Unobserved Worker Quality and Inter-Industry Wage Differentials†

Suqin Ge
Virginia Tech

João Macieira
Virginia Tech

February 15, 2014

Abstract

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

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

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, the Latin American Meetings of the Econometric Society in Santiago, and the AEA meetings in Philadelphia 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 because these differentials 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 worker-specific productive abilities not measured in data (Murphy and Topel, 1987a, 1987b). 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) empirically assess both 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.

Debate persists over how much observed inter-industry wage differentials can be explained by unobserved worker or firm characteristics. To disentangle simultaneous worker- and firm-level heterogeneity in wage determination, microdata that match the characteristics of firms to those of their workers are preferred. Several studies (Abowd, Kramarz and Margolis, 1999; Abowd et al., 2005) have decomposed inter-industry wage differences into a worker fixed effect and a firm fixed effect by using extensive matched employer–employee panel data.[2] However, such matched employer–employee panels are rarely accessible to researchers. Moreover, the decomposition of inter-industry wage differences by using a worker fixed effect assumes unobserved worker characteristics to be time-invariant, but this assumption may not hold in practice. For example, if labor quality evolves over time as a result of learning-

[1] Thaler (1989) reviews the debate on whether 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 wage structure.

[2] In a related paper, Fox and Smeets (2011) use matched employer–employee panel data from Denmark to explain productivity dispersion across firms.

by-doing, a worker fixed effect cannot fully capture the effects of unmeasured quality on wages.

In this study, 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 nonparametrically recover unobserved worker quality by using an estimator based on Bajari and Benkard (2005), Imbens and Newey (2009), and Norets (2010). This estimator exploits both the uniqueness of the 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 nonparametrically infer firm-specific willingness to pay (WTP) with respect to both observed and unobserved worker attributes by using model results relating WTP and first-order conditions for profit maximization. Once unobserved worker and firm effects are identified, we can quantitatively assess their importance in accounting for inter-industry wage differentials on the basis of 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 estimating demands for differentiated products in industrial organization.[3] We model labor demand as a discrete choice of a set of worker attributes. Worker quality is modeled as a worker attribute unobserved by econometricians but valued by employers. Literature on industrial organization has proposed nonparametric methods to identify product characteristics observed by consumers but not by researchers (e.g., Bajari and Benkard, 2005). We apply these methods to recover worker quality. As in hedonic literature, the marginal prices of worker characteristics are estimated as random coefficients in a hedonic

[3] Most of hedonic literature considers a market with a continuum of products and perfect competition and assumes all product characteristics to be perfectly observed. Rosen's estimation strategy is criticized by Brown and Rosen (1982), Epple (1987), and Bartik (1987), who argue that preference estimates are biased because consumers who strongly prefer a product characteristic purchase more of that characteristic. Bajari and Benkard (2005) relax these assumptions and propose a hedonic model of demand for differentiated products; this model accounts for unobserved product characteristics and heterogeneous consumers. Ekeland, Heckman, and Nesheim (2004) and Heckman, Matzkin, and Nesheim (2010) thoroughly discuss identification issues in estimating hedonic models.

wage function.

Our labor demand model is estimated on the basis of individual data from NLSY79 to explore the importance of worker and firm effects in wage determination. We estimate the model separately for two different years and identify unobserved worker quality and firm WTP for productive characteristics in each given year. Our estimates show that the worker effect captured by unobserved worker quality is statistically more important in explaining wages than the firm effect measured by firm WTP. Unmeasured worker quality accounts for approximately two thirds of inter-industry wage differentials. Although worker quality is persistent, it evolves over time and cannot be captured by a worker fixed effect alone.[4]

Observed worker characteristics that are supposed to account for productivity differences typically explain no more than 30 to 40 percent of wage variations across workers. Considerable residual variance suggests differences in unmeasured worker ability: highly skilled workers earn high wages. Our empirical analysis reveals that the percentage of explained wage differentials across workers nearly doubles when log wage regressions on observed worker attributes are augmented by estimated unobserved worker quality.

This paper is organized as follows. In section 2 we present a hedonic model of labor demand 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 the estimation results. Section 6 concludes and outlines possible extensions for future research. All derivations and auxiliary results can be found in the appendices.

[4]Using matched employer–employee panel data from France, Abowd, Kramarz, and Margolis (1999) also find that worker effects are more important than firm effects in explaining inter-industry wage differentials. However, those authors assume both work and firm effects to be fixed over time, whereas we allow them to vary.

2 A Model of Labor Demand

This section describes a labor demand model for heterogeneous workers. Consider an economy in which 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. Each market has $j = 1, ..., J_t$ 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 set of characteristics that potential employers value differently. M 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 unobserved worker characteristics valued by the employer but unobserved by the researcher, such as productive abilities, communication skills, and career ambition. For simplicity, we interpret the variable ξ_{jt} as representing unmeasured worker quality rewarded in labor markets.

The output of 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 unit 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 the efficiency unit.[5] 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

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

[5] 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, a 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) simplifies the employer's problem and yields 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 cost; that is

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

We model a worker's labor efficiency units as a function of his or her characteristics such that $E_{jt} = E(X_{jt}, \xi_{jt})$. The employer's decision then 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 that the employer randomly picks only one worker to fill the vacancy. The option of not hiring is denoted by $j = 0$.

In the proposed heterogeneous labor demand model, a unique equilibrium wage function $\mathbf{w}_t(X_{jt}, \xi_{jt})$ in each market t maps the set of worker characteristics onto the set of wages under weak assumptions. The equilibrium wages have the following properties: (1) there is one wage for each set of worker characteristics, and (2) the wage function increases in 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.[6]

The wage function in each market is an equilibrium function dependent on market primitives, such as the production function. The wage function is not additively separable a priori because we have limited information about its form. The proposition is based on demand-side arguments only; thus, our estimation strategy is robust to various supply-side assumptions.[7]

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

$$\frac{\partial R_{it}(X_{j^*t}, \xi_{j^*t})}{\partial x^c_{j,m,t}} = \frac{\partial \mathbf{w}_t(X_{j^*t}, \xi_{j^*t})}{\partial x^c_{j,m,t}}, \tag{6}$$

$$\frac{\partial R_{it}(X_{j^*t}, \xi_{j^*t})}{\partial \xi_{jt}} = \frac{\partial \mathbf{w}_t(X_{j^*t}, \xi_{j^*t})}{\partial \xi_{jt}}. \tag{7}$$

Thus, with a firm's optimal labor demand, the value the firm derives from the last unit of each worker characteristic is equal to the implicit price it had to pay for that unit. Otherwise, the firm can increase its profits by employing an alternative worker with a different set of worker attributes.

[6]We follow a similar strategy taken by Bajari and Benkard (2005) in their model of the demand for differentiated products.

[7]An equilibrium wage function with the same properties as ours can be derived from models based on different assumptions on firm behavior.

Some restrictions on the revenue-per-worker function $R_{it}(X_{jt}, \xi_{jt})$ will be required for model identification. Therefore, we use the following linear specification for the revenue function:

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

In this specification, each firm i's revenue is linear in terms of worker attributes (X_{jt}, ξ_{jt}). Coefficients $\boldsymbol{\beta}_{i,X}$ and $\beta_{i,\xi}$ represent employer i's preference for characteristic vector X_j and ξ_j, respectively. We allow each firm to have a unique set of preference parameters. When the optimal choice is suspension of hiring, all coefficients in the revenue function are equal to zero. Similar specifications are commonly used to estimate preference parameters in literature on estimating demand in differentiated product markets (Berry, 1994; Berry, Levinsohn and Pakes, 1995; Petrin, 2002; Bajari and Benkard, 2005; Bajari and Kahn, 2005). These random coefficient models are considerably more flexible than standard logit or probit models, where preference parameters are assumed to be identical across individuals. Although seemingly arbitrary, the linearity assumption can be derived under mild conditions on model primitives.[8] Appendix B shows how the linear revenue function can be derived from common specifications of labor efficiency and the production function.

Given the parametric form in (8), the employer's problem in Equation (5) 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

[8] The proposed functional form is not required for identification. Other parametric specifications may be considered; these include a linear function in which continuous variables are indicated in logarithms rather than in levels similar to those in Bajari and Benkard (2005) and Bajari and Khan (2005). We have tried this linear-in-logs specification, but its performance in explaining inter-industry wage differentials is not significantly different from that of the linear-in-levels specification used in the present study. The linear-in-levels case clearly interprets $\boldsymbol{\beta}_i$ as a WTP vector for worker characteristics; thus, we focus our analysis on this specification.

$x^c_{j,m,t}$ and ξ_{jt} evaluated at the observed optimal choice j^* become

$$\beta_{i,x^c_{j,m,t}} = \frac{\partial \mathbf{w}_t(X_{j^*t}, \xi_{j^*t})}{\partial x^c_{j,m,t}}, \tag{10}$$

$$\beta_{i,\xi} = \frac{\partial \mathbf{w}_t(X_{j^*t}, \xi_{j^*t})}{\partial \xi_{jt}}. \tag{11}$$

Therefor, parameter vector $\boldsymbol{\beta}_i$ is intuitively interpreted as firm i's WTP for worker characteristics.

3 Estimation of 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 function in which X and ε are nonseparable. 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; Chernozhukov, Imbens and Newey, 2007). Whereas 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 considers multiple variables in X to be correlated with unobserved attributes in ε.

Our estimation strategy proceeds in two stages. In the first stage, we recover unobserved worker quality up to a normalization by using nonparametric methods based on the identification results of Matzkin (2003).[9] To consider 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 stage, we use the first-order conditions in Equations (10) and (11) to infer firm-specific parameters on their WTP for continuous worker characteristics.

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

3.1 Estimation of Unobserved Worker Quality

Because 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 distribution function (CDF) of ξ_{jt}. If observed characteristics X_{jt} are uncorrelated with ξ_{jt}, then $F_\xi(\xi_{jt}) = F_{w|x}(w_{jt}|X_{jt})$, where $F_{w|x}(.|.)$ denotes the CDF of wages conditional on worker characteristics (Bajari and Benkard, 2005). In the context of our labor demand model, however, observable worker characteristics, such as education and experience, are likely correlated with unobserved worker quality. To confront the endogeneity problem, we develop an estimator that allows for multiple endogenous variables, following Bajari and Benkard (2005) and Imbens and Newey (2009).

A control variable V is a variable conditional on which X and ε are independent. The first step of our estimation builds on estimators conditional 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 the sub-vectors of the vector of the observed characteristics such that $X = (X_0, X_1)$.[10] In addition, let $X_0 = (x_{01}, .., x_{0M_0})$ represent the variables in X that may be correlated with unobserved quality ξ, where M_0 denotes the number of endogenous variables in X_0. Sub-vector X_1 represents the vector of exogenous variables. We assume that the researcher also observes a vector (Z) of instruments correlated with X_0 but uncorrelated with ξ and that Z has a dimension of $G \geq M_0$.

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 regression

$$x_{0m} = h_m(X_1, Z, \eta_m), \quad m = 1, ..., M_0, \tag{12}$$

[10] To simplify notation, we suppress both the individual sub-index j and the market sub-index t.

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 to estimate 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*

$$\xi = \int_{\boldsymbol{\eta} \in [\mathbf{0},\mathbf{1}]^{M_0}} F_{w|x,\boldsymbol{\eta}}(w|X, \eta) d\eta. \tag{13}$$

Our proof (Appendix C) extends Theorem 1 of Imbens and Newey (2009) and Theorem 4 of Bajari and Benkard (2005) by allowing for multiple endogenous characteristics.

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 by using an empirical analog of $F_{x_{0m}|X_1,Z}(.|.)$. Second, we use the recovered series of η_m to nonparametrically estimate $F_{w|x,\boldsymbol{\eta}}(.|.)$, the integrand function in Equation (13). Third, worker quality is estimated by integrating η out by using Halton draws of an M_0-dimensional unit cube.[11] The same procedure is applied to all workers $j = 1, ..., J_t$ and markets $t = 1, ..., T$.

Several nonparametric methods, such as the kernel method and series estimators, have been proposed to estimate conditional CDFs. Imbens and Newey (2009) find that series estimators are preferable in empirical frameworks similar to ours. Among series estimators, mixtures of normal distributions are frequently used nonparametric estimators (e.g., Bajari,

[11] 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 eight numbers in the sequence corresponding to base 3 are 1/3, 2/3, 1/9, 4/9, 7/9, 2/9, 5/9, and 8/9. To span the domain of the M_0-dimensional unit cube, Halton draws can be formed by using different prime numbers for each dimension. Halton draws exhibit advantages over random draws from $U[0, 1]$ in terms of low variance and few draws (Bhat, 2001; Petrin and Train, 2010).

Fox and Ryan, 2007; Bajari et al., 2011) because of their desirable approximation and consistency properties (e.g., Norets, 2010). We use this type of estimator because it fits the data well and is computationally more tractable for the numeric integration in Equation (13) than other methods.

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:

$$\hat{f}(Y|U;\boldsymbol{\theta}) \equiv \sum_{r=1}^{R(N)} \alpha_r(U,\boldsymbol{\theta}^\alpha)\phi(Y|\mu_r,\sigma_r), \tag{14}$$

where $R(N)$ represents the (integer) number of normal densities as an (increasing) function of sample size N, $\boldsymbol{\theta}$ is the vector of the 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

$$\hat{F}(Y|U;\boldsymbol{\theta}) \equiv \sum_{r=1}^{R(N)} \alpha_r(U,\boldsymbol{\theta}^\alpha)\Phi(Y|\mu_r,\sigma_r), \tag{15}$$

where $\Phi(.|\mu_r,\sigma_r)$ denotes the CDF of the same normal distribution. Each normal density in Equation (14) is weighted by a multinomial logit function $\alpha_r(U,\boldsymbol{\theta}^\alpha)$ with an $(H+1)\times 1$ parameter vector $\boldsymbol{\theta}^\alpha$ defined as

$$\alpha_r(U,\boldsymbol{\theta}^\alpha) = \begin{cases} \dfrac{1}{1+\sum_{l=2}^{R(N)} \exp\left(\theta_{0,l}^\alpha + U\cdot\boldsymbol{\theta}_{U,l}^\alpha\right)} & if\ r=1, \\ \dfrac{\exp\left(\theta_{0,r}^\alpha + U\cdot\boldsymbol{\theta}_{U,r}^\alpha\right)}{1+\sum_{l=2}^{R(N)} \exp\left(\theta_{0,l}^\alpha + U\cdot\boldsymbol{\theta}_{U,l}^\alpha\right)} & if\ r=2,...,R(N). \end{cases} \tag{16}$$

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

In each market $t=1,...,T$, our maximum likelihood estimator for the PDF of an endogenous attribute $x_{0,m}$ conditional on exogenous worker characteristics X_1 and an instrument set Z is defined as

$$\hat{\boldsymbol{\theta}}_{x_{0,m}} \equiv \arg\max_{\boldsymbol{\theta}} \sum_{j=1}^{J_t} \log\{\hat{f}(x_{0,m,j,t}|X_{1,j,t}, Z_{jt}; \boldsymbol{\theta})\}.^{12} \tag{17}$$

After $\hat{\boldsymbol{\theta}}_{x_{0,m}}$ is estimated for each $m = 1, ..., M_0$, the corresponding estimate for the control variable for each worker j in market t is

$$\eta_{m,j,t} = \hat{F}(x_{0,m,j,t}|X_{1,j,t}, Z_{jt}; \hat{\boldsymbol{\theta}}_{x_{0,m}}).^{13} \tag{18}$$

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

$$\hat{\boldsymbol{\theta}}_w \equiv \arg\max_{\boldsymbol{\theta}} \sum_{j=1}^{J_t} \log\{\hat{f}(w_{jt}|X_{jt}, \eta_{jt}; \boldsymbol{\theta})\}. \tag{19}$$

With control variable estimates of $\eta_{m,j,t}$ for all m, $\hat{\boldsymbol{\theta}}_w$ is obtained by solving Equation (19). We can then estimate the unobserved quality of each worker j in market t by using Equation (13):

$$\hat{\xi}_{jt} = \int\limits_{\boldsymbol{\eta} \in [\mathbf{0},\mathbf{1}]^{M_0}} \hat{F}(w_{jt}|X_{jt}, \eta; \hat{\boldsymbol{\theta}}_w) d\eta. \tag{20}$$

[12] 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). We then decrease the value for $R(N)$ to obtain a model as parsimonious as possible provided that it does not change the estimated distribution significantly. Intuitively, this corresponds to eliminating those normal distributions in the mixture with weights close to zero.

[13] Although there could be efficiency gains from pooling observations from different markets in the estimation, we choose to estimate the model for each market t seperately for several reasons. First, if wages differ across markets due to differences in market equilibrium, then pooling observations from different markets is invalid. Second, if the market index t represents years, estimating the model by year allows one to identify unmeasured worker quality without imposing structure on its evolution over time.

3.2 Estimation of Firm WTP Parameters

The labor demand problem described in Equation (5) is characterized by the revenue-per-worker function $R_i(X_j, \xi_j)$. As discussed in the previous section, we consider a linear function of $R(X_j, \xi_j; \beta_i)$ (Equation 8). Under this model specification, Equation (10) suggests that if we recover an estimate of $\partial \mathbf{w}_t(X_{j^*t}, \xi_{j^*t})/\partial x^c_{j,m,t}$, then we can learn a firm's WTP for worker characteristic m. As we observe each worker's characteristics in our data, 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 unobserved worker quality, we can also estimate a firm's WTP for unobserved quality according to $\partial \mathbf{w}_t(X_{j^*t}, \xi_{j^*t})/\partial \xi_{jt}$, following Equation (11).

A practical, flexible way to quantify wage function derivatives at each point in data is to apply local linear regression methods to data on wages, observed worker attributes, and unobserved quality estimates. Bajari and Khan (2005) use this approach to estimate a hedonic price function in the housing market and quantify derivatives of the pricing function. However, two important differences are observed. First, Bajari and Khan assume that ξ is independent of all observed characteristics X. Although this assumption is acceptable in their housing demand model, it is unreasonable for our application because of 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 ξ_{jt}. We separate the two values by first quantifying unobserved worker quality through the methods described above and then treating the estimated $\xi_{j,t}$ as an extra regressor for local linear regression.

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

$$w_{j^*,t} = b_{j^*,0} + b_{j^*,1} x_{j^*,1,t} + ... + b_{j^*,M} x_{j^*,M,t} + b_{j^*\xi} \xi_{j^*,t}, \tag{21}$$

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}.[14]

In the context of nonparametric regression, Fan and Gijbels (1996) provide a formula for the coefficients in Equation (21) for each observation j^*. The $J_t \times 1$ vector of all wages is denoted by $\mathbf{w}_t$, and the vector that stacks all coefficients is denoted by $\mathbf{b}_{j^*}$, which is solved according to

$$\mathbf{b}_{j^*} = \left(Z_t^T \Omega_t Z_t\right)^{-1} Z_t^T \Omega_t \mathbf{w}_t, \tag{22}$$

where Z_t and Ω_t are matrices defined as

$$\begin{aligned} Z_t &= [\mathbf{1} \;\; \mathbf{z}_t] = \begin{bmatrix} 1 & (x_{1,1,t} - x_{j^*,1,t}) & \ldots & (x_{1,M,t} - x_{j^*,M,t}) & (\xi_{1,t} - \xi_{j^*,t}) \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & (x_{J_t,1,t} - x_{j^*,1,t}) & \ldots & (x_{J_t,M,t} - x_{j^*,M,t}) & (\xi_{J_t,t} - \xi_{j^*,t}) \end{bmatrix}, & (23) \\ \Omega_t &= \mathbf{diag}\left(K_{\mathbf{H}}(\mathbf{z}_t)\right). & (24) \end{aligned}$$

$K_{\mathbf{H}}(\mathbf{z}_t)$ is a multivariate kernel function with smoothing parameter matrix $\mathbf{H}$, and $K_{\mathbf{H}}$ is a multivariate standard normal density of dimension $M+1$. As in other practical applications of local linear regression with several covariates, the bandwidth matrix $\mathbf{H}$ is selected by visual inspection of estimates.[15]

The values of $\mathbf{b}_{j^*}$ are estimates of wage function derivatives. According to the first-order conditions in (10) and (11), these values consist of estimates of firm-specific WTP parameters $\beta_{i,x^c_{j,m,t}}$ and $\beta_{i,\xi}$ for observed continuous attributes and unobserved worker quality, where firm i is the employer of worker j^*.

For worker characteristics that take on discrete values, point identification of the random

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

[15] Fan and Gijbels (1996) provide asymptotically optimal methods for bandwidth matrix choice. However, these approaches are unreliable for applications that use several covariates, such as ours and Bajari and Khan (2005). We use this matrix as our starting point for selecting bandwidth parameters and make subsequent adjustments based on goodness-of-fit heuristics.

coefficients of these characteristics cannot be achieved by using first-order conditions similar to those in Equation (10). Instead, we can establish bound estimates for these coefficients by using the condition that firm i's choice of the discrete characteristic observed in the data maximizes profit in Equation (5). For example, suppose that firm i hires worker j^*. Let $\hat{X}_{j^*t}$ and $\bar{X}_{j^*t}$ denote the vectors of observed characteristics with $female = 1$ and $female = 0$, respectively, and all other elements equal the corresponding observed attributes in vector X_{j^*t}. The implicit price faced by employer i for a female worker is then $w_t(\hat{X}_{j^*t}, \xi_{j^*t}) - w_t(\bar{X}_{j^*t}, \xi_{j^*t})$. $\beta_{i,f}$ is denoted as the coefficient for the female dummy in the revenue function. Profit maximization implies that $\beta_{i,f} > w_t(\hat{X}_{j^*t}, \xi_{j^*t}) - w_t(\bar{X}_{j^*t}, \xi_{j^*t})$ if worker j^* is female and $\beta_{i,f} \leq w_t(\hat{X}_{j^*t}, \xi_{j^*t}) - w_t(\bar{X}_{j^*t}, \xi_{j^*t})$ otherwise. That is, if employer i hires a female worker, then i's WTP for this characteristic exceeds the implicit price for the characteristic.[16]

A firm's WTP for a discrete worker characteristic is not point-identified even if the researcher assumes a parametric distribution. This lack of point identification precludes the usage of firm WTP for discrete attributes in our statistical analysis of inter-industry wage differentials. Thus, we focus on firm WTP on continuous attributes, including, education, work experience, and unobserved worker quality.[17]

4 Data

The micro data used 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

[16] 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 applications of hedonic models.

[17] To remedy for the lack of point identification of WTP for discrete attributes, we estimate the mean WTP for gender, race, and marital status conditional on firm characteristics on the assumption of 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 analysis of wage differentials. Not only are the conditional means statistically insignificant, they also jointly explain less than 1% of the wage variation in the wage regressions. For this reason and for the sake of exposition, we focus our analysis on WTP for continuous characteristics.

demographic characteristics. For each individual, the NLSY79 reports age, gender, race, education, marital status, region of residence, employment status, occupation, and earnings. In addition, the NLSY79 asks questions on individual background and employer characteristics. We obtain information on parental education, Armed Force Qualification Test (AFQT) score, and each worker's industrial affiliation.

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, or per week) yields the hourly wage rate. The earnings variable used in this study is the hourly wage for the CPS job, that is, the current or 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. The usual hours worked per week at all jobs are available from 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 industrial wage premiums. Columns (1) and (5) of Table 1 report the raw differences in log hourly wages by industry for both the 1990 and 1993 observations. These differences are computed from cross-section regressions of log wage on a set of industry dummy variables by using one digit Census Industry Classification (CIC) Codes.[18] 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 weighted and unweighted

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

standard deviations of estimates of the industry coefficient. Unweighted standard deviation measures the difference in wages between a randomly chosen industry and the average industry, whereas weighted standard deviation (by employment) measures the difference in wages between a worker in a given industry and the average worker. Both statistics demonstrate 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 industry dummy variables. We estimate industry wage differentials from the cross-section wage function

$$w = X\zeta + D\tau + \varepsilon, \tag{25}$$

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, work experience, gender, race, marital status, occupation, location dummies, union status, veteran status, and several interaction terms. The industry dummy variables are statistically significant in both years, substantial in magnitude, and similar to those estimated with data from the 1970s and 1980s (e.g., Blackburn and Neumark, 1992; Krueger and Summers, 1988). For example, earnings in construction, transportation, communication, and public utilities, are substantially higher than those in the wholesale and retail trade and service industries, even with controls for years of schooling, experience, gender, and race. Adding human capital controls reduces inter-industry wage differentials, as measured by their standard deviations, by 8%–10% in 1990 and 15%–20% in 1993.

Even after various human capital controls are included, the coefficient estimates on industry dummies in Equation (25) may pick up the differences in unobserved worker quality across industries. Previous research has attempted to correct unobserved quality bias in estimated industry effects by including proxies for worker quality, such as test scores in wage regressions (Blackburn and Neumark, 1992). In Columns (3) and (7), we include 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 decline slightly for both the 1990 (from 0.136 to 0.133, unweighted) and 1993 regressions (from 0.115 to 0.114, unweighted). Furthermore, including parental education in the wage regressions only slightly affects the standard deviations of the industry effects, as shown in Columns (4) and (8) of Table 1. These results fail to support the unobserved quality explanation for industry wage differentials, consistent with the conclusion reached by Blackburn and Neumark (1992).

Another approach to solving the problem of unobserved labor quality is to analyze longitudinal data and estimate the 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% pay increase. These results are consistent with the findings by Krueger and Summers (1988), who interpret their findings as evidence that differences in labor quality cannot explain inter-industry wage differentials.[19]

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 rewarded in labor markets. The ability to do well in standard tests may differ from the motivation and perseverance necessary to succeed in the workplace. On the other hand, first-difference estimates rely on the assumption that unobserved quality is time

[19] One notable difference between our first-difference results and those of 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 a sample of displaced workers is not available from NLSY79. However, our estimates yield similar results to those of analyzing non-displaced longitudinal data in Krueger and Summers (1988).

invariant and equally rewarded in all industries and can therefore be differenced out as an individual fixed effect. If labor quality evolves over time, perhaps through learning-by-doing, or if it is valued differently across firms, then an 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 variations in unobserved labor quality.

5 Empirical Results

This section presents estimates of our hedonic labor demand model. We first outline estimation results for the unobserved worker quality recovered in our first stage of estimation. We then present firm WTP distribution estimates based on our model specification. Finally, we assess how much unobserved worker quality and firm WTP for education, work experience, and quality account for inter-industry wage differentials.

5.1 Unobserved Worker Quality

We use NLSY79 data on wages and observed worker characteristics to estimate unobserved worker quality based on Equation (20). Our approach is flexible enough to allow unobserved worker quality to evolve over time and allow firms to reward both observed worker attributes and unobserved labor quality differently. The variables of observed worker characteristics, represented by the vector X, include years of schooling, years of work experience, and dummy variables on gender, race, and marital status.[20] Out of these variables, years of schooling and experience are potentially correlated with unobserved worker quality and constitute the sub-vector X_0. To estimate the control variables for education and experience, we use an instrument vector Z that includes worker age and location dummies. All other observed characteristics are included in sub-vector X_1.

[20] We experiment with alternative specifications of vector X containing other worker characteristics observed in NSLY79. The results on quality estimates and subsequent wage differential analysis do not change significantly. Given that more variables in X would increase computational costs drastically, we focus our analysis on the current parsimonious specification of X without loss of generality.

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

The top two panels of Table 3 report 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. The estimated quality is less significantly correlated with experience than AFQT score and education in both years, but worker quality becomes significantly correlated with experience over time. Learning-by-doing may explain the increasing correlation between worker quality and experience. The correlations between the estimated quality and AFQT score are 0.361 and 0.352 in 1990 and 1993, respectively. The relatively low correlations imply that worker quality rewarded in labor markets may not reflect completely 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. Worker quality is by no means fixed over time according to our estimates, even though it is highly persistent. The evolution of labor quality over the career path may be related to post-school human capital investment, such as learning-by-doing. Thus, standard first-difference estimators cannot difference out the effects of unobserved quality on wages.

5.2 Distributions of WTP Parameters

We estimate the structural model of labor demand presented in Section 2 for both 1990 and 1993. This estimation 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 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. WTP considerably varies for both observed education and work 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. On average, a firm is willing to pay 15.4 cents per hour for each additional year of education.[21] Panels B to H present 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, whereas the mining industry has the lowest mean WTP for education of 14.5. All industry-specific distributions are right-skewed. Specifically, the distribution in the service industry has the longest tail with a standard deviation of 5.0, and the distribution in the construction industry is the least dispersed with a standard deviation of 4.2.

Figure 2 presents the histograms of 1990 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 cents per hour) is lower than the average WTP for a year of education (15.4 cents per hour), but WTP for experience is highly dispersed with a standard deviation of 6.1. Firms in the finance, insurance, and real estate industry and the service industry value work experience the most, with a mean WTP of 6.9, whereas experience is the least valued in the construction industry with a mean WTP of 5.5. In terms of dispersion, the service 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

[21] The estimated WTP for education is low because the estimation is conditional on all other observable characteristics and unobserved worker quality. If we do not consider unobserved worker quality explicitly, the WTP for education would be higher because of the correlation between education and labor quality.

1990 is presented in Figure 3. Because worker quality has no intrinsic units and is normalized between 0 and 1, the values of WTP parameters for quality are unimportant; thus, we focus on their relative levels across industries. The distribution of WTP for worker quality across all industries depicted in Panel A appears bimodal because most firms in several industries (e.g., mining) do not value worker quality much, whereas most firms in several others (e.g., finance, insurance, and real estate) highly value worker quality. The two modes in the distributions in the construction and service industries also contribute to the overall bimodal distribution. Based on Panels B to H, (unobserved) worker quality is less valuable to firms in the mining, construction, and wholesale and retail trade industries than to firms in the finance, insurance and real estate, service, and transportation, communication and public utilities industries. Similar to the distribution of WTP for education, that for quality is most dispersed in the service 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 Figures 4 to 6. Firms in all industries value education more in 1993 than in 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 highly in 1993 than in 1990, and the distributions of WTP for quality are more dispersed in 1993. These results are consistent with the increasing returns to both education and unobserved ability documented in literature. On the contrary, work experience is less valued by firms, and firms' valuation of experience is less dispersed in 1993 than in 1990, as indicated by the lower means and variances of WTP parameters in Figure 5 than those in Figure 2.

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

5.3 Inter-industry Wage Differentials

Columns (2) and (6) of Table 5 present estimates of coefficient τ in Equation (25) by adding recovered worker quality as an extra 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 high and statistically significant. The magnitude of the coefficients on industry dummies declines, 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 declines by a similar magnitude. These results suggest that unmeasured worker quality is an important driving force of inter-industry wage differentials. Worker quality also accounts for a large portion of the overall wage variation as the adjusted R^2 of the log wage regression increases from 0.356 to 0.861 in 1990 and from 0.376 to 0.857 in 1993 once worker quality is included in the regressions.[22]

Columns (3) and (7) of Table 5 present estimates of τ coefficients in Equation (25) by adding recovered firm-specific WTP to education, experience, and quality as additional control variables. The industry wage premiums in both years decrease but remain 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 increases slightly from 0.356 to 0.390 in 1990. Compared with worker quality (columns 2 and 6), firm WTP can account for only 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 regressions in columns (4) and (8), the standard deviations of industry wage differentials almost stay the same as in the regressions that control only for worker quality.

[22] Using a different dataset and different methodology, Abowd, Kramarz and Margolis (1999) also find that wage regressions that include person effects can explain between 77% to 83% of wage variance, whereas regressions that exclude person effects can explain only between 30% to 55% of the variance.

We further decompose the contribution of worker heterogeneity (in terms of unobserved labor quality) and firm heterogeneity (measured by WTP for human capital attributes) to inter-industry wage differentials. We estimate inter-industry wage differentials by regressing (25) with two-digit industry dummies while controlling for education, years of experience and its square, gender, race, marital status, union and veteran status, region dummies, occupation, parental education, AFQT test score, and several interaction terms.[23] Table 6 uses the industry-level averages of worker quality and firm-specific WTP parameters to account for the industry wage differentials. The first column of Table 6 shows the separate influence of worker heterogeneity on explaining industry effects by regressing the estimated industry wage premiums on industry-average worker quality alone. Similarly, column (2) of Table 6 presents industry-level regressions using industry-average firm WTP parameters alone. Industry-average worker quality alone accounts for approximately two thirds of observed inter-industry wage variation, whereas the explanatory power of industry-average firm WTP parameters is relatively low. Therefore, individual effects, as measured by average worker quality, are more important than firm effects, as measured by WTP parameters, for explaining inter-industry wage differentials.[24] The combination of worker quality and firm WTP can explain close to 80% of the overall variations in inter-industry wage differentials in both years.

6 Concluding Remarks

In this paper we propose an alternative approach to explain inter-industry wage differentials by using a hedonic model of labor demand. The model allows the nonparametric identifica-

[23] These results are available from the authors upon request.

[24] Using matched employer–employee data from France, Abowd, Kramarz, and Margolis (1999) find that individual heterogeneity alone explains 84%–92% of the inter-industry wage variation, whereas firm heterogeneity alone explains only 7%–25%. Thus, they reach the same conclusion as ours that individual effects are more important than firm effects for explaining inter-industry wage differentials. However, our approach does not require the use of matched employer–employee data and does not impose the assumption that unobserved labor quality is fixed over time.

tion of unobserved worker quality as well as employer-specific WTP for worker attributes. Our approach does not require the use of matched employer–employee panels to separate 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 and that unmeasured worker quality varies over one's career despite its high degree of persistence.

An important caveat to the effects 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, and marital status, even if the researcher makes strong assumptions about the distribution of WTP parameters. Our framework shares this feature with other related models (e.g., Bajari and Benkard, 2005; Bajari and Khan, 2005). Therefore, we cannot identify which portion of inter-industry wage differentials can be explained by WTP for discrete attributes. Finding a set of mild assumptions that can point-identify employer WTP for discrete attributes is beyond the scope of this study and is thus left for future work.

As in the hedonic model of differentiated products proposed by Bajari and Benkard (2005), supply-side assumptions on worker behavior are not required in our model. An interesting extension of our framework is to explicitly model labor supply behavior and allow workers to choose which firm to work for. In such an equilibrium model, compensating differences may be separately identified from WTP parameters, but such exercise involves various challenges in identification and remains an important topic for future research.

Appendix A: Proof of Proposition 1

Proposition 1 is illustrated as follows. 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 that $w_{jt} > w_{j't}$ 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 observation implies that no one would hire worker j in market t and is thus a contradiction.

Suppose that $w_{jt} \leq w_{j't}$ for some market t in which both workers j and j' are employed and $X_{jt} = X_{j't}$ and $\xi_{jt} > \xi_{j't}$. Given that $R_{it}(X_{jt}, \xi_{jt})$ strictly increases in ξ_{jt}, $R_{it}(X_{jt}, \xi_{jt}) - w_{jt} > R_{it}(X_{j't}, \xi_{j't}) - w_{j't}$ for all employers $i = 1, ..., V_t$. This observation implies that no one would hire worker j' in market t and is thus 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 $M < \infty$. Given that $|R_{it}(X_{jt}, \xi_{jt}) - R_{it}(X_{j't}, \xi_{j't})| = |[(R_{it}(X_{jt}, \xi_{jt}) - w_{jt}) - (R_{it}(X_{j't}, \xi_{j't}) - w_{j't})] + (w_{jt} - w_{j't})|$,

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

Assuming that without loss of generality $w_{jt} > w_{j't}$, then the second term on the left-hand side, $w_{jt} - w_{j't}$, is positive. Because 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,

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

In this instance, we use the fact that both workers have positive demand to limit how much their wages can vary.

Appendix B: An Example of Deriving Linear Revenue Function

In what follows, we illustrate how a linear revenue function can be derived from common specifications of labor efficiency and production function. We suppress the market subindex t in our notation for ease of exposition.

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

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

In addition, consider a CES production function:

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

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 demand takes the form of $K_i^* = \delta_i E_j$, where

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

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

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

Therefore, under the model specification, the employer's problem is simplified to

$$\pi_i = \max_{j \in \{0,1,\ldots,J)} \{R(E_j) - w_j, 0\},$$

where the revenue function $R(E_j) = \gamma_i E_j$ and 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

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

Given the specification for labor efficiency, the revenue per worker function has the following parametric form

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

where the coefficient vector $\boldsymbol{\beta}_i$ is the product of the vector of efficiency unit coefficients $\rho' s$ and γ_i.

Appendix C: Proof of Proposition 2

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 $h_m(X_1, Z, \eta_m)$. According to the proof of Lemma 1 of

Matzkin (2003), for each $m = 1, ..., M_0$,

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

where the third equality follows from the monotonicity assumption, the fourth equality follows from the independence between (X_1, Z) and $\eta = (\eta_1, ..., \eta_{M_0})$, and the last equality is the result of normalizing η_m so that it lies in $U[0,1]$.

Next, we show that vector η consists of control variables conditional on which X and ξ are independent by adapting the proof of Theorem 1 of Imbens and Newey (2009) for multiple endogenous variables. For any bounded function $a(x_0, x_1)$, it follows from the independence of (X_1, Z) and (ξ, η) that

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

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

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

which indicates the independence between X and ξ conditional on η.

Finally, given that ξ and each η_m are normalized such that all of them follow $U[0,1]$, for each market t and worker j,

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

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

References

[1] Abowd, J., F. Kramarz, P. Lengermann, K.L. McKinney, and S. Roux (2012), "Persistent Inter-industry Wage Differences: Rent Sharing and Opportunity Costs," *IZA Journal of Labor Economics*, 1:7, doi:10.1186/2193-8997-1-7.

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

[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., 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.

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

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

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

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

[11] Blackburn, M. and D. Neumark (1992), "Unobserved Ability, Efficiency Wages, and Interindustry Wage Differentials," *Quarterly Journal of Economics* 107: 1421–1436.

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

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

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

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

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

[20] 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.

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

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

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

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

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

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

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

[28] 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.

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

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

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

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

[33] Murphy, K.M., and R. Topel (1987a), "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.

[34] Murphy, K.M. and R. Topel (1987b), "Unemployment, Risk and Earnings," in *Unemployment and the Structure of Labor Markets*, K. Lang and J. Leonard (eds.), Oxford: Basil Blackwell.

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

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

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

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

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

[40] 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.

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

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

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

	1990 Cross Section				1993 Cross Section				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 service.

[b] Other control variables are education, years of experience and its square, gender dummy, race dummy, ever married dummy, union and veteran status, four region dummies, three 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	1990 Mean	1990 (Std.)	1993 Mean	1993 (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+ years	0.547	(0.248)	0.551	(0.253)
By AFQT percentile scores				
AFQT $\leq$ 25	0.403	(0.248)	0.394	(0.245)
25 < AFQT $\leq$ 50	0.507	(0.267)	0.499	(0.265)
50 < AFQT $\leq$ 75	0.580	(0.265)	0.560	(0.267)
AFQT > 75	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)
Service	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 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 service. Other control variables are education, years of experience and its square, gender dummy, race dummy, ever married dummy, union and veteran status, four region dummies, three 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.

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 parameters	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 parameters	No	Yes	Yes
R squared	0.697	0.423	0.811
Adjusted R squared	0.689	0.377	0.790

Figure 1: Firm WTP for Education Across Industries, 1990

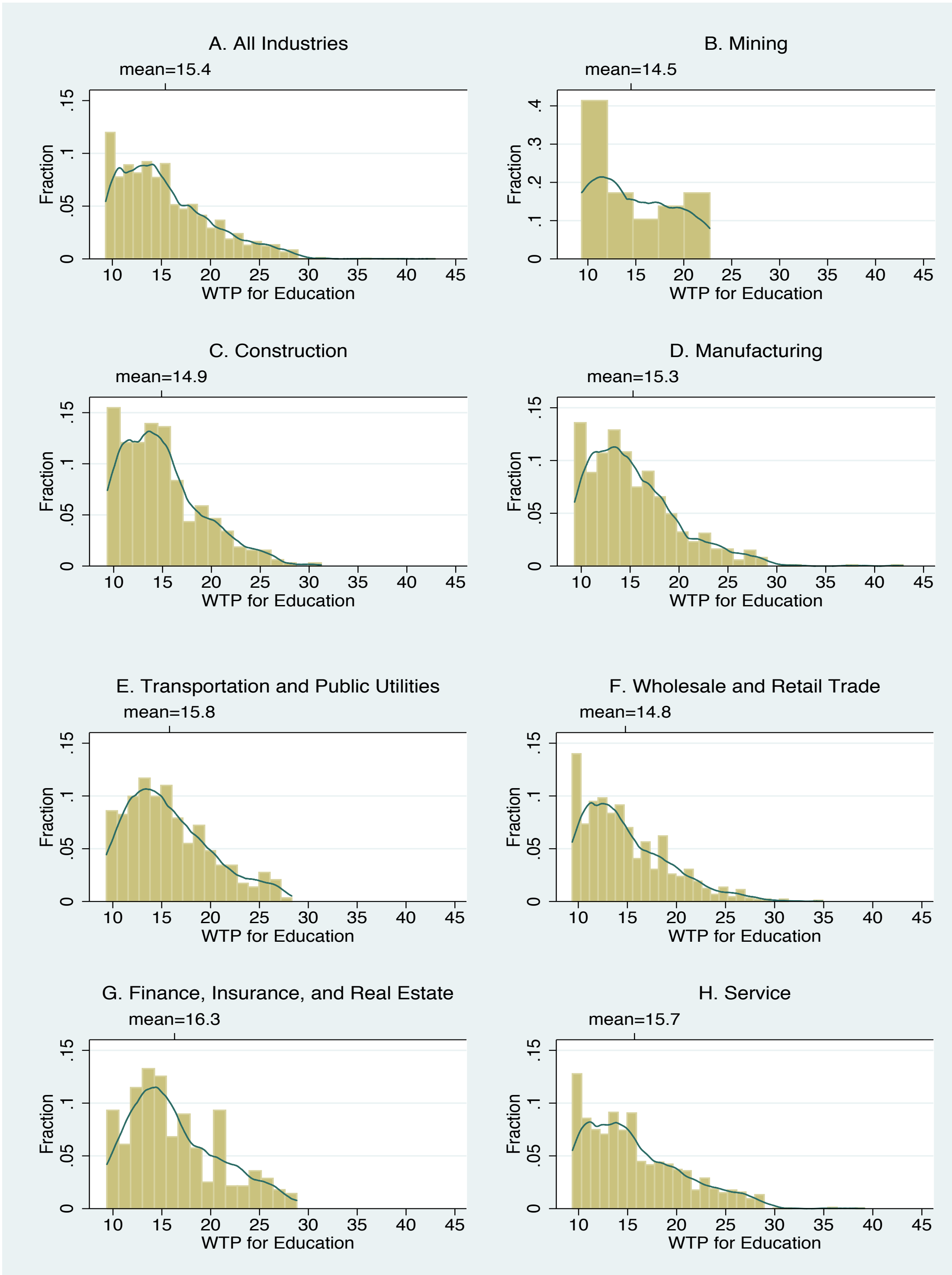

Figure 2: Firm WTP for Work Experience Across Industries, 1990

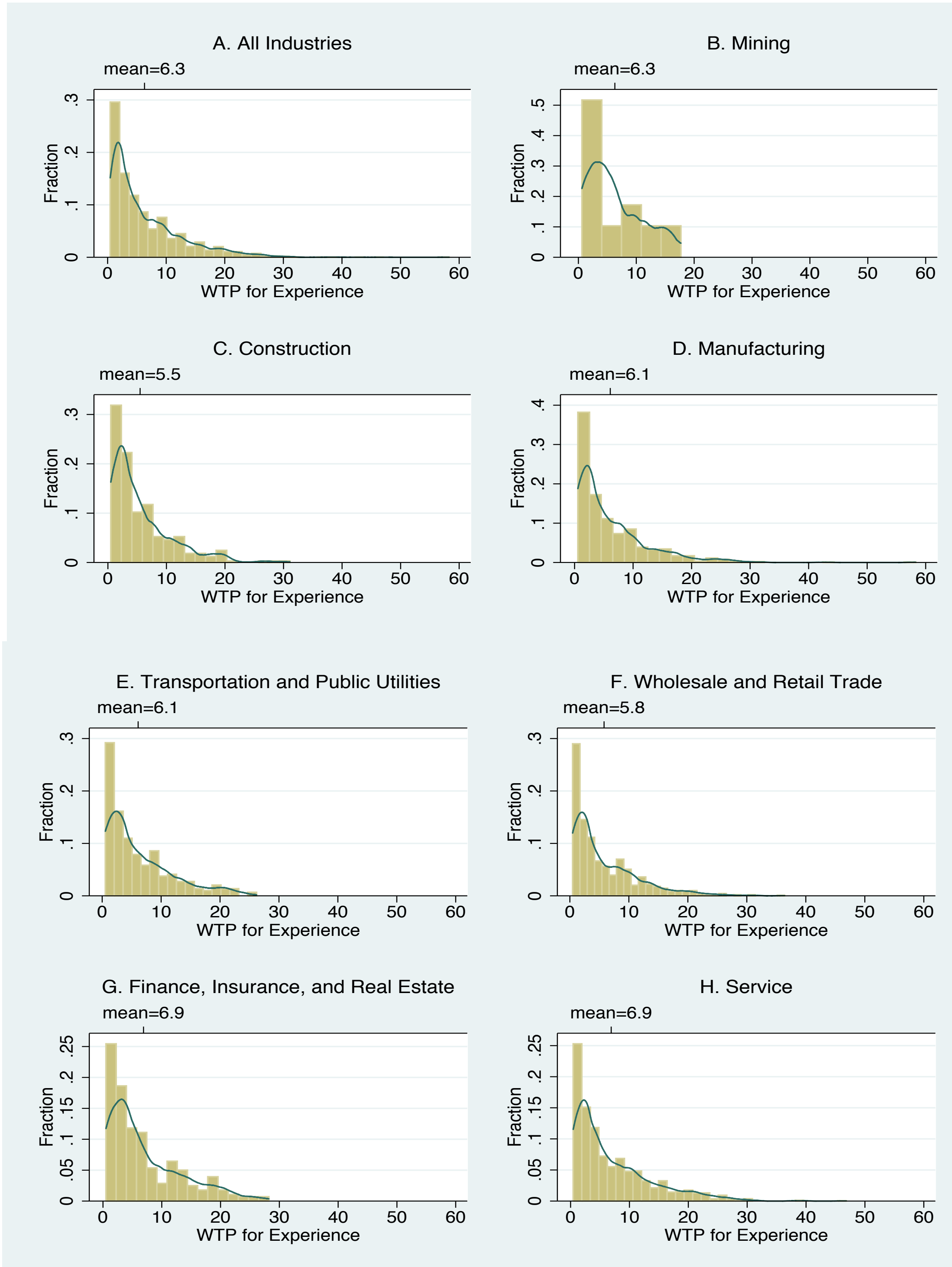

Figure 3: Firm WTP for Worker Quality Across Industries, 1990

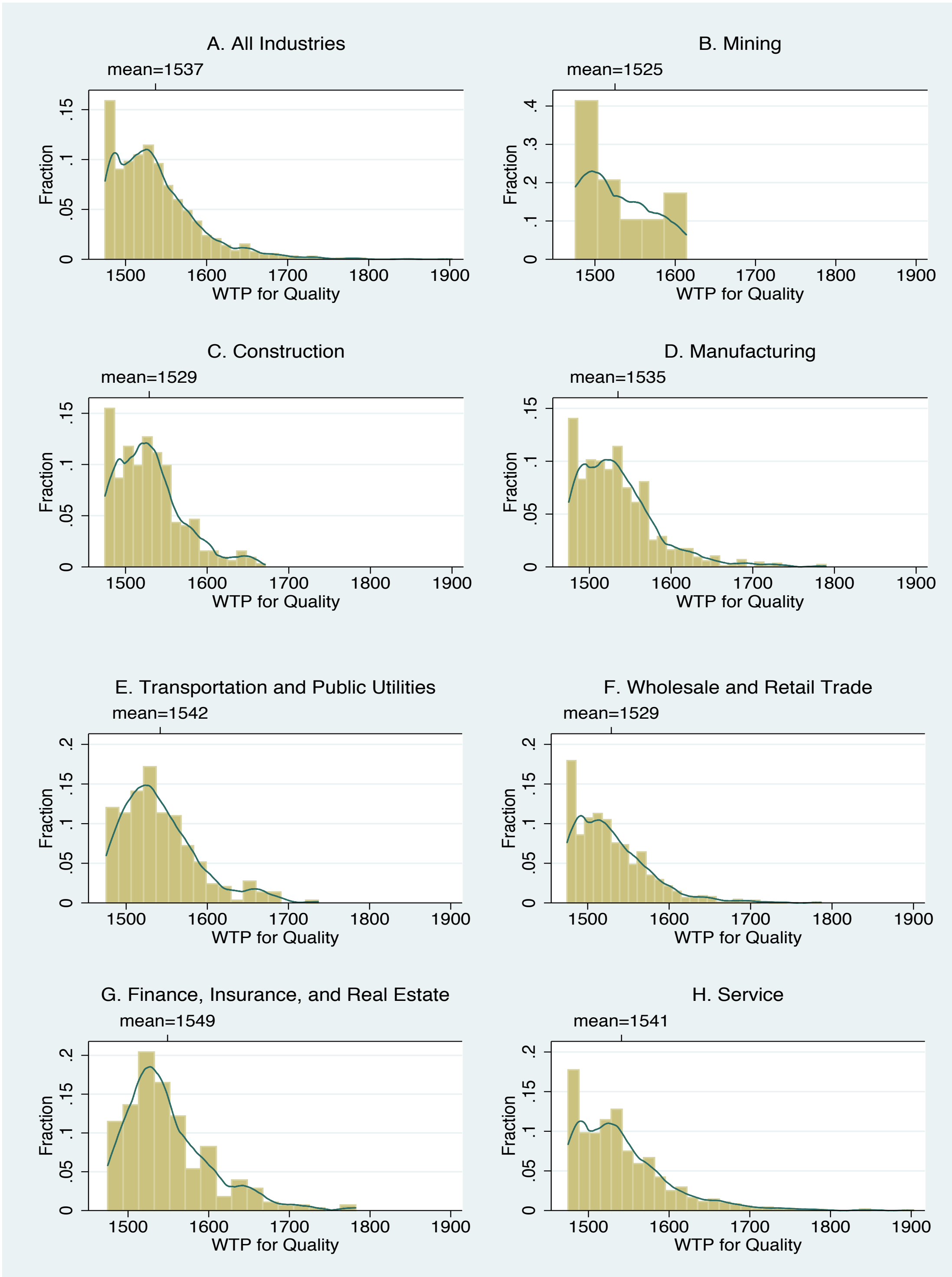

Figure 4: Firm WTP for Education Across Industries, 1993

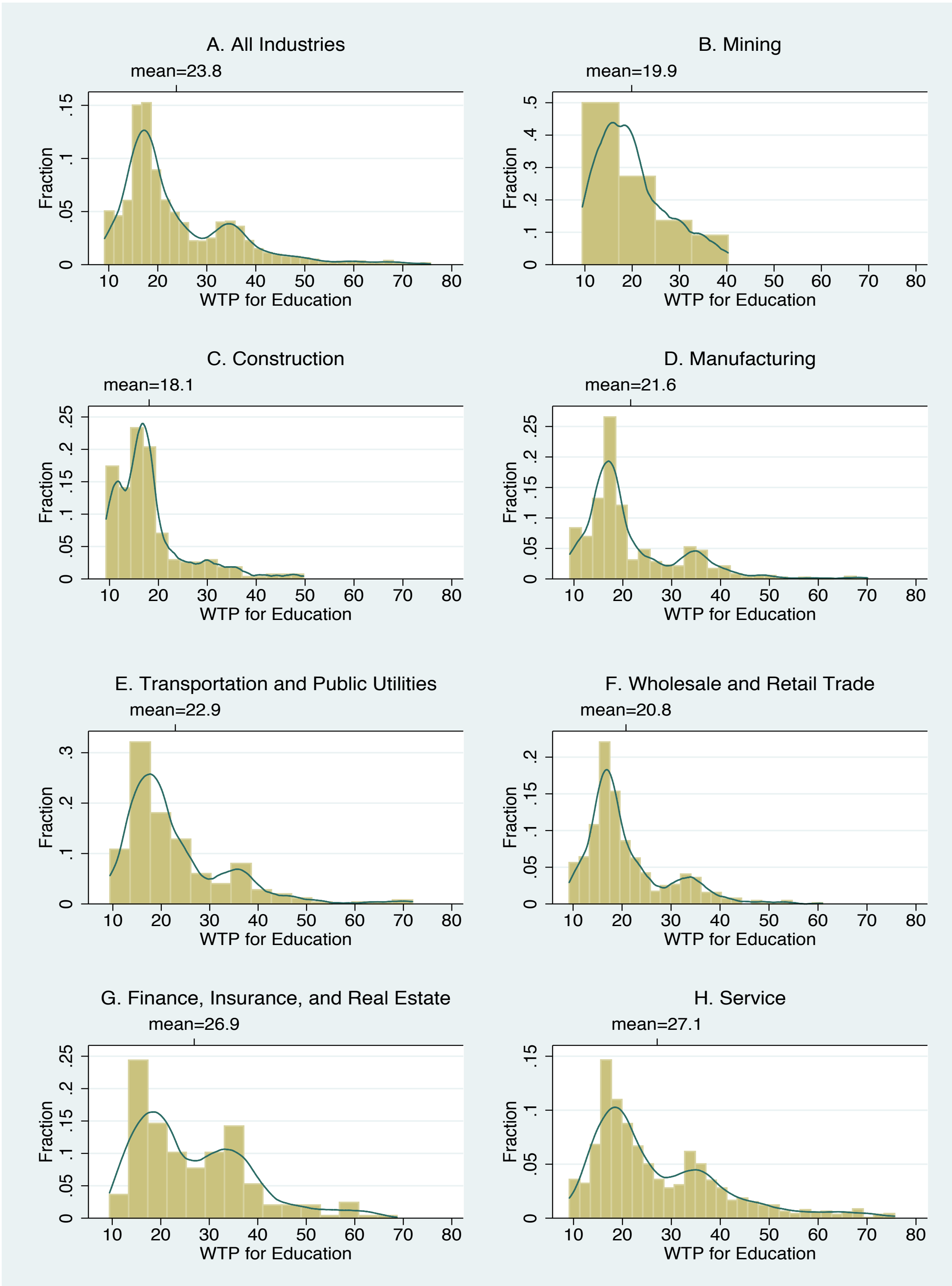

Figure 5: Firm WTP for Work Experience Across Industries, 1993

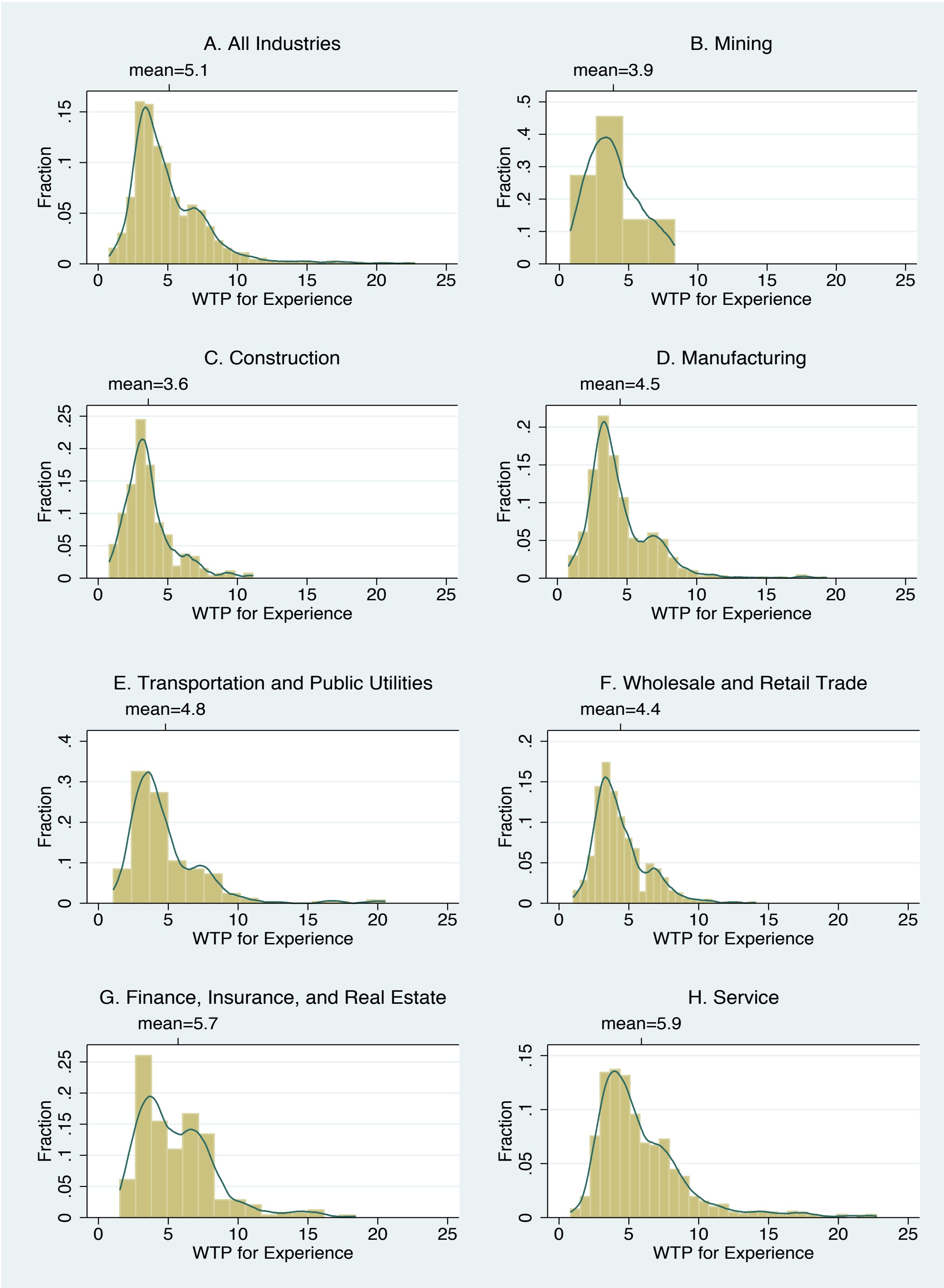

Figure 6: Firm WTP for Worker Quality Across Industries, 1993