T$, our maximum likelihood (ML) estimator for the pdf of an endogenous attribute $x_{0,m}$ conditional on exogenous worker characteristics X and an instrument set Z is defined as

$$\hat{\boldsymbol{\theta}}_{x_{0,m}} \equiv \arg\max_{\boldsymbol{\theta}} \sum_{j=1}^{J_t} \hat{f}(x_{0,m,j,t} | X_{1,j,t}, Z_{jt}; \boldsymbol{\theta}).^{10} \tag{24}$$

Our maximum likelihood estimator for the pdf of wages conditional on observed worker attributes X and control variables η is

$$\hat{\boldsymbol{\theta}}_w \equiv \arg\max_{\boldsymbol{\theta}} \sum_{j=1}^{J_t} \hat{f}(w_{jt} | X_{jt}, \eta_{jt}; \boldsymbol{\theta}) \tag{25}$$

After obtaining estimates of $\hat{\boldsymbol{\theta}}_{x_{0,m}}$ for each $m = 1, ..., M_0$, the corresponding estimate for the control variable value for each worker j in market t is $\eta_{m,j,t} = \hat{F}_{x_{0,m}|X_1,Z}(x_{0,m,j,t} | X_{1,j,t}, Z_{jt}; \hat{\boldsymbol{\theta}}_{x_{0,m}})$. With control variable estimates of $\eta_{m,j,t}$ for all m, $\hat{\boldsymbol{\theta}}_w$ is obtained by solving Equation (25). Then we can estimate the unobserved quality of each worker j in market t by using Equation (20), and we have

$$\hat{\xi}_{jt} = \int_{\boldsymbol{\eta} \in [\mathbf{0}, \mathbf{1}]^{M_0}} \hat{F}(w_{jt} | X_{jt}, \eta; \hat{\boldsymbol{\theta}}_w) d\eta. \tag{26}$$

3.2 Estimation of Firm WTP Parameters

The firm's labor demand problem described in Equation (5) is characterized by the revenue per worker function $R_i(X_j, \xi_j)$, which can be derived from given model primitives on

[10] We need to select $R(N)$ in order to obtain estimates of distribution parameters. This is analogous to the selection of smoothing parameters of other nonparametric estimators such as kernels or local linear regressions. Following Bajari and Benkard (2005) and Bajari and Khan (2005), among others, we guide our choice by visual inspection of the estimates. Our starting point for choosing the number of normal distribution in the mixture is $R(N) = int(\sqrt{N/2})$,a rule of thumb proposed by Mardia, Kent and Bibby (1979).

labor efficiency and the production function. We consider a revenue per worker function $R(X_j, \xi_j; \beta_i)$ that is linear in β_i.[11]

Equation (10) suggests that if we could recover an estimate of $\partial \mathbf{w}_t(X_{j^*t}, \xi_{j^*t})/\partial x^c_{j,m,t}$, then we could learn a firm's random coefficient or WTP for worker characteristic m. In our micro data, we do observe the worker characteristics employed by each firm. We can flexibly estimate $\partial \mathbf{w}_t(X_{j^*t}, \xi_{j^*t})/\partial x^c_{j,m,t}$ by using nonparametric methods. After we recover the unobserved worker quality, we can estimate a firm's WTP for unobserved quality based on $\partial \mathbf{w}_t(X_{j^*t}, \xi_{j^*t})/\partial \xi_{jt}$.

A practical, flexible way to quantify wage function derivatives at each point in the data is to apply local linear regression methods to data on wages, observed worker attributes and unobserved quality estimates. Bajari and Khan (2005) also resort to this approach to estimate the hedonic price function and quantify derivatives of this function. However, two important differences apply. First, they assume that ξ is independent of all observed characteristics X. While this is an acceptable assumption in their housing demand model, it is not reasonable for our application due to endogeneity concerns about schooling and experience. Second, their direct application of local linear regression to housing data does not separate the derivative $\partial \mathbf{w}_t(X_{j^*t}, \xi_{j^*t})/\partial \xi_{jt}$ from the value of ξ_{jt}. We separate the two values by first quantifying unobserved worker skill using the methods described above, and then treat the estimated values of $\xi_{j,t}$ as an extra regressor for local linear regression.

Specifically, for a given t, the wage function at each data observation $j^* \in \{1, ..., J_t\}$ (locally) satisfies the equation

$$w_{j^*,t} = b_{j^*,0} + b_{j^*,1} x_{j^*,1,t} + ... + b_{,j^*,M} x_{j^*,M,t} + b_{j^*\xi} \xi_{j^*,t}, \tag{27}$$

where each coefficient $b_{j^*,m}$ represents the derivative of w with respect to characteristic m at point j^*. Intuitively, this corresponds to the fact that, by a first-order Taylor expansion argument, a function w at point (X_{j^*t}, ξ_{j^*t}) is well approximated by a tangent hyperplane in a neighborhood of the function value at that point, w_{j^*t}.[12]

In the context of nonparametric regression, Fan and Gijbels (1996) provide a formula for the coefficients in Equation (27) for each observation j^*. Denote the $J_t \times 1$ vector of all wages by w_t and the vector stacking all coefficients by $\mathbf{b}_{j^*}$, the solution to the latter is

$$\mathbf{b}_{j^*} = \left(Z_t^T \Omega_t Z_t\right)^{-1} Z_t^T \Omega_t \mathbf{w}_t, \tag{28}$$

[11] Recall that this assumption nests the simple specifications on labor efficiency in Equation (12) and the CES production function in Equation (13) as especial cases.

[12] See Judd (1998) and Fan and Gijbels (1996) for a discussion.

where Z_t and Ω_t are matrices defined as

$$Z_t = [\mathbf{1}\ \ \mathbf{z}_t] = \begin{bmatrix} 1 & (x_{1,1,t} - x_{j^*,1,t}) & \dots & (x_{1,M,t} - x_{j^*,M,t}) & (\xi_{1,t} - \xi_{j^*,t}) \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & (x_{J_t,1,t} - x_{j^*,1,t}) & \dots & (x_{J_t,M,t} - x_{j^*,M,t}) & (\xi_{J_t,t} - \xi_{j^*,t}) \end{bmatrix}, \quad (29)$$

$$\Omega_t = \mathbf{diag}\left(K_{\mathbf{h}}(\mathbf{z}_t)\right). \quad (30)$$

$K_{\mathbf{h}}(\mathbf{z}_t)$ is a multivariate kernel function with smoothing parameter vector h, and we let $K_{\mathbf{h}}$ be the product of $M+1$ standard normal densities. Similar to other practical application of local linear regression with several covariates, the bandwidth vector h is selected by visual inspection of estimates.[13]

As the values of $\mathbf{b}_{j^*}$ are estimates of the derivatives of the wage function, they consist also of estimates to the values of β_i for continuous attributes implied by our linear profit specification. This allows us to recover the unobserved, firm-specific WTP parameters $\beta_{i,x^c_{j,m,t}}$.

For worker characteristics that take on discrete values, point identification of random coefficients on these characteristics cannot be achieved by using first-order conditions similar to Equation (10). Instead, we can only establish bound estimates for these coefficients using the condition that firm i's choice of the discrete characteristic observed in the data maximizes profit in Equation (**??**). To illustrate this point, suppose that firm i hires worker j^*. Let $\hat{X}_i$ and $\bar{X}_i$ denote the vectors of observed characteristics with $female = 1$ and $female = 0$, respectively, and all other elements equal to the corresponding observed attributes in vector X_{j^*}. The implicit price faced by employer i for a female worker is then $w_t(\hat{X}_i, \xi_{j^*t}) - w_t(\bar{X}_i, \xi_{j^*t})$. Denote $\beta_{i,f}$ as the coefficient for female dummy in the revenue function. Profit maximization problem in Equation (**??**) implies that $\beta_{i,f} > w_t(\hat{X}_i, \xi_{j^*t}) - w_t(\bar{X}_i, \xi_{j^*t})$ if worker j^* is female, and $\beta_{i,f} \leq w_t(\hat{X}_i, \xi_{j^*t}) - w_t(\bar{X}_i, \xi_{j^*t})$ otherwise. That is, if employer i hires a female worker, then we can infer that i's WTP for this characteristic exceeds the implicit price for this characteristic.[14]

A firm's WTP for a discrete worker characteristic is not point identified even if the researcher is willing to assume a parametric distribution for the parameter. This lack of point identification precludes the usage of firm WTP for discrete attributes in our statistical

[13] Fan and Gijbels (1996) provide aymptotically optimal methods for bandwidth choice, yet for applications with several covariates such as ours this approach may not be reliable. Relying bandwidth choice on fit quality heuristics is as reasonable here as in related work where local linear regression is also used (e.g. Bajari and Khan 2005, Bajari and Benkard 2005).

[14] Bajari and Khan (2005) provide a similar example in the context of their hedonic housing demand model where similar identification concerns arise. Thus, the lack of point identification of WTP for discrete attributes is an issue that our framework has in common with other approaches of the hedonic model literature.

work on inter-industry wage differentials. So we will focus on firm WTP on continuous attributes including education, work experience and unobserved worker quality.[15]

4 Data

The micro data in our empirical analysis come from the 1990 and 1993 waves of the National Longitudinal Survey of Youth 1979 (NLSY79). The NLSY79 is a nationally representative sample of 12,686 young men and women who were 14-22 years old when they were first surveyed in 1979. The NLSY79 data contain rich information on employment and demographic characteristics. For each individual, the NLSY79 reports age, gender, race, education, marital status, region of current residence, employment status, occupation, and earnings. In addition, the NLSY79 ask questions on individual background and employer characteristics. We have information on parental education, AFQT score, industry affiliation, and firm size.

Data on individuals' usual earnings (inclusive of tips, overtime, and bonuses but before deductions) have been collected during every survey year on the first five jobs since the last interview date in NLSY79. Combining the amount of earnings with information on the applicable unit of time, e.g. per hour, per day, per week, etc., an hourly wage rate was collected. The earnings variable used in this paper is the hourly wage for the job identified as the CPS job, i.e., current or the most recent job. We consider hourly wage less than \$1.00 and greater than \$250.00 to be outliers and eliminate them from the sample.

We construct the work experience variable from the week-by-week NLSY79 Work History Data. Usual hours worked per week at all jobs are available beginning January 1, 1978. Annual hours are computed by aggregating weekly hours in each calendar year. An individual accumulates one year of experience if she works for at least 1,000 hours a year. We restrict our sample to those with complete history of work experience. The sample we analyze contains 4,266 observations from the 1990 survey and 3,522 observations from the 1993 survey.

We use our NLSY data to estimate a standard cross-section Mincer wage equation to examine the importance of industry affiliation in explaining wage variance. Columns (1) and (5) of Table 1 report raw differences in log hourly wages by industry, for both the 1990 and 1993 observations. These are computed from cross-section regressions of log wage on a set of

[15] In an attempt to remedy for the lack of point identification of WTP for discrete attributes, we have tried to estimate the mean WTP for gender, race and marital status conditional on firm characteristics assuming a probit specification, as in Bajari and Khan (2005). We then use the estimated mean WTP in lieu of the true WTP in our statistical work on wage differentials. Not only the conditional means are statistically insignificant, but they jointly explained less than 1% of wage variation in the wage regressions discussed below. For this reason and for the sake of exposition, we focus our analysis on WTP for continuous characteristics.

industry dummy variables using one digit Census Industry Classification (CIC) Codes.[16] We use two cross-section wage observations so that we can check the consistency of our results over time and across different points in the career path. A simple summary measure of the importance of industry coefficients is their standard deviation. We report both unweighted and weighted standard deviations of the industry coefficient estimates. The unweighted standard deviation measures the difference in wages between a randomly chosen industry and the average industry, while the weighted standard deviation (by employment) measures the difference in wages between a worker in a given industry and the average worker. There is substantial variation in wages across industries.

In Columns (2) and (6) we examine the extent to which the raw inter-industry wage differentials persist once the usual human capital controls are added. Our strategy is to control for worker characteristics as well as possible, and then analyze the effects of the industry dummy variables. We estimate industry wage differentials from the cross-section wage function

$$w = X\zeta + D\tau + \varepsilon, \tag{31}$$

where w is the logarithm of the hourly wage, X is a vector of individual attributes, D is a vector of industry dummy variables, and ε is a random error term. The controls are education, age, sex, race, marital status, 4 location dummies, union status, veteran status, and several interaction terms. The industry dummy variables are statistically significant in both years,[17] substantial in magnitude, and quite similar to those estimated with other data sets (e.g. Blackburn and Neumark, 1992; Krueger and Summers, 1988) using data from the 1970s and 1980s. Earnings in construction, mining, transportation, communication and public utilities, for example, are substantially greater than those in wholesale and retail trade and service industries, even with controls for years of schooling, potential experience, gender, race, etc. The addition of human capital controls reduces inter-industry wage differentials, as measured by their standard deviation, by 8–10% in 1990 and 15–20% in 1993.

Even after including various human capital controls, the coefficient estimates on industry dummies in Equation (31) may pick up the unobserved worker quality differences across industries. Previous research has attempted to correct unobserved quality bias in estimated industry effects by including proxies of worker quality such as test scores in wage regressions (Blackburn and Neumark, 1992). In Columns (3) and (7), we include the AFQT scores as additional independent variables in the wage equations. Compared with the estimates from Columns (2) and (6), the standard deviations of the industry effects fall slightly for

[16] The service industry is used as the reference industry. Since the wage regressions include a constant, we treat the service industry as having a zero effect on wages.

[17] The only notable exception is the coefficient on the mining industry in 1993.

both the 1990 (from 0.136 to 0.133, unweighted) and the 1993 regressions (from 0.115 to 0.114, unweighted). Furthermore the inclusion of parental education in the wage regressions has little effect on the standard deviations of the industry effects, as shown in Columns (4) and (8) of Table 1. These results appear to provide no support for the unobserved quality explanation of industry wage differentials, consistent with the conclusion reached by Blackburn and Neumark (1992).

Another approach to address the problem of unobserved labor quality is by analyzing longitudinal data and by estimating first difference specification of wage equations (Gibbons and Katz, 1992; Krueger and Summers, 1988; Murphy and Topel, 1987a, 1987b). When we pool the 1990 and 1993 samples, 877 of the workers report changes in their one digit industry from 1990 to 1993. Column (9) of Table 1 reports the first difference estimates of the wage regression. The industry variables are jointly significant. For example, the first difference results show that workers who join the construction sector gain a 23.1 percent pay increase. These results are consistent with the findings by Krueger and Summers (1988), and they interpret them as evidence that differences in labor quality cannot explain inter-industry wage differentials.[18]

One potential problem with using test scores and family background as proxies to remove omitted-quality bias is that test scores and family background are only partly correlated with the types of ability that are rewarded in labor markets. The ability to do well in standard tests may be very different from the motivation and perseverance necessary to succeed in the workplace. On the other hand, the first-difference estimates rely on the assumption that unobserved quality is time invariant and equally rewarded in all industries, and therefore it can be differenced out as individual fixed effect. If labor quality evolves over time, perhaps though learning, and it is valued differently by industries, then individual fixed effect can no longer capture its effect on wages. Therefore, we cannot conclude from Table 1 that inter-industry wage differentials are not attributable to variation in unobserved labor quality.

5 Empirical Results

In this section we present estimates of our hedonic labor demand model. We first outline estimation results for the unobserved worker quality recovered in our first stage estimation. Then we present firm WTP distribution estimates based on our model specification. Fi-

[18] One notable difference between our first difference results and the previous studies (Gibbons and Katz, 1992; Krueger and Summers, 1988) is that they attempt to correct for selection bias from industry changes by using samples of displaced workers. Such sample of displace workers is not available from NLSY79. But our estimates yield similar results as those from analyzing non-displaced longitudinal data in Krueger and Summers (1988).

nally we assess how much the unobserved worker quality and firm WTP to education, work experience and quality can account for the inter-industry wage differentials.

5.1 Unobserved Worker Quality

We use NLSYdata on wages and observed worker characteristics to estimate unobserved worker quality using Equation (26). Our approach is flexible enough to allow the unobserved worker quality to evolve over time and allow firm to reward both observed and unobserved worker quality differently. The variables of observed worker characteristics, represented by the vector X, include years of schooling, experience, and dummy variables on gender, race and marital status. Out of these variables, years of schooling and experience are potentially correlated with unobserved worker quality. We include them in the sub-vector X_0 and the other observed characteristics in the sub-vector X_1.

Table 2 shows the joint distribution between some of the observed worker characteristics and the worker quality. As for the worker attributes on human capital, we find that both average worker quality increases in educational attainment, work experience and AFQT scores. Across industries, we also observe substantial differences in average worker quality, with transportation and public utilities, finance, and construction have higher average worker quality than wholesale and retail trade and services.

Table 3 reports correlations between the estimated quality and human capital variables in each year. The correlations of these variables are positive but relatively low; all six correlations are less than 0.40. In particular, the correlations between the estimated quality and AFQT score are 0.361 and 0.352 in 1990 and 1993, respectively. These estimates imply that worker quality rewarded in labor markets may not reflect in the AFQT score. Therefore explicitly incorporating AFQT scores into wage regressions cannot fully account for variations in unobserved worker quality across industries. The bottom panel of Table 3 reports the correlation between the quality estimates in 1990 and 1993 to be fairly high at 0.712. Although worker quality is highly persistent, it is by no means fixed over time according to our estimates. Thus standard first-difference estimators cannot account for the unobserved quality. Our estimates indicate that worker quality becomes more correlated with experience over time, which may be explained by the theory of learning-by-doing.

5.2 Distributions of WTP Parameters

For both years, we estimate the structural model of labor demand presented in Section 2. This yields for each firm, a WTP parameter for schooling, experience and unobserved worker quality, respectively. We present histograms of WTP parameters for these attributes for the

1990 and 1993 firms, respectively, with the estimated kernel densities. In each figure, we plot the distribution of WTP parameters for firms across all industries, followed by the distribution of the same parameters in each one-digit industry. There appears to be large variation in WTP for both observed education and experience and unobserved worker quality. All the distributions are right-skewed and are not normally distributed.

Panel A of Figure 1 presents the histogram of firm-specific WTP for one year of education in all industries in 1990. The distribution has a long right tail, with a mean of 15.4 and a standard deviation of 4.7. Panels B to H present the histograms of firm WTP for education in each one-digit industry. Finance, insurance and real estate industry has the highest mean WTP for education at 16.3, while the mining industry has the lowest mean WTP for education of 14.5. All industry-specific distributions are right-skewed. More specifically, the distribution in the services industry has the longest tail with a standard deviation of 5.0, and the distribution in the construction industry is least dispersed with a standard deviation of 4.2.

Figure 2 present the histograms of firm-specific WTP for work experience in all industries in Panel A, and in each one-digit industry in Panels B to H. The average WTP for a year of work experience (6.3) is lower than the average WTP for a year of education (15.4), but WTP for experience is more dispersed with a standard deviation of 6.1. Firms in the finance, insurance and real estate industry and the services industry value work experience most, with a mean WTP of 6.9, while experience is least valued in the construction industry with a mean WTP of 5.5. In terms of dispersion, the services industry has the longest right tail, and the distribution of WTP for experience is most concentrated in the mining industry.

Firm-specific WTP for worker quality in all industries and in each one-digit industry in 1990 are presented in Figure 3. As worker quality has no intrinsic units and is normalized between 0 and 1. The values of WTP parameters for quality are not important and we will focus on their relative level across industries. The distribution of WTP for worker quality in Panel A appears to be bimodal. This is due to the fact that most firms in some industry (e.g. mining) do not value worker quality as much whereas a majority of firms in other industry (e.g. finance, insurance and real estate) derive much higher value for worker quality. The two modes in the distributions in the construction and the services industries also contribute to the overall bimodal distribution. Based on Panels B to H, the (unobserved) worker quality is less valuable to firms in mining, construction, and wholesale and retail trade industries compared to firms in finance, insurance and real estate, services, and transportation, communication and public utilities industries. Similar to the distribution of WTP for education, the distribution of WTP for quality is most dispersed in the services industry and least dispersed in the construction industry.

Similarly we present the distributions of WTP for education, work experience, and worker quality from 1993 in Figure 4 to 6. Firms in all industries value education more in 1993 compared to 1990. The 1993 distributions of WTP for education in Figure 4 are also more dispersed than the 1990 distributions in Figure 1, and they show two modes. Likewise, Figure 6 shows that firms in all industries value worker quality more, and the distributions of WTP to quality are more dispersed in 1993. These results are consistent with the increasing return to both education and unobserved ability documented in the literature. On the contrary, work experience is less valued by firms, and firms' valuation on experience is less dispersed in 1993, as indicated by the lower mean and variance of WTP parameters in Figure 5 relative to those in Figure 2.

Firm WTP across workers' human capital attributes are not independently distributed. Table 4 reports, in each year, the correlation matrix of WTP across worker attributes on education, experience and quality. In both years, firm WTP for all human capital attributes have strong positive correlation with each other.

5.3 Inter-industry Wage Differentials

Columns (2) and (6) of Table 5 present estimates of coefficients τ in Equation (31) by adding recovered worker quality as additional control variable in the 1990 and 1993 cross-section wage regressions. For comparison, Columns (1) and (5) report the same estimates with all controls including AFQT scores and family background, but without estimated quality. The coefficient on worker quality is large and statistically significant. The magnitude of the coefficients on industry dummies declined and many of them become statistically insignificant after worker quality is included. The standard deviation of the unweighted inter-industry wage differentials decreases by 82% from 0.133 to 0.024 in 1990 and by 90% from 0.114 to 0.011 in 1993. The weighted standard deviation of wage differentials fall by a similar magnitude. In addition, the adjusted R^2 of the log wage regression increase from 0.356 to 0.861 in 1990 and from 0.376 to 0.857 in 1993. These results suggest that the unmeasured worker quality is an important driving force of inter-industry wage differentials and overall wage dispersion.

Columns (3) and (7) of Table 5 present estimates of τ coefficients in equation (31) by adding recovered firm WTP as additional control variables. The industry wage premiums in both years become smaller in size, but stay significant. The standard deviation of the unweighted inter-industry wage differentials decreases from 0.133 to 0.122 in 1990 and barely changes in 1993. The adjusted R^2 of the log wage regression increased slightly from 0.356 to 0.390 in 1990. Compared to worker quality (columns 2 and 6), firm WTP can only account for a small portion of the inter-industry wage differentials and overall wage dispersion. When both worker quality and firm WTP are included in the OLS wage regression in columns (4) and (8), the standard deviations of industry wage differentials almost stay the same as in the regressions that only control for worker quality.

We further decompose the contribution of worker heterogeneity and firm heterogeneity to observed inter-industry wage differentials by using the estimated worker quality and firm WTP. First, we estimate inter-industry wage differentials by regressing (31) with two-digit industry dummies. Then we regress the estimated industry wage premiums on recovered average worker quality at the industry level. Column (1) of Table 6 shows that unobserved worker quality alone can account for approximately two-third of the observed inter-industry wage differentials in both years. When we regress industry wage premium with firm WTP parameters in column (2), their explanatory power on industry premiums is relatively low in 1990, but higher in 1993. Combining worker quality and firm WTP in column (3), they can account for close to 80% of the overall variations in inter-industry wage premiums in both years.

6 Concluding Remarks

In this paper we propose an alternative approach to explain inter-industry wage differentials by recovering unobserved worker skill from an hedonic model of labor demand. The model allows nonparametric identification of unobserved worker skill as well as employer-specific WTP for worker attributes. Our approach does not require the usage of matched employer-employee panels to disentangle the worker effect and the firm effect in inter-industry wage differentials. Instead, we can rely on widely available household or individual micro data sets. Using data from NLSY79, we find that unmeasured worker quality accounts for most of inter-industry wage differentials.

An important caveat to the effect of firm WTP on industry wage premiums is that the hedonic labor demand model does not point identify employer-specific WTP for discrete worker characteristics, such as gender, race or marital status, even if the researcher poses strong assumptions about the distribution of WTP parameters. This is a feature that our framework shares with other related models (e.g. Bajari and Benkard 2005, Bajari and Khan 2005). We are therefore unable to identify what portion of inter-industry wage differentials can be explained by WTP for discrete attributes. Finding a set of mild assumptions that could point-identify employer WTP for discrete attributes is beyond the scope of this paper, and it is left for future research.

As in the hedonic model of differentiated product proposed by Bajari and Benkard (2005), supply-side assumptions on worker behavior are not required in our model. An interesting extension to our framework is to explicitly model labor supply behavior and allow workers to choose which firm to work for. In such an equilibrium model, one may separately identify compensating differences from WTP parameters, but it involves various challenges in identification (Ekeland, Heckman and Nesheim 2004, Heckman, Matzkin and Nesheim 2010).

Appendix A:

Proof of Proposition 1

To show Proposition 1, it suffices to demonstrate that for any two workers j and j' employed in market t, three conditions hold:

(1) If $X_{jt} = X_{j't}$ and $\xi_{jt} = \xi_{j't}$ then $w_{jt} = w_{j't}$.

(2) If $X_{jt} = X_{j't}$ and $\xi_{jt} > \xi_{j't}$ then $w_{jt} > w_{j't}$.

(3) $|w_{jt} - w_{j't}| \leq M(|X_{jt} - X_{j't}| + |\xi_{jt} - \xi_{j't}|)$ for some $M < \infty$.

Suppose $w_{jt} > w_{j't}$ for some market t in which both workers j and j' are employed and $X_{jt} = X_{j't}$ and $\xi_{jt} = \xi_{j't}$. Then $R_{it}(X_{jt}, \xi_{jt}) - w_{jt} < R_{it}(X_{j't}, \xi_{j't}) - w_{j't}$ for all employers $i = 1, ..., V_t$. This implies that no one would hire worker j in market t, which is a contradiction.

Suppose $w_{jt} \leq w_{j't}$ for some market t in which both workers j and j' are employed and $X_{jt} = X_{j't}$ and $\xi_{jt} > \xi_{j't}$. Since $R_{it}(X_{jt}, \xi_{jt})$ is strictly increasing in ξ_{jt}, it follows that $R_{it}(X_{jt}, \xi_{jt}) - w_{jt} > R_{it}(X_{j't}, \xi_{j't}) - w_{j't}$ for all employers $i = 1, ..., V_t$. This implies that no one would hire worker j' in market t, which is a contradiction.

The assumption that $R_{it}(X_{jt}, \xi_{jt})$ is Lipschitz continuous in (X_{jt}, ξ_{jt}) implies that, for any two workers j and j' differing in at least one characteristic,

$$|R_{it}(X_{jt}, \xi_{jt}) - R_{it}(X_{j't}, \xi_{j't})| \leq M(|X_{jt} - X_{j't}| + |\xi_{jt} - \xi_{j't}|),$$

for some $M < \infty$. As $|R_{it}(X_{jt}, \xi_{jt}) - R_{it}(X_{j't}, \xi_{j't})| = |[(R_{it}(X_{jt}, \xi_{jt}) - w_{jt}) - (R_{it}(X_{j't}, \xi_{j't}) - w_{j't})] + (w_{jt} - w_{j't})|$, we have

$$\begin{aligned} & |[(R_{it}(X_{jt}, \xi_{jt}) - w_{jt}) - (R_{it}(X_{j't}, \xi_{j't}) - w_{j't})] + (w_{jt} - w_{j't})| \\ \leq \; & M(|X_{jt} - X_{j't}| + |\xi_{jt} - \xi_{j't}|). \end{aligned}$$

Assume without loss of generality that $w_{jt} > w_{j't}$, then the second term on the right hand side, $w_{jt} - w_{j't}$, is positive. Since the demand for worker j is positive, the first term must be positive for some employer i. For these employers, we can ignore the absolute sign.

$$\begin{aligned} & |[(R_{it}(X_{jt}, \xi_{jt}) - w_{jt}) - (R_{it}(X_{j't}, \xi_{j't}) - w_{j't})] + (w_{jt} - w_{j't})| \\ = \; & [(R_{it}(X_{jt}, \xi_{jt}) - w_{jt}) - (R_{it}(X_{j't}, \xi_{j't}) - w_{j't})] + (w_{jt} - w_{j't}) > w_{jt} - w_{j't}. \end{aligned}$$

Therefore,

$$w_{jt} - w_{j't} \leq M(|X_{jt} - X_{j't}| + |\xi_{jt} - \xi_{j't}|) \text{ for employer } i \text{ that prefers } j \text{ to } j'.$$

Here we use the fact that both workers have positive demand to limit how much their wages can vary.

Proof of Proposition 2:

First, we use the assumption that each function $h_m(., \eta_m)$ is strictly monotonic in η_m to define $h_m^{-1}(x_{0,m}, x_1, z)$ as the inverse of $h_m(x_1, z, \eta_m)$. Then, following the proof of Lemma 1 of Matzkin (2003), for each $m = 1, ..., M_0$ we have

$$
\begin{aligned}
F_{X_{0,m}|X_1,Z}(x_{0,m}|x_1, z) &= \Pr(X_{0,m} \leq x_{0,m}|X_1 = x_1, Z = z) &(32)\\
&= \Pr(h_m(x_1, z, \eta_m) \leq x_{0,m}|X_1 = x_1, Z = z) &(33)\\
&= \Pr(\eta_m \leq h_m^{-1}(x_{0,m}, x_1, z)|X_1 = x_1, Z = z) &(34)\\
&= \Pr(\eta_m \leq h_m^{-1}(x_{0,m}, x_1, z)) &(35)\\
&= F_{\eta_m}(h_m^{-1}(x_{0,m}, x_1, z)) = h_m^{-1}(x_{0,m}, x_1, z) = \eta_m, &(36)
\end{aligned}
$$

where (34) follows from the monotonicity assumption, (35) follows from the independence between (X_1, Z) and $\eta = (\eta_1, ..., \eta_{M_0})$, and (36) is a result of normalizing η_m so that it follows an $U[0, 1]$).

We now show that the vector η consists of control variables such that X and ξ are independent conditional on η. Adapting the proof of Theorem 1 of Imbens and Newey (2009) for multivariate X_0, for any bounded function $a(x_0, x_1)$, it follows from the independence of (X_1, Z) and (ξ, η) that

$$
\begin{aligned}
E[a(x_0, x_1)|\xi, \boldsymbol{\eta}] &= E[a(h_1(x_1, z, \eta_1), ..., h_{M_0}(x_1, z, \eta_{M_0}), x_1)|\xi, \boldsymbol{\eta}]\\
&= \int a(h_1(x_1, z, \eta_1), ..., h_{M_0}(x_1, z, \eta_{M_0}), x_1)dF_{X_1,Z}(x_1, z)\\
&= E[a(x_0, x_1)|\boldsymbol{\eta}].
\end{aligned}
$$

Thus, for any bounded function $b(\xi)$, it follows from law of iterated expectations that

$$
\begin{aligned}
E[a(x_0, x_1)b(\xi)|\boldsymbol{\eta}] &= E[b(\xi)E[a(x_0, x_1)|\xi, \boldsymbol{\eta}]|\boldsymbol{\eta}]\\
&= E[b(\xi)E[a(x_0, x_1)|\boldsymbol{\eta}]|\boldsymbol{\eta}]\\
&= E[b(\xi)|\boldsymbol{\eta}]E[a(x_0, x_1)|\boldsymbol{\eta}],
\end{aligned}
$$

which means by definition independence between X and ξ conditional on η.

Finally, given that both ξ each η_m are normalized such that each follow a $U[0, 1]$, for each market t and worker j we have

$$
\begin{aligned}
\int F_{w|x,\boldsymbol{\eta}}(w_{jt}|X_{jt},\boldsymbol{\eta})d\boldsymbol{\eta} &= \int \Pr(\mathbf{w}_t(X,\xi) \leq w_{jt}|X = X_{jt},\boldsymbol{\eta})d\boldsymbol{\eta} \\
&= \int \Pr(\xi \leq \mathbf{w}_t^{-1}(X,w_{jt})|X = X_{jt},\boldsymbol{\eta})d\boldsymbol{\eta} \\
&= \int \Pr(\xi \leq \mathbf{w}_t^{-1}(X_{jt},w_{jt})|\boldsymbol{\eta})d\boldsymbol{\eta} \\
&= \int \Pr(\xi \leq \xi_{jt}|\boldsymbol{\eta})d\boldsymbol{\eta} = F_{\xi}(\xi_{jt}) = \xi_{jt},
\end{aligned}
$$

where we exploited the fact that the equilibrium wage function is strictly monotonic in ξ.

Appendix B

Construction of one-digit industry aggregates from 1970 Census Industry Classification Codes

One-digit industry	1970 Census Industry Classification Codes
Mining	47-57
Construction	67-77
Manufacturing	107-398
Transportation, Communication, Public Utilities	407-479
Wholesale and Retail Trade	507-698
Finance, Insurance and Real Estate	707-718
Services	727-897

References

[1] Abowd, J., and F. Kramarz (2000), "Inter-industry and Firm-size Wage Differentials: New Evidence from Linked Employer-Employee Data," Unpublished manuscript, Cornell University.

[2] Abowd, J., F. Kramarz, and D. Margolis (1999), "High Wage Workers and High Wage Firms," *Econometrica* 67: 251–334.

[3] Bajari, P., and C.L. Benkard (2005), "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach," *Journal of Political Economy* 113: 1239–1276.

[4] Bajari, P., J. Fruehwirth, K. Kim and C. Timmins (2012), "A Rational Expectations Approach to Hedonic Price Regressions with Time-Varying Unobserved Product Attributes: The Price of Pollution," *American Economic Review* 102: 1898–1926.

[5] Bajari, P., J. Fox and S. Ryan (2007), "Linear Regression Estimation of Discrete Choice Models with Nonparametric Distributions of Random Coefficients," *American Economic Review Papers and Proceedings* 97: 459–463.

[6] Bajari, P., J. Fox, K. Kim and S. Ryan (2011), "A Simple Nonparametric Estimator for the Distribution of Random Coefficients," *Quantitative Economics* 2: 381–418.

[7] Bajari, P., and M.E. Kahn (2005), "Estimating Housing Demand With an Application to Explaining Racial Segregation in Cities," *Journal of Business and Economic Statistics* 23: 20–33.

[8] Bartik, T.J. (1987), "The Estimation of Demand Parameters in Hedonic Price Models," *Journal of Political Economy* 95: 81–88.

[9] Berry, S. (1994), "Estimating Discrete Choice Models of Product Differentiation," *Rand Journal of Economics* 25: 242–262.

[10] Berry, S. J. Levinsohn and A. Pakes (1995), "Automobile Prices in Market Equilibrium," *Econometrica* 63: 841–890.

[11] Bhat, C. (2001), "Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model," *Transportation Research Part B: Methodological* 35: 677–693.

[12] Blundell, R. and J. Powell (2003), "Endogeneity in Nonparametric and Semiparametric Regression Models," in *Advances in Economics and Econometrics*, Vol. II, ed. by Dewatripont, L. Hansen and S. Turnovsky. Cambridge: Cambridge University Press, pp. 312-357.

[13] Blundell, R. and J. Powell (2004), "Endogeneity in Semiparametric Binary Response Models," *Review of Economic Studies* 71: 581–913.

[14] Brown, J.N. and H.S. Rosen (1982), "On the Estimation of Structural Hedonic Price Models," *Econometrica* 50: 765–768.

[15] Chernozhukov, V. and C. Hansen (2005), "An IV Model of Quantile Treatment Effects," *Econometrica* 73: 245–261.

[16] Chernozhukov, V., G. Imbens and W. Newey (2007), "Instrumental Variable Estimation of Nonseparable Models," *Journal of Econometrics* 139: 4–14.

[17] Chesher, A. (2003), "Identification in Nonseparable Models," *Econometrica* 71: 1405–1441.

[18] Ekeland, I., J. Heckman and L. Nesheim (2004), "Identification and Estimation of Hedonic Models," *Journal of Political Economy* 112: S60–S109.

[19] Epple, D. (1987), "Hedonic Prices and Implicit Markets: Estimating Supply and Demand Functions for Differentiated Products," *Journal of Political Economy* 95: 59–80.

[20] Fan, J. and I. Gijbels (1996), *Local Polynomial Modelling and its Applications: Monographs on Statistics and Applied Probability*. London, Chapman & Hall.

[21] Farre, L. R. Klein and F. Vella (2010), "A Parametric Control Function Approach to Estimating the Returns to Schooling in the Absence of Exclusion Restrictions: an Application to the NSLY," *Empirical Economics*, forthcoming.

[22] Fox, J. (2009), "Firm-Size Wage Gaps, Job Responsibility, and Hierarchical Matching," *Journal of Labor Economics* 27: 83–126.

[23] Fox, J. and V. Smeets (2011), "Does Input Quality Drive Measured Differences in Firm Productivity?" *International Economic Review* 52: 961–989.

[24] Gibbons, R. and L. Katz (1992), "Does Unmeasured Ability Explain Inter-Industry Wage Differentials," *Review of Economic Studies* 59: 515–535.

[25] Halton, J. (1960), "On the Efficiency of Evaluating Certain Quasi-Random Sequences of Points in Evaluating Multi-dimensional Integrals," *Numerische Mathematik* 2: 84–90.

[26] Heckman, J., R. Matzkin and L. Nesheim (2010), "Nonparametric Identification and Estimation of Nonadditive Hedonic Models," *Econometrica* 78: 1569–1591.

[27] Hwang, H., W.R. Reed, and C. Hubbard (1992), "Compensating Wage Differentials and Unobserved Productivity," *Journal of Political Economy* 100: 835–858.

[28] Imbens, G. and W. Newey (2009), "Identification and Estimation of Triangular Simultaneous Equations Models without Additivity," *Econometrica* 77: 1481–1512.

[29] Judd, K. (1998), *Numerical Methods in Economics*. MIT Press, Cambridge MA.

[30] Katz, L. (1986), "Efficiency Wage Theories: A Partial Evaluation," *NBER Macroeconomics Annual* 1: 235–276.

[31] Katz, L. and D. Autor (1999), "Changes in the Wage Structure and Earnings Inequality," in *Handbook of Labor Economics*, O. Ashenfelter and D. Card, edt. vol. 3, Part A, 1463–1555. Elsevier.

[32] Katz, L. and L. Summers (1989), "Industry Rents: Evidence and Implications," *Brookings Papers on Economic Activity: Microeconomics*, 209–275.

[33] Krueger, A.B., and L.H. Summers (1988), "Efficiency Wages and the Inter-Industry Wage Structure," *Econometrica* 56: 259–294.

[34] Mardia, J., J. Kent and M. Bibby (1979), *Multivariate Analysis*. Academic Press, New York.

[35] Matzkin, R. (2003), "Nonparametric Estimation of Nonadditive Random Functions," *Econometrica* 71: 1332–1375.

[36] Murphy, K.M., and R. Topel (1987), "Efficiency Wages Reconsidered: Theory and Evidence," in *Advances in the Theory and Measurement of Unemployment,* Y. Weiss and G. Fishelson eds. New York: St. Martin's Press.

[37] Nickell, S. and S. Wadhwani (1990), "Insider Forces and Wage Determination," *Economic Journal* 100: 496–509.

[38] Norets, A. (2010), "Approximation of Conditional Densities by Smooth Mixtures of Regressions," *Annals of Statistics* 38: 1733–1766.

[39] Petrin, A. (2002), "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy* 110: 705–729.

[40] Petrin, A. and K. Train (2010), "A Control Function Approach to Endogeneity in Consumer Choice Models," *Journal of Marketing Research* 47: 3–13.

[41] Rosen, S. (1974), "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy* 82: 34–55.

[42] Rosen, S. (1986), "The Theory of Equalizing Differences," in *Handbook of Labor Economics*, vol. 1, edited by Orley Ashenfelter and Richard Layard. Amsterdam: North-Holland.

[43] Sattinger, M. (1980), *Capital and the Distribution of Labor Earnings*. North-Holland.

[44] Thaler, R.H. (1989), "Anomalies: Interindustry Wage Differentials," *Journal of Economic Perspectives* 3: 181–193.

[45] Willis, R. and S. Rosen (1979), "Education and Self-Selection," *Journal of Political Economy* 87: S7–S36.

Table 1: Estimated Wage Differentials for One-Digit Industries, NLSY79[a]
(Standard Errors in Parentheses)

	Cross-section, 1990				Cross-section, 1993				Fixed Effects
Industry	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Mining	0.211	0.287	0.275	0.276	0.112	0.126	0.118	0.120	0.159
	(0.097)	(0.082)	(0.081)	(0.081)	(0.119)	(0.098)	(0.097)	(0.097)	(0.211)
Construction	0.215	0.273	0.266	0.265	0.153	0.216	0.214	0.212	0.231
	(0.032)	(0.028)	(0.028)	(0.028)	(0.037)	(0.033)	(0.033)	(0.033)	(0.064)
Manufacturing	0.101	0.160	0.158	0.159	0.103	0.139	0.138	0.139	0.161
	(0.022)	(0.019)	(0.019)	(0.019)	(0.026)	(0.022)	(0.022)	(0.022)	(0.051)
Transportation, Communication,	0.208	0.178	0.174	0.172	0.224	0.168	0.164	0.163	0.065
Public Utilities	(0.033)	(0.028)	(0.028)	(0.028)	(0.038)	(0.032)	(0.032)	(0.032)	(0.065)
Wholesale and Retail Trade	-0.159	-0.083	-0.084	-0.086	-0.178	-0.118	-0.118	-0.120	-0.036
	(0.022)	(0.019)	(0.019)	(0.019)	(0.026)	(0.022)	(0.022)	(0.022)	(0.048)
Finance, Insurance and Real Estate	0.228	0.173	0.166	0.166	0.233	0.142	0.143	0.142	0.133
	(0.034)	(0.029)	(0.029)	(0.029)	(0.038)	(0.033)	(0.033)	(0.033)	(0.071)
Other Control variables[b]	No	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
AFQT	No	No	Yes	Yes	No	No	Yes	Yes	No
Parental Education	No	No	No	Yes	No	No	No	Yes	No
Unweighted St.d. of Differentials	0.147	0.136	0.133	0.133	0.143	0.115	0.114	0.114	0.096
Weighted St.d. of Differentials	0.052	0.047	0.046	0.046	0.052	0.044	0.043	0.043	0.038
Adjusted R squared	0.056	0.345	0.355	0.356	0.048	0.370	0.376	0.376	0.044

[a] The dependent variable is log (hourly wage). The reported estimates are the coefficient values for the industry dummy variables. The reference industry is Services.

[b] Other control variables are education, years of experience and its square, gender dummy, race dummy, ever married dummy, union and veteran status, 4 region dummies, 3 occupation dummies, marriage x gender interaction, education x gender interaction, education squared x gender interaction, age x gender interaction, and a constant.

Table 2: Conditional Worker Quality Distribution

Normalized Worker Quality	Cross-section, 1990 Mean	(Std.)	Cross-section, 1993 Mean	(Std.)
All workers	0.507	(0.274)	0.495	(0.272)
By education				
High school incomplete	0.403	(0.249)	0.382	(0.240)
High school graduates	0.448	(0.262)	0.433	(0.260)
Some college	0.533	(0.267)	0.511	(0.269)
College graduates	0.674	(0.241)	0.654	(0.238)
By work experience				
0-4 years	0.414	(0.272)	0.343	(0.253)
5-9 years	0.534	(0.275)	0.466	(0.280)
10-15 years	0.547	(0.248)	0.551	(0.253)
By AFQT percentile scores				
1-25	0.403	(0.248)	0.394	(0.245)
26-50	0.507	(0.267)	0.499	(0.265)
51-75	0.580	(0.265)	0.560	(0.267)
76-100	0.660	(0.256)	0.644	(0.253)
By industry				
Mining	0.583	(0.264)	0.530	(0.229)
Construction	0.581	(0.268)	0.556	(0.265)
Manufacturing	0.529	(0.256)	0.522	(0.262)
Transportation, Communication, Public Utilities	0.605	(0.263)	0.596	(0.265)
Wholesale and Retail Trade	0.397	(0.250)	0.389	(0.249)
Finance, Insurance and Real Estate	0.608	(0.252)	0.588	(0.251)
Services	0.503	(0.283)	0.482	(0.276)

Table 3. Correlations of Estimated Quality and Observed Human Capital Variables

		1990 cross-section	
	Education	Experience	AFQT
Estimated quality	0.338	0.196	0.361
		1993 cross-section	
	Education	Experience	AFQT
Estimated quality	0.348	0.251	0.352
		1990 and 1993 pooled	
		Estimated quality in 1993	
Estimated quality in 1990		0.712	

Table 4. Correlation Matrix of WTP Parameters by Year

	Education	Experience
1990		
Experience	0.894	
Quality	0.925	0.774
1993		
Experience	0.978	
Quality	0.958	0.965

Table 6. Decomposition of Inter-Industry Wage Differentials

	(1)	(2)	(3)
	1990 two-digit industry premiums		
Quality	1.018		1.543
	(0.114)		(0.139)
Firm WTP	No	Yes	Yes
R squared	0.664	0.235	0.824
Adjusted R squared	0.656	0.174	0.805
	1993 two-digit industry premiums		
Quality	1.000		1.369
	(0.104)		(0.157)
Firm WTP	No	Yes	Yes
R squared	0.697	0.423	0.811
Adjusted R squared	0.689	0.377	0.790

Table 5. Estimated Wage Differentials for One-Digit Industries with Quality and WTP Estimates, NLSY79[a]
(Standard Errors in Parentheses)

	1990 Cross-section				1993 Cross-section			
Industry	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mining	0.276	0.056	0.232	0.050	0.120	0.006	0.119	0.006
	(0.081)	(0.038)	(0.079)	(0.037)	(0.097)	(0.047)	(0.097)	(0.046)
Construction	0.265	0.060	0.247	0.059	0.212	0.016	0.212	0.012
	(0.028)	(0.013)	(0.028)	(0.013)	(0.033)	(0.016)	(0.033)	(0.016)
Manufacturing	0.159	0.042	0.147	0.042	0.139	0.017	0.139	0.019
	(0.019)	(0.009)	(0.019)	(0.009)	(0.022)	(0.011)	(0.022)	(0.010)
Transportation, Communication,	0.172	0.008	0.155	0.009	0.163	-0.008	0.163	-0.007
Public Utilities	(0.028)	(0.013)	(0.027)	(0.013)	(0.032)	(0.015)	(0.032)	(0.015)
Wholesale and Retail Trade	-0.086	0.016	-0.087	0.017	-0.120	-0.011	-0.118	-0.006
	(0.019)	(0.009)	(0.018)	(0.009)	(0.022)	(0.011)	(0.022)	(0.011)
Finance, Insurance and Real Estate	0.166	0.026	0.153	0.026	0.142	0.011	0.142	0.013
	(0.029)	(0.013)	(0.028)	(0.013)	(0.033)	(0.016)	(0.033)	(0.016)
Worker Quality	No	Yes	No	Yes	No	Yes	No	Yes
Firm's Willingness to Pay	No	No	Yes	Yes	No	No	Yes	Yes
Unweighted St.d. of Differentials	0.133	0.024	0.122	0.022	0.114	0.011	0.114	0.010
Weighted St.d. of Differentials	0.046	0.008	0.043	0.008	0.043	0.004	0.043	0.004
Adjusted R squared	0.356	0.861	0.390	0.865	0.376	0.857	0.376	0.860

[a] The dependent variable is log (hourly wage). The reported estimates are the coefficient values for the industry dummy variables. The reference industry is Services. Other control variables are education, years of experience and its square, gender dummy, race dummy, ever married dummy, union and veteran status, 4 region dummies, 3 occupation dummies, marriage x gender interaction, education x gender interaction, education squared x gender interaction, age x gender interaction, mother's schooling, father's schooling, AFQT test score, and a constant.

Figure 1: Firm WTP for Education Across Industries, 1990

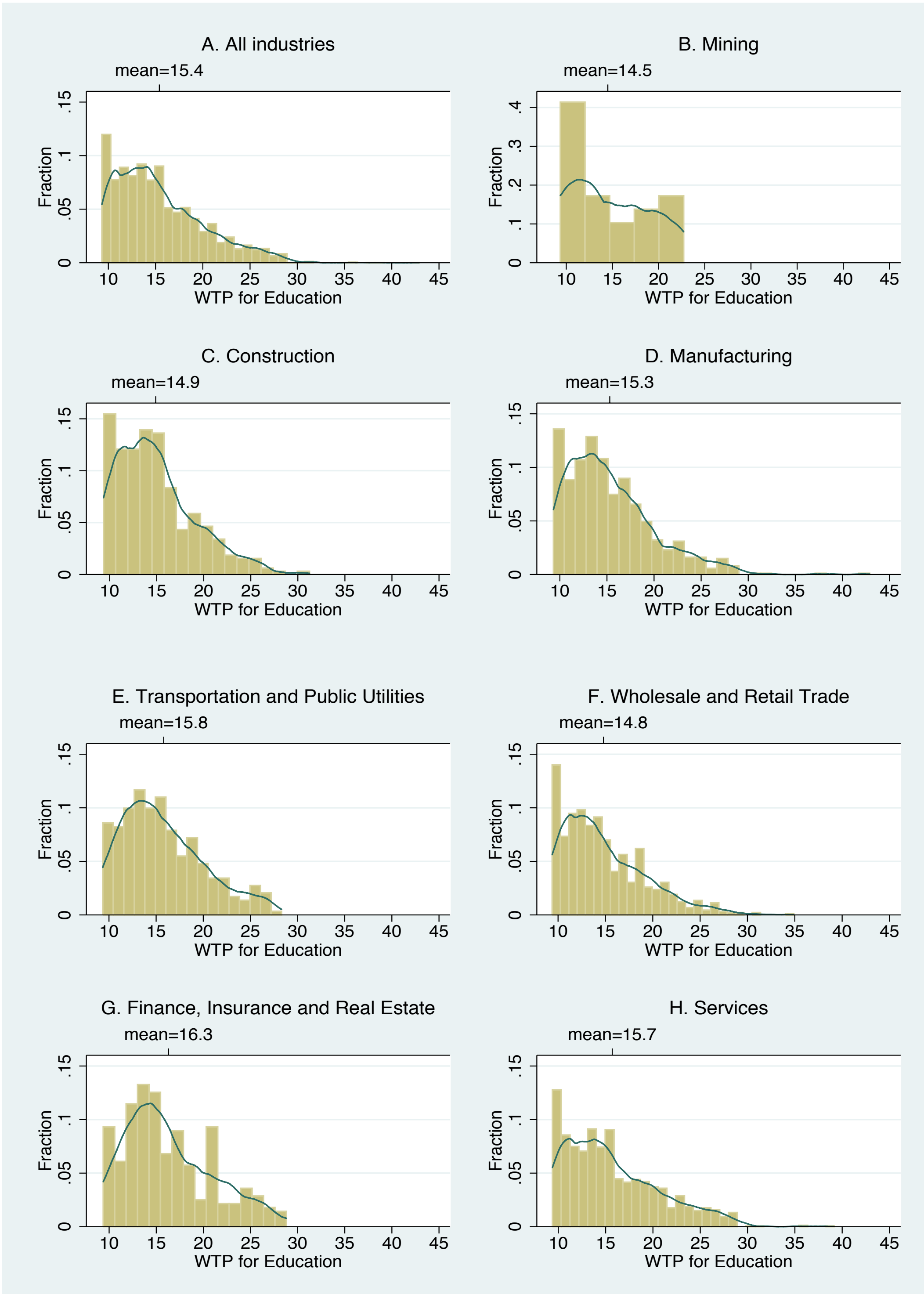

Figure 2: Firm WTP for Work Experience Across Industries, 1990

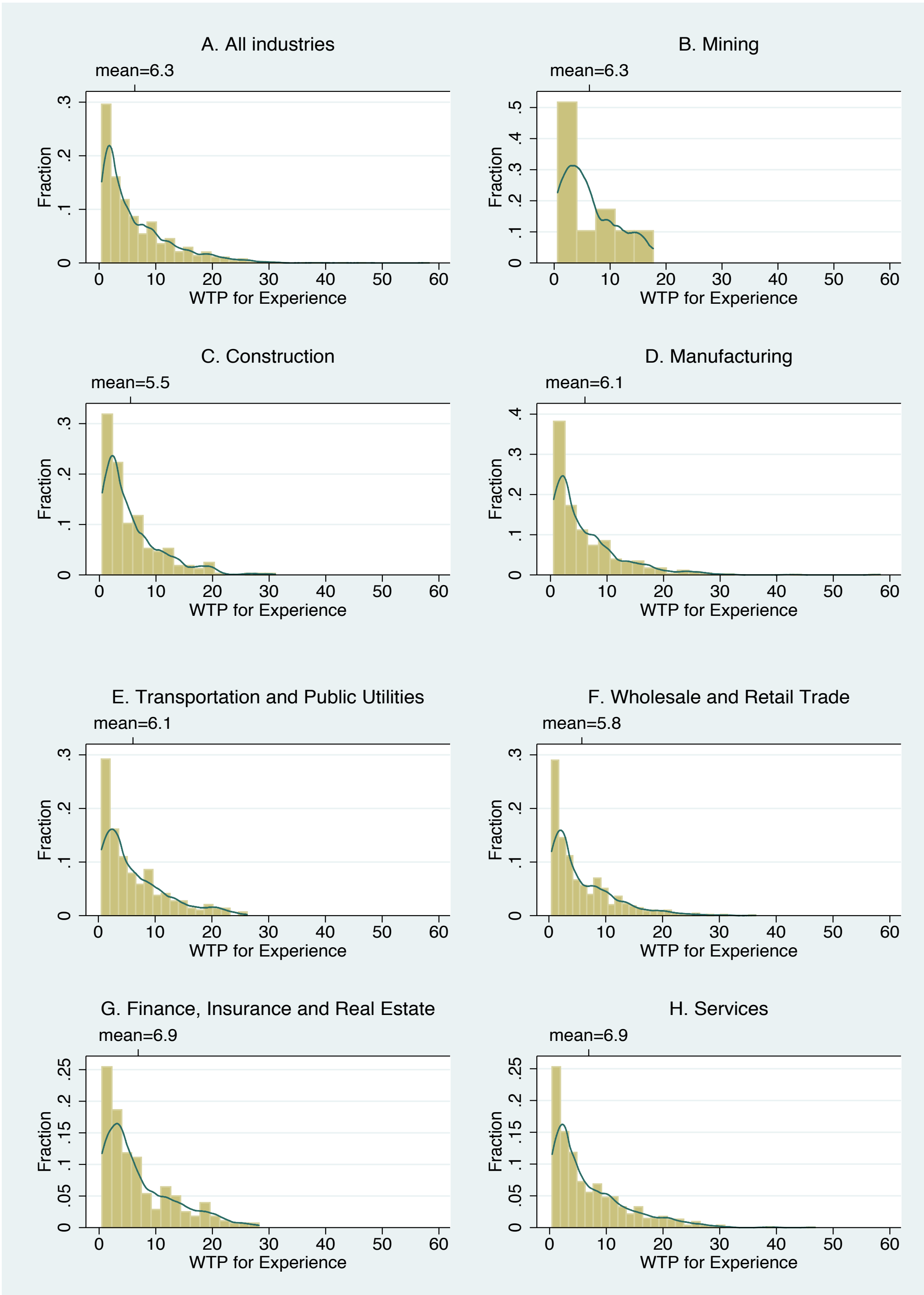

Figure 3: Firm WTP for Worker Quality Across Industries, 1990

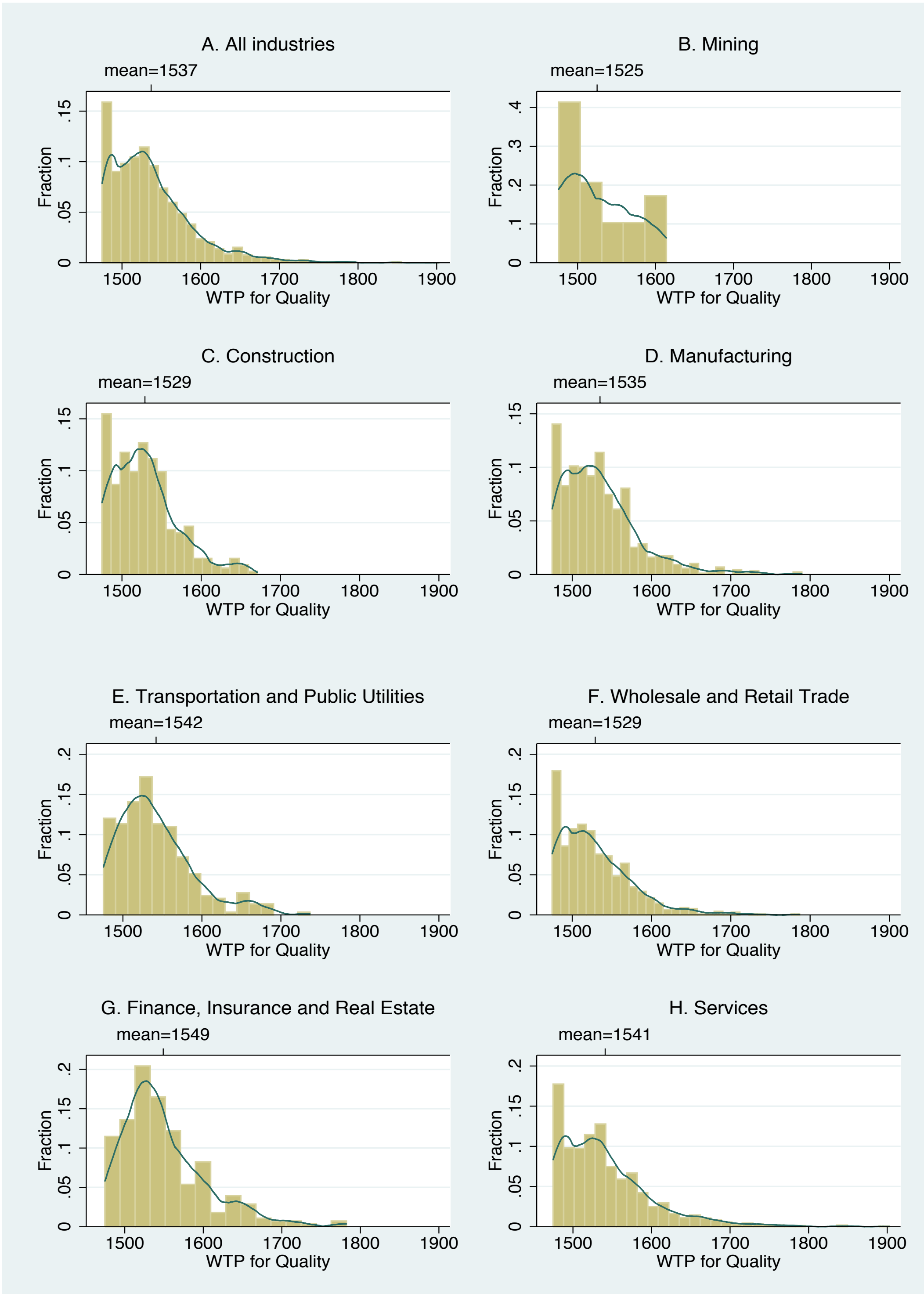

Figure 4: Firm WTP for Education Across Industries, 1993

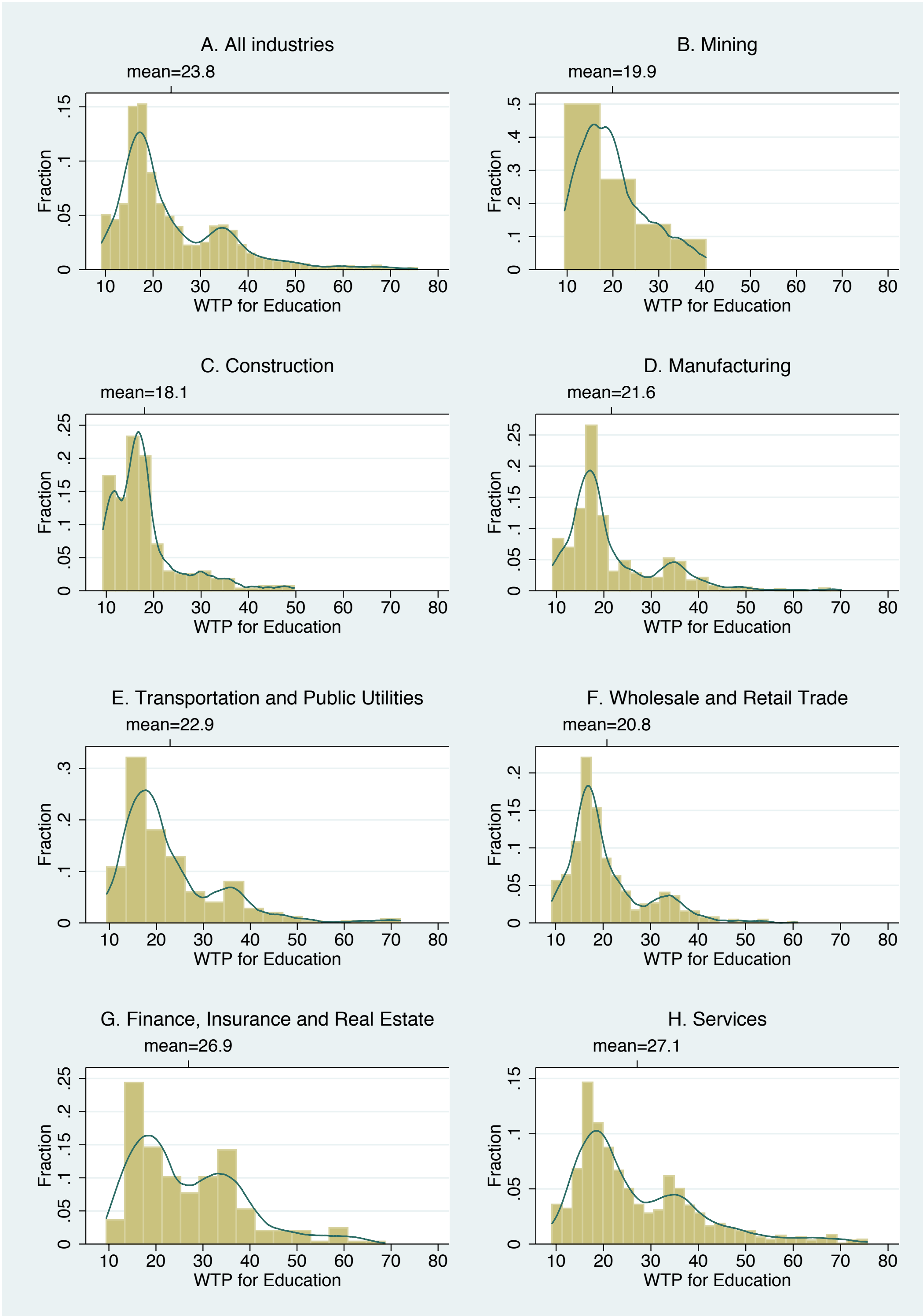

Figure 5: Firm WTP for Work Experience Across Industries, 1993

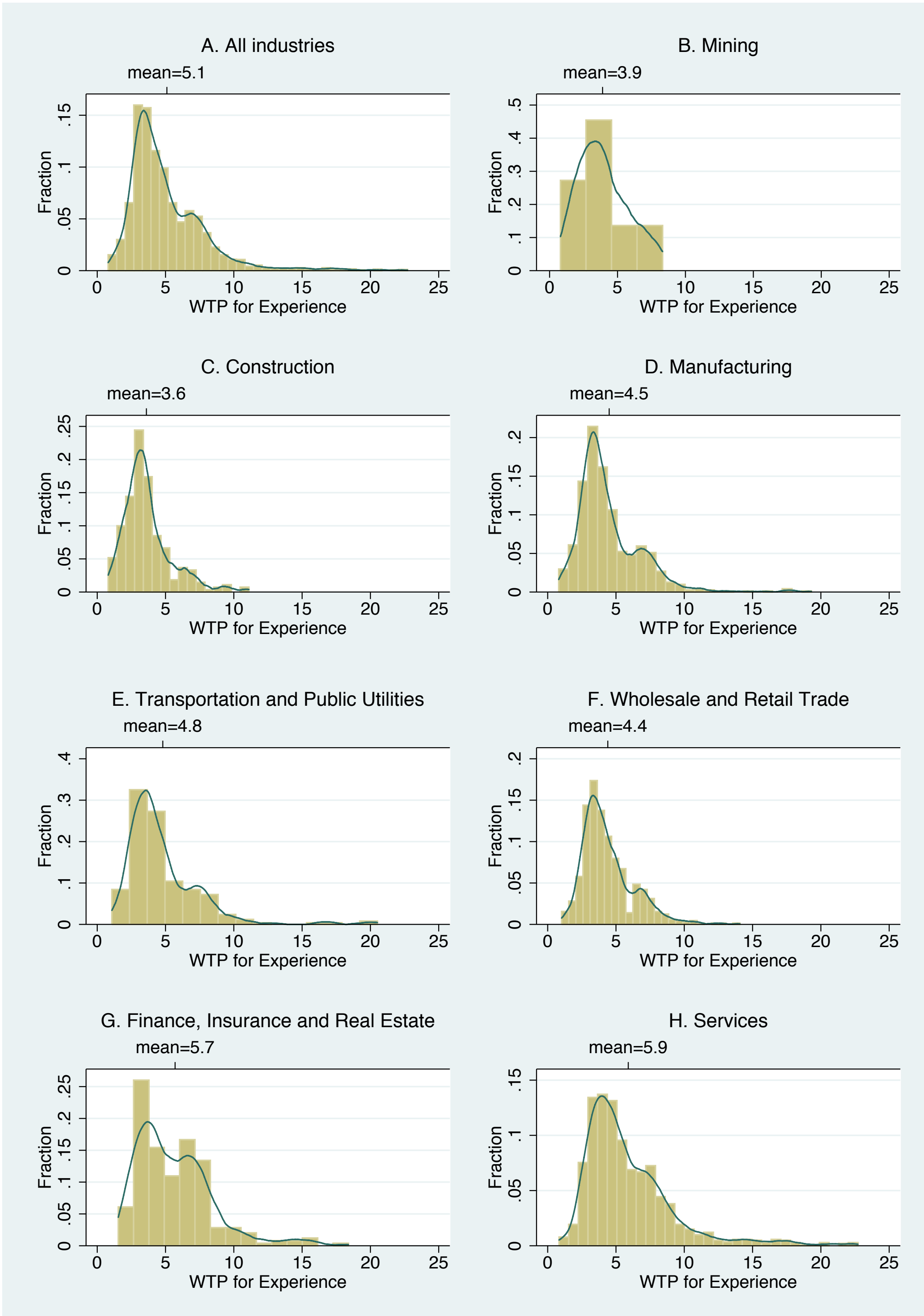

Figure 6: Firm WTP for Worker Quality Across Industries, 1993