

# Comparative Advantage, Learning, and Sectoral Wage Determination, *Journal of Labor Economics*, 2005

Wenzhi Wang \*

August 26, 2024

Gibbons et al. (2005)

## 1. Introduction

We analyze the theoretical and econometric implications of comparative advantage and learning for the wages and sector affiliations of individuals and for changes in these variables over workers' careers. After developing the theory and econometrics, we turn to two empirical applications of our methodology concerning the wages and allocations of workers across occupations and across industries.

Our focus on comparative advantage is motivated by a large and established literature. Many have found that the average characteristics of individuals vary by sector. Furthermore, several have found that the measured returns to individuals' observed characteristics vary by sector.

Our focus on learning is motivated by a smaller and more recent literature. While Jovanovic (1979) and others showed long ago that learning models could provide new interpretations for important features of the data (such as return to seniority and the increase in the variance of wages with experience), recent work has built on these foundations to derive and test novel implications, many of which have survived confrontations with data.

Our theoretical model emphasizes the role of worker skills that cannot be measured by an econometrician. To clarify the exposition of the econometrics, we develop the theory in stages. We begin with two models in which workers' skills are equally valued in all sectors. In the first of these models, all labor market participants have perfect information about workers' skills; in the second, information is initially imperfect, but output observations convey additional information over time and so endogenize wage changes. We then develop two other models in which different sectors place different values on workers' skills and workers sort themselves into different sectors on the basis of comparative advantage. In the first of these latter two models, labor market participants again have perfect information

---

\*This note is written in my MPhil period at the University of Oxford.

about workers' skills; in the second, information is again imperfect, so learning endogenizes not only wage changes but now also sector mobility.

As is well known, in our simplest theoretical model (in which worker skills are equally valued in all sectors and there is no learning by labor market participants), the returns to time-varying worker characteristics can be estimated using first differences to eliminate individual fixed effects that are unmeasured by the econometrician. Similarly, in this simplest model, first differences can be used to estimate sectoral wage differentials without bias from unmeasured fixed effects. Unfortunately, first-difference estimation is not appropriate for any of the three other theoretical models we develop. Simply put, in these models, a worker's fixed ability does not translate into a fixed effect in a wage equation, so first-differencing the wage equation does not correct ability bias.

Our theoretical model rely heavily on the assumption of normality. Many models that rely on normality can be estimated by maximum likelihood or two-step methods, but estimating our dynamic model of wage determination and sector affiliation would be computationally difficult because it entails more than two sectors and more than two periods. In addition, it is not necessary to estimate the full model when the parameters of interest are those that determine the returns to skills and wage differences across sectors. We therefore undertake the more modest task of estimating the wage equations in each sector.

We show that our richest theoretical model produces a random coefficients econometric model in a panel data setting, which can be estimated using a nonlinear instrumental variables technique. Even in this richest model, consistent estimates of the effects of both measured and unmeasured skills on wage require neither distributional assumptions nor standard exclusion restrictions. Instead, the estimation strategy utilizes natural restrictions available in panel data with three or more observations per person.

After developing the theory and econometrics, we undertake two empirical investigations concerning sorting and wage differentials across occupations and industries using individual-level panel data from NLSY. Our richest theoretical model is consistent with several familiar facts about wage determination: *a typical individual's wage increases with experience, the variance of the wage distribution across individuals increases with experience, and the skewness of the wage distribution increases with experience.* But other models can also explain these basic facts, so we focus on our model's further predictions concerning the returns to skills and the resulting allocation of workers across sectors. For both occupations and industries, we find important variation in sector-specific returns to observed and unobserved skills. Furthermore, in both cases, high-wage sectors employ high-skill workers and offer high returns to workers' skills, so estimates of sectoral wage differences that do not account for sector-specific returns to skill and the sorting of workers across sectors on the basis of unmeasured skills are misleading and difficult to interpret.

## 2. Theory and Econometrics

The four theoretical models analyzed below are special cases of the following model. If worker  $i$  is employed in sector  $j$  at time  $t$ , the worker's output is

$$y_{ijt} = \exp(X_{it}\beta_j + \psi_{ijt}), \quad (1)$$

where  $X_{it}$  is a vector of human capital and demographic variables measured by the econometrician and  $\psi_{ijt}$  represents determinants of productivity that are not measured by the econometrician. The worker characteristics  $X_{it}$  and the slope vector  $\beta_j$  are known by all labor market participants at the beginning of period  $t$ ; the realized output  $y_{ijt}$  is observed by all labor market participants at the end of period  $t$ . The error term  $\psi_{ijt}$  has the components

$$\psi_{ijt} = Z_i + b_j(\eta_i + \varepsilon_{ijt}) + c_j, \quad (2)$$

where  $Z_i$  denotes the portion of worker  $i$ 's productive ability that is equally valued in all sectors,  $\eta_i$  denotes the portion that is differentially valued across sectors, and  $\varepsilon_{ijt}$  is a random error. The coefficients  $\{b_j, c_j; j = 1, \dots, J\}$  are fixed and known to all labor market participants. The noise terms  $\varepsilon_{ijt}$  are normal with zero mean and precision  $h_{ve}$  (i.e., variance  $\sigma_\varepsilon^2 = 1/h_\varepsilon$ ) and are independent of each other and of all the other random variables in the model.

In developing the theory and econometrics, we treat  $Z_i$  and  $\eta_i$  differently. We assume throughout that  $Z_i$  is observed by all labor market participants; that is the standard case of a fixed effect that the econometrician cannot observe but market participants can. For  $\eta_i$ , however, we consider two cases: perfect information (no learning by market participants, as with  $Z_i$ ) and imperfect information (learning).

In the imperfect information case, all labor market participants share symmetric but imperfect information about  $\eta_i$ . In particular, given their initial information ( $Z_i$  and  $X_{i1}$ ), all participants in the labor market share the prior belief that  $\eta_i$  is normal with mean  $m$  and precision  $h$ . Subsequent productivity observations,  $y_{ijt}$ , refine this belief. Information in the labor market therefore remains symmetric and improves over time. For simplicity, we assume that subsequent realizations of measured skills,  $X_{it}$ , are conditionally independent of  $\eta_i$  given  $Z_i$  and  $X_{i1}$ . This assumption is not only convenient but realistic because the major time-varying element of  $X_{it}$  is experience. Thus, market participants can compute

$$s_{it} = \frac{\ln y_{ijt} - X_{it}\beta_j - Z_i - c_j}{b_j}, \quad (3)$$

which yields  $s_{it} = \epsilon_i + \varepsilon_{ijt}$ , a noisy signal about the worker's ability that is independent of the worker's sector during period  $t$ . We call  $s_{it}$  the worker's normalized productivity observation for period  $t$ . Let  $s_i^t = (s_{i1}, \dots, s_{it})$  denote the history of the worker's normalized productivity observations through period  $t$ . Then the posterior distribution of  $\eta_i$  given history  $s_i^t$  is normal with mean

$$m_t(s_i^t) = \frac{hm + h_\varepsilon(s_{i1} + \dots + s_{it})}{h + th_\varepsilon} \quad (4)$$

and precision  $h_t = h + th_\epsilon$ .

To close this model, we assume that workers are risk neutral and that there is no cost to firms or workers at the beginning or end of a job (i.e., no hiring, firing, or mobility costs), so we can restrict our attention to single-period compensation contracts. For simplicity, we further restrict attention to contracts that specify the period's wage before the period's production occurs (as opposed to piece-rate contracts). Competition among firms causes each firm in a given sector to offer a given worker a wage equal to the expected value of the worker's output in that sector, given the worker's observed characteristics and history of previous output realizations.

It is not controversial that workers' productive abilities are imprecisely measured in standard micro data sets. But if unmeasured skills are to explain estimated sector wage differentials, then these skills must be non-randomly allocated across sectors. This, too, is plausible, for example because different sectors use different technologies that require workers' skills in different proportions. But if this unmeasured skill explanation for measured sectoral wage differentials is correct, it suggests that the few skills that are measured in standard micro data sets (hereafter "measured skills") should be systematically related to the sector in which the worker is employed. We investigate this prediction about measured skills in our empirical work on occupations and industries in later sections. In this section's discussion of econometric issues, however, we confine our attention to estimating the role of unmeasured skills.

## 2.1. Sorting without Comparative Advantage

In this subsection, we ignore the possibility of comparative advantage by assuming that  $b_j = b$  for every  $j$ , so that a worker's unmeasured ability is  $Z_i + b\eta_i$  and is equally valued in every sector. Continuing in this vein, we also assume in this section that  $\beta_j = \beta$  for every  $j$ . But we allow the intercepts  $c_j$  to vary by sector, in keeping with the possibility that measured sector premia may reflect true sector effects. Of course, all else constant, jobs in sectors with high values of  $c_j$  may be more attractive (depending on the source of  $c_j$ , such as rent sharing vs. compensating differentials). If some sectors are more attractive, issues such as queuing and rationing arise. Because our main interest is in the richer model with comparative advantage in the next section, we do not formally address queuing or rationing here.

In the perfect information case without comparative advantage, all firms know that the worker's ability is  $Z_i + b\eta_i$ . As always, the wage offered to worker  $i$  by firms in sector  $j$  in period  $t$  is the worker's expected output in that sector, but the only uncertainty in this case is the error term  $b\varepsilon_{ijt}$  in (2). Recall that if  $\log \theta$  is normally distributed with mean  $\mu$  and variance  $\sigma^2$ , then  $\mathbb{E}(\theta) = \exp\{\mu + (1/2)\sigma^2\}$ . Therefore, the log wage offered to worker  $i$  in sector  $j$  in period  $t$  is

$$\ln w_{ijt} = X_{it}\beta + Z_i + b\eta_i + c_j + (1/2)b^2\sigma_\varepsilon^2. \quad (5)$$

Turning to the imperfect information case without comparative advantage, in each period, firms in sector  $j$  bid worker  $i$ 's wage up to the worker's expected output in that sector (conditional on the

publicly observable information available at that date), so the log wage is

$$\ln w_{ijt} = X_{it}\beta + Z_i + bm_{i,t-1} + c_j + (1/2)b^2\sigma_t^2, \quad (6)$$

where  $m_{i,t-1}$  is shorthand for  $m_{t-1}(s_i^{t-1})$  and  $\sigma_t^2 = (h + th_\varepsilon) / \{h_\varepsilon [h + (t-1)h_\varepsilon]\}$ . Note that  $\sigma_t^2$  converges to  $\sigma_\varepsilon^2$ , the corresponding variance term in equation (5), as the number of periods  $t$  goes to infinity. Also, in both the perfect and the imperfect information cases, the worker's ability  $\eta_i$  is unmeasured by the econometrician (as in  $Z_i$ ); in the latter case,  $\eta_i$  is also unobserved by labor market participants (unlike  $Z_i$ ). Note that, since  $t$  represents the number of years of experience in the model, the error component  $(1/2)b^2\sigma_t^2$  will be captured by a function in labor market experience that we include in all estimated models.

## 2.2. Estimation without Comparative Advantage

In the absence of both learning and comparative advantage, the source of possible bias in conventional cross-section estimates of sectoral log wage differentials is the potential partial correlation between sector affiliation and unmeasured skills ( $Z_i$  and  $\eta_i$ ) conditional on measured skills ( $X_{it}$ ). In this simplest case, the worker's fixed ability ( $Z_i + b\eta_i$ ) creates a worker fixed effect in the wage regression, which can be eliminated in standard fashion. For example, a first-differenced regression eliminates the fixed effects  $Z_i + b\eta_i$  in (5).

Even without comparative advantage, however, learning implies that fixed ability is not a fixed effect in the earnings equation. The key property of our learning model is that Bayesian beliefs are a martingale. That is, the conditional expectation  $m_t(s_i^t)$  in (4) obeys the law of motion

$$m_{it} = m_{i,t-1} + \xi_{it}, \quad (7)$$

where  $\xi_{it}$  is a noise term orthogonal to  $m_{i,t-1}$ . In somewhat more intuitive terms, the market begins period  $t$  with the information contained in  $s_i^{t-1}$  and then extracts new information about  $\eta_i$  from the output observation  $y_{ijt}$  (or, equivalently,  $s_{it}$ ). But the new information that can be extracted from  $y_{ijt}$  is precisely the part that could not be forecasted from  $s_i^{t-1}$ . Hence, the innovation  $\xi_{it}$  is orthogonal to the prior belief  $m_{i,t-1}$ .

Formally, a first-differenced regression eliminates  $Z_i$  but not  $m_{i,t-1}$  from (6). Instead, first-differencing (6) for a worker who switches from sector  $j$  to sector  $j'$  yields

$$\begin{aligned} \ln w_{it} - \ln w_{i,t-1} &= (X_{it} - X_{i,t-1})\beta + b(m_{i,t-1} - m_{i,t-2}) \\ &\quad + (c_{jj} - c_j) + (1/2)b^2(\sigma_t^2 - \sigma_{t-1}^2), \end{aligned} \quad (8)$$

where  $m_{i,t-1} - m_{i,t-2} = \xi_{i,t-1}$ . But  $\xi_{i,t-1}$  may be correlated with the change in sector affiliation through whatever (unmodeled) process led unmeasured ability to be correlated with sector affiliation in the first place. Thus, with learning, first-differenced estimates of sectoral wage differentials are biased if the change in the residual is correlated with the change in sector affiliation.

Fortunately, this endogeneity problem is simple to correct because the new information summarized in  $\xi_{i,t-1}$  is not related to wage, skill, or sector information in period  $t-1$  or earlier. For example,

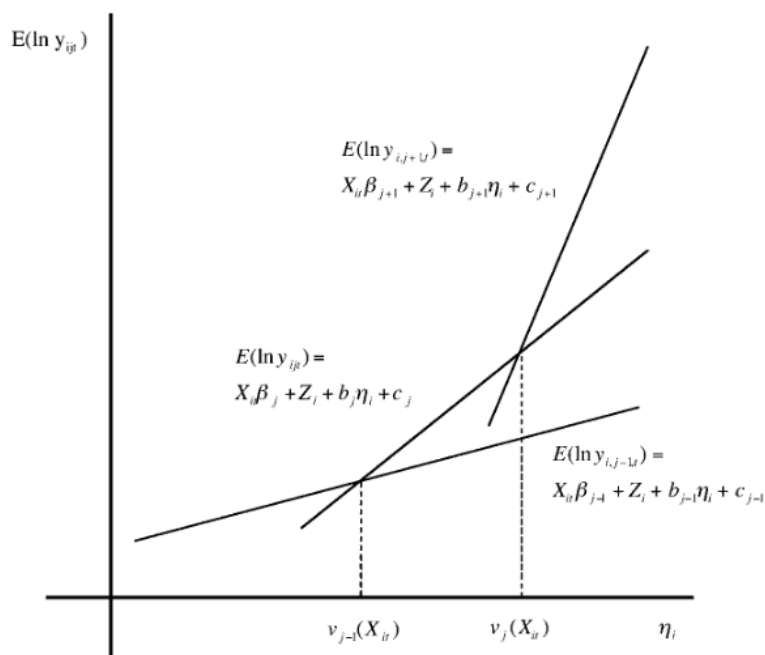
equation (8) can be estimated by two-stage least squares using the interaction of the worker's (publicly observable) score on an ability test (taken before the worker entered the labor market) and the worker's sector affiliation at  $t - 1$  as a valid instrumental variable for changes (between  $t - 1$  and  $t$ ) in sector affiliation.

### 2.3. Sorting with Comparative Advantage

In this section, we relax the assumption that a worker's ability is equally valued in every sector. By introducing comparative advantage, we endogenize sector affiliation. By subsequently introducing learning, we endogenize not only wage changes but also sector mobility.

To analyze comparative advantage, we now return to the production function specified in (1) and (2), where the slope coefficients  $\beta_j$  in (1) and  $b_j$  in (2) vary by sector. We index the  $J$  sectors so that  $b_j$  strictly increases in  $j$ : sector  $j + 1$  values the worker's ability  $\eta_j$  more than does sector  $j$ . In keeping with the notion that ability is productive, we assume that  $b_1 > 0$ . Given a fixed  $X_i$ , there exist critical values of  $\eta_i$  that determine the efficient assignment of workers to sectors. Denoting these critical values by  $\{v_j(X_{it}); j = 0, 1, \dots, J\}$ , the efficient assignment rule assigns worker  $\eta_i$  to sector  $j$  if and only if  $v_{j-1}(X_{it}) < \eta_i < v_j(X_{it})$ , where  $v_0(X_{it}) = -\infty$ ,  $v_n(X_{it}) = \infty$ , and  $v_j(X_{it})$  strictly increases in  $j$ . See Figure 1.

Figure 1: Efficient assignment with comparative advantage



We again analyze first perfect and then imperfect information. In the perfect information case, firms in sector  $j$  bid worker  $i$ 's wage up to the expected output in that sector:

$$\ln w_{ijt} = X_{it}\beta_j + Z_i + b_j\eta_i + c_j + (1/2)b_j^2\sigma_\varepsilon^2, \quad (9)$$

analogous to (5) but with the sector-specific returns  $\beta_j$  and  $b_j$ .<sup>1</sup> *If the worker faces no mobility constraints, worker  $i$  will choose to work in sector  $j$  if  $v_{j-1}(X_{it}) < \eta_i < v_j(X_{it})$ .* Thus, taking the model literally, sector mobility in the perfect information case is driven entirely by changes in  $X_{it}\beta_j$ . One could envision exogenous shocks to sector demand that produce additional sector mobility in this model, but we will not formally model such shocks for the same reason that we did not model queues or rationing in Section II.A: our ultimate interest is in the model with comparative advantage and learning, which gives a coherent account of sectoral mobility without reference to queues, rationing, or sectoral shocks. Whatever the reason that worker  $i$  is employed in sector  $j$  in period  $t$ , in the perfect information case with comparative advantage, we assume that the worker's wage is given by (9).

In the imperfect information case, we again assume that information in the labor market is symmetric but imperfect. As in the model of learning without comparative advantage, all participants in the labor market share the prior belief that  $\eta_i$  is normal with mean  $m$  and precision  $h$ , conditional on their initial information  $Z_i$  and  $X_{it}$ . Inferences from the productivity observations,  $y_{ijt}$ , are greatly simplified because the information content of an output observation is constant across sectors; that is, (2) involves  $b_j(\eta_i + \varepsilon_{ijt})$  rather than  $b_j\eta_i + \varepsilon_{ijt}$ . This functional form is what allows us to define the normalized productivity observation for worker  $i$  in period  $t$ ,  $s_{it}$  from (3), to be a noisy signal about the worker's ability that is independent of the worker's sector during period  $t$ . Relaxing this assumption about the functional form of (2) would complicate the analysis because workers' sector choices would then depend on the benefit from faster learning, as well as on the benefit from increased expected output given current beliefs. Relaxing the assumption that all labor market participants observe  $Z_i$  (so that there could be learning about both  $Z_i$  and  $\eta_i$ ) would cause similar complications. *Under our assumptions, the posterior distribution of  $\eta_i$  given the history  $s_{it}$  is normal with mean  $m_{it}$  given by (4) and precision  $h_t = h + th_\varepsilon$ , regardless of the worker's history of sector affiliations.*

In this fourth model, with learning and comparative advantage, we finally have an internally consistent account for sector affiliation, wages, sector mobility, and wage changes, as follows. In each period, firms in sector  $j$  bid a worker's wage up to the worker's expected output in that sector, conditional on the publicly observable information about the worker available at that date:

$$\ln w_{ijt} = X_{it}\beta_j + Z_i + b_jm_{i,t-1} + c_j + (1/2)b_j^2\sigma_t^2 \quad (10)$$

which is analogous to (6) but with sector-specific returns  $\beta_j$  and  $b_j$ . The model also includes sector-specific effects since the posterior variance  $\sigma_t^2$ , which declines with time (labor market experience), is interacted with  $b_j^2$ . The worker chooses to work in the sector offering the highest current wage. Thus, worker  $i$  chooses sector  $j$  in period  $t + 1$  if  $v_{j-1}(X_{i,t+1}) < m_{it} < v_j(X_{i,t+1})$ . In all of our models,

<sup>1</sup>In this model, the sector-specific intercepts are now  $c_j$  plus the term  $(1/2)b_j^2\sigma_\varepsilon^2$ . This additional term accounts for the fact that when wages are paid in advance, differences in the variance of productivity across sectors (due to differences in the returns to skill  $b_j$ ) lead to systematic differences in mean log wages because of the log normal transformation mentioned earlier. We view this as just one among several other possible sources of systematic wages differences across sectors. Other possible factors include compensating differentials, efficiency wages, and rents.



including this richest one, if the parameters  $\{\beta_j, b_j, c_j; j = 1, \dots, J\}$  and the measured characteristics  $X_{it}$  take on certain values, then one or more sectors may lie below the upper envelope in Figure 1 for all values of the unmeasured characteristics, in which case no workers with such measured characteristics should be employed in such sectors.

## 2.4. Estimation with Comparative Advantage

We are now ready to develop a nonlinear instrumental variables procedure to estimate the parameters  $\{\beta_j, b_j, c_j; j = 1, \dots, J\}$  in (9) and (10). This procedure does not rely on normality. To discuss the estimation of the model, define the sector indicators  $D_{ijt}$ , where  $D_{ijt}$  equals one if person  $i$  is employed in sector  $j$  at time  $t$  and zero otherwise.

The wage equation (10) for each sector  $j$  can then be written as a single equation where measurement error is assumed to be independent of sector affiliation:

$$\begin{aligned} \ln w_{it} = & \sum_j D_{ijt} c_j + \sum_j D_{ijt} X_{it} \beta_j + Z_i \\ & + \sum_j D_{ijt} b_j m_{i,t-1} + (1/2) \sum_j D_{ijt} b_j^2 \sigma_t^2 + \mu_{it}. \end{aligned} \quad (11)$$

Estimates of the sector slopes and intercepts  $\{\beta_j, c_j; j = 1, \dots, J\}$  obtained by estimating equation (11) with OLS are inconsistent. The problem is that expected ability influences sector affiliation, so  $m_{i,t-1}$  is correlated with the set of sector dummies  $\{D_{ijt}, j = 1, \dots, J\}$ . If the worker's ability were fixed, known, and equally valued in all sectors, then a first-differenced regression would eliminate this ability bias. But the endogeneity problem in equation (11) is different from the usual fixed-effect case for two reasons. First,  $m_{i,t-1}$  is a martingale rather than a fixed effect. This martingale property does not depend on the normality assumptions in our theoretical model; all Bayesian beliefs are martingales. In the absence of comparative advantage, we could handle this martingale problem as described in Section 2.2. But, second, comparative advantage causes  $m_{i,t}$  to be interacted with the set of sector dummies  $\{D_{ijt}, j = 1, \dots, J\}$ .

*SKIPPED FOR NOW!*

## 3. Conclusion

Broadly speaking, the results suggest that the measured occupational wage differentials in a cross-section regression are largely due to unmeasured and unobserved worker skills. We find evidence that the sorting of skills into “high-wage” occupations is explained by high returns to skills in these occupations. Although comparative advantage appears to play a fundamental role in occupational wage differences, the role of learning is more limited. One possible explanation for this finding is that though



learning may be quite important in the first few years in the labor market, it may not be as important later on.

## References

- Gibbons, Robert, Lawrence F Katz, and Thomas Lemieux (2005) “Comparative Advantage, Learning, and Sectoral Wage Determination,” *Journal of Labor Economics*, 23 (4), 681–723, [10.1086/491606](#).
- Jovanovic, Boyan (1979) “Job Matching and the Theory of Turnover,” *Journal of Political Economy*, 87 (5, Part 1), 972–990, [10.1086/260808](#).