Connecting third-degree price and labor discrimination literatures: the importance of concavity of demand/supply curves and search∗

Alexei Alexandrov†

January 20, 2017

Abstract

I show that theoretical models of third-degree price discrimination and discrimination in labor markets are identical in many cases, and I show how researchers in industrial organization and price theory can use the formal models, findings, and intuition from the labor literature, and vice versa. I analyze two applications: one applying the existing results in the industrial organization literature to a labor setting (after extending the generality of the existing results) and another applying existing results in the labor literature to an industrial organization setting. I derive conditions on when wage discrimination is inefficient when two groups of workers have different labor supply elasticities and I show that charging higher prices to consumers with higher search cost is inefficient under mild theoretical assumptions. I also discuss other potential applications.

INTRODUCTION

Welfare effects of third-degree price discrimination in industrial organization and of discrimination in labor economics had been of long-standing interest to economists. However, these two research literatures had largely been developing independently. I show that theoretical models in these two literatures are identical in many cases, and I also show how researchers in industrial organization and price theory can use formal models, findings, and intuition from the labor literature, and vice versa. I hope that this work initiates successful cross-pollination between the heretofore largely unconnected research literatures.

From a firm’s perspective, in either scenario the ability to discriminate allows the firm to tailor its offer to different markets, whether that offer is a price for its product or wage for its workers or any other characteristic.1 This tailoring increases profit and typically increases the number

∗Keywords: price discrimination, statistical discrimination, race discrimination, gender discrimination, search.
JEL codes: D40, L11, J70.
†Email: alexei01@gmail.com. The views expressed are those of the author and do not necessarily represent those of the author’s current or previous employers. I would like to thank the participants of EARIE conference for their comments.

1I focus on statistical discrimination (as opposed to taste-based discrimination); however, see the exercise applying the results from labor literature to industrial organization.
of transactions in the advantaged market, while decreasing the number of transactions in the disadvantaged market, to the extent that these markets are elastic. However, there are two more welfare effects.

First, discrimination dampens the incentive to invest or to search for the right match for the disadvantaged group. While this dampening of investment incentive has been the focus of the statistical discrimination literature in labor, it had not been nearly as emphasized in the third-degree price discrimination literature (see below on suppliers pricing to retailers, a setting much closer to the labor literature).

Second, discrimination shifts transactions (either on the supply or demand side) from the disadvantaged group to the advantaged group. The shift is inefficient since some unlucky consumers/workers with higher willingness to pay/accept in the disadvantaged group are passed over in favor of some lucky consumer/worker with lower willingness to pay/accept in the advantaged group. While this inefficient shift had been the focus of the industrial organization literature, it had not been nearly as emphasized in the labor literature (see below for labor discrimination search models, a setting much closer to the industrial organization literature).

I show that the two frameworks are often equivalent, and results from one could be applied to another. For example, I use the framework from the industrial organization to show when wage discrimination can be efficient, depending on the shape of the labor supply function. Then, I use the labor framework to derive consumer surplus implications of firms competing for consumers who search for the best match, both in terms of price and of an idiosyncratic vertical utility. I show that if firms can price discriminate against a group of consumers with a higher search cost, that discrimination is inefficient.

I briefly outline the progress of the two literatures immediately below, only focusing on several papers, while skipping many important contributions for conciseness. I then outline the two applications in this paper: one applying the existing results in the industrial organization literature to a labor setting (after extending the generality of the existing results) and another applying existing results in the labor literature to an industrial organization setting. I also discuss several other applications.

Third-degree price discrimination occurs if a firm is selling to two (or more) groups of consumers that the firm can distinguish between, and rationally decides to charge the groups different prices. Assuming that the marginal costs are the same regardless of the group, difference in elasticities of demand immediately implies that the firm should charge these two groups different prices. Pigou (1920) showed that as long as demand functions are linear and all the markets remain served, then total output remains the same and such price discrimination decreases social welfare due to the aforementioned inefficient shift of transactions from the disadvantaged to the advantaged group.

---

2For example, suppose that the optimal non-discriminating price in the market is 5 dollars, while the discriminating prices are 6 dollars for group A and 4 dollars for group B. In this case, some consumers in group A valued the product at between 5 and 6 dollars, and will not buy the product if the firm discriminates. Meanwhile, some consumers in group B valued the product at between 4 and 5 dollars, and will buy the product only if the firm discriminates. Thus, discrimination leads to consumers with lower willingness-to-pay getting the product instead of consumers with higher willingness-to-pay.
Robinson (1933) showed that the curvature of demands of the two groups determines whether the total output increases, see also Schmalensee (1981). Varian (1985) and Schwartz (1990) showed that in many cases total output increase in a necessary condition for price discrimination to be socially beneficial. For price discrimination to increase social welfare, this inefficient shift has to be outweighed by a beneficial overall output expansion. Aguirre, Cowan, and Vickers (2010) provided sufficient conditions on curvatures of the demand functions that guarantee either a positive or a negative effect of discrimination.

Statistical discrimination occurs if a firm is employing from two groups of workers that the firm can distinguish between, and rationally decides to offer the groups different wage. Assuming perfectly competitive labor markets and that the two groups are apriori the same, either a difference in prior beliefs or different precisions of signals about the ability of workers that firms receive for two different groups result in rational-for-firms policies of offering different wages, see Arrow (1971) and Phelps (1972). Lundberg and Startz (1983) formalized earlier intuition that such statistical discrimination results in lower ex-ante investment incentives for the disadvantaged group, implying that a no discrimination constraint could increase social welfare by adjusting the ex-ante investment incentives, a finding resting on the convexity of the investment function of each individual. Coate and Loury (1993) go further by analyzing when affirmative action (effectively an equally likely promotion policy) can increase social welfare or, alternatively, result in a patronizing equilibrium where the disadvantaged group’s incentives are further lowered by affirmative action. See Fang and Moro (2010), Lang and Lehmann (2012), and Guryan and Charles (2013) for a review of more recent literature.

The starting point in my exercise of applying industrial organization results to labor is Aguirre, Cowan, and Vickers (2010), who show that the effects of third-degree price discrimination on social welfare and output can be often characterized by comparing the curvatures of demand curves in the two markets. I extend their findings in two ways. First, I allow the firm to endogenously choose any characteristic (not just price) to allow for discrimination in hiring, promotion, or admission policies, for example. Second, I allow consumers to endogenously alter their ex-ante investment choice based on the market prices for their group, for example investment in unobserved by the firms training. I show how the results apply in a labor setting of an upward sloping labor supply curve, where the supply curves are different for the two groups of workers. I outline sufficient conditions on the supply function and on worker investments that determine the effects of discrimination on social welfare, and as in Aguirre, Cowan, and Vickers (2010) these conditions revolve around the relative curvatures of the supply functions in the two markets, adjusted by the workers’ investment decisions. In the Appendix, I extend the model to analyze the case of many firms in the market, with the difference being that, with many firms, an individual firm’s choice to discriminate does not

---

3 The key variable does not have to be the wage: The same intuition applies to, for example, admission policies in schools and promotion policies.

4 I am agnostic on reasons behind an upward-sloping supply curve (as opposed to a perfectly competitive labor market), but the existing literature provides several reasons, for example heterogeneous match quality and search costs on the worker side, see Bhaskar, Manning, and To (2002).
affect workers’ apriori investment decisions, but the overall market equilibrium does affect workers’ investment decisions. I also analyze an extension where the two groups are not separable: the disadvantaged group is affected by the firm’s offer to the advantaged group, and vice versa. While restrictive assumptions on the shape of cross-effects must be made to utilize the same method, it is possible to derive some results even in that scenario.

This application nests many of the statistical discrimination frameworks. The special case without investment incentives fits a more static model where workers are choosing between employment and leisure, with different groups having either different outside options or employers having different beliefs regarding the groups’ abilities. This special case can also incorporate workers choosing endogenously whether to go into self-employment, similar to the one analyzed by Schwab (1986). Something akin to this special case was also analyzed by Robinson in the second edition of her book. The more involved case of incorporating ex-ante incentives can also include workers investing in their unobserved characteristics prior to entering the job market. I do not assume perfectly competitive labor markets in either case. Insights on curvature comparisons of appropriate functions suggest similar insights for when statistical discrimination is welfare enhancing. Similar arguments for increased output as a necessary condition can be made.

The starting point in my exercise of applying labor results to industrial organization is Black (1995), who analyzes an equilibrium employment model of costly search, where workers search for best match and wage, and where some of the firms engage in taste-based discrimination as in Becker (1957). I show that simple relabeling of variables sheds light on a similar issue in third-degree price discrimination, where consumers from two different groups engage in costly search for best match and price, while some firms do not like selling to one of the groups (either by not selling at all or by charging a much higher price). In this case, while the prices to the advantaged group are the same across firms, regardless of whether a firm engages in taste-based discrimination, consumers from the disadvantaged group pay more and get worse matches even when buying from firms that do not engage in taste-based discrimination. I continue by analyzing the case where no firm engages in taste-based price discrimination; however, the two groups of consumers differ in search costs. In this case discrimination is profit-maximizing, but welfare-reducing.

Discrimination based on certain characteristics can be illegal or de-facto prohibited due to the possible public relation concerns. Thus, the following discussion is either about discriminating on other characteristics or about discriminating on these characteristics when it is much harder to detect discrimination, either for law enforcement or for the public.

I analyze two applications in the text, but many more can be analyzed. For example, Salop (1977) shows that if a group of consumers has higher search costs, then a firm might find it profitable to introduce many products and price dispersion to facilitate price discrimination: consumers

---

5See Lang, Manove, and Dickens (2005) whose model does not allow firms to offer different wages to different groups, thus any effect of discrimination is through hiring preferences as opposed to wage differences once hired.

6See, for example, Bertrand and Mullainathan (2004) for employment context (documenting lower call-back rates for resumes with African-American names and see Ayres (1991) and Ayres and Siegelman (1995) documenting race and gender discrimination in car bargaining at dealerships. However, see Goldberg (1996) qualifying the earlier findings based on data from the Consumer Expenditure Survey.
with low search costs will find the cheaper products, while consumers with high search costs will end up buying the more expensive products. This intuition should be portable to labor discrimination where firms can design their hiring websites and implement hiring practices that would effectively discriminate against a group of workers if that particular group of workers’ cost of navigating through these practices is higher. Another example from industrial organization is strategic competition, with some of the competitive third-degree price discrimination literature mentioned in the Appendix. Non-compete clauses in labor contracts should be ripe for analyzing using those techniques.

An example from labor literature that could be useful to researchers in industrial organization is the analysis of discrimination when there are complementarities between the two types of workers in the firm’s production function, see Moro and Norman (2004). These results might advance the industrial organization literature on wholesalers/manufacturers discriminating against retailers/downstream firms, with clear complementarities or substitutabilities due to competition between retailers/downstream firms for consumers. In particular, these results could be useful in analyzing most-favored nation clauses that prohibit manufacturers from charging different prices to different retailers or different prices at different marketplaces. I hope that the two applications in the text and this discussion convince researchers that there is much to be learned from the literature in the other field (even for the purposes of advancing one’s own field) and that there might be fruitful areas of potential collaboration.

APPLYING INDUSTRIAL ORGANIZATION RESULTS TO LABOR

Model

I outline a simple model of a monopolist firm that is offering its product (or hiring from) two different markets or two different groups of workers. In the text below, I made several important simplifying assumptions. I relax some of these assumptions in the Appendix, at least partially. In particular, (1) I analyze the case where there are many different firms so that while no firm’s choices can affect consumer/worker investments, the investments are affected by the market equilibrium, (2) I analyze the case and show which simplifying assumptions need to be made for the method to apply when the two markets/groups of workers are interconnected, (3) I discuss the presence of endogenous variables that the firm cannot discriminate with, and (4) I discuss incorporating strategic competition. I focus on social welfare throughout this Section; however the results can be easily adjusted to focus on output or on consumer surplus, see Cowan (2012) for the latter.

Assume that a firm is dealing with two markets, weak and strong. The firm can choose a characteristic, \( x \), for each of the markets. Based on the characteristic chosen for their market, consumers/workers in that market make an investment decision, \( v(x) \), to maximize consumer welfare (or any other function).\(^7\) This investment decision might then affect the firm’s profit. The firm’s

\(^7\)I am agnostic regarding the microeconomic foundations of \( v(x) \).
overall profit function is thus:

\[ \pi(x_w, x_s) = \pi_w(x_w, v^*_w(x_w)) + \pi_s(x_s, v^*_s(x_s)), \]  

(1)

where \(v^*_i(x_i)\) denotes the optimal choice of \(v\) by consumers.

I am assuming that the two markets are effectively independent of each other on consumer side: the consumers/workers in the weak market do not care about \(x_s\) and the consumers/workers in the strong market do not care about \(x_w\). The firm chooses \(x_w\) and \(x_s\) to maximize overall profit. WLOG, \(x^*_i = x^*_w + r^*\), where \(r^* \geq 0\). Finally, I assume that the firm’s optimization is well-defined (profit is concave) – the assumption that includes the firm taking into account consumers’ response to the characteristic choice, \(v^*(x)\). The assumption also implies that profit is concave separately in each market.

Suppose, as suggested by Schmalensee (1981), that there is an exogenously imposed binding constraint, \(x_s \leq x_w + r\): the firm cannot choose characteristics that differ by more than \(r \in [0, r^*]\). In that case, \(x_s = x_w + r\). The two limiting scenarios are \(r = 0\) (meaning no discrimination is allowed) and \(r = r^*\) (the firm is allowed to do no more discrimination than the firm would choose to do optimally, meaning the constraint is not binding). The firm’s profit function becomes:

\[ \pi(x_w, x_s) = \pi_w(x_w, v^*_w(x_w)) + \pi_s(x_w + r, v^*_s(x_w + r)). \]  

(2)

We can also redefine the firm’s profit function to directly analyze the firm’s profit after consumers’ investment choice:

\[ \Pi(x_w, x_s) = \Pi_w(x_w) + \Pi_s(x_w + r), \]  

(3)

where \(\Pi_i(x_i) = \pi_i(x_i, v^*_i(x_i))\). At this point we can analyze how the firm changes its optimal choice of \(x_w\) in response to changes in \(r\) (effectively the extent of discrimination). The optimal choice of \(x_w\) is determined by the firm’s first-order condition (FOC): \(\Pi'_w(x_w) + \Pi'_s(x_w + r) = 0\). Totally differentiating the FOC (and dropping the arguments that are clear):

\[ \frac{\partial x_w}{\partial r} = \frac{-\Pi''_s}{\Pi'_w + \Pi'_s} < 0. \]  

(4)

Second derivatives are negative via second-order conditions, and that guarantees that allowing more discrimination results in an even lower characteristic \(x_w\) in the weak market. Since \(x_s = x_w + r\),

\[ \frac{\partial x_s}{\partial r} = \frac{\partial x_w}{\partial r} + 1 = \frac{\Pi''_w}{\Pi'_w + \Pi'_w} > 0. \]  

(5)

In other words, allowing more discrimination results in an even higher characteristic \(x_s\) in the stronger market.

Define \(W\) as the social welfare function. We can then analyze the response of \(W\) to changes in
Define, dropping unnecessary arguments and subscripts,
\[ z(x) \equiv \frac{W'}{\Pi'}. \]  
(7)  

Then,
\[ W'(r) = \frac{\Pi'\Pi''}{(\Pi' + \Pi'')} (z'(x_w(r)) - z'(x_s(r))). \]  
(8)  

As in Aguirre, Cowan, and Vickers (2010), the following assumption, Generalized Increasing Ratio Condition (GIRC), simplifies the analysis considerably.

**Assumption 1 (Generalized Increasing Ratio Condition)** Function (ratio) \( z(x) \equiv \frac{W'}{\Pi'} \) increases in each of the two markets.

**Proposition 1** Suppose GIRC is satisfied. Then social welfare is strictly quasi-concave in \( r \), and

- if the ratio is higher in the stronger market at non-discriminatory prices \( (z_s(x_w(0)) \geq z_w(x_w(0))) \), then allowing additional discrimination reduces social welfare (regardless of the starting point),
- if the ratio is higher in the weaker market at the optimal (for the firm) discriminatory price \( (z_s(x_s^*) \leq z_w(x_w^*)) \), then allowing additional discrimination increases social welfare (regardless of the starting point), and
- if neither of the two conditions is satisfied, then social welfare first increases in the amount of discrimination allowed (some discrimination is better than none) and then decreases (the optimal amount of discrimination from the firm’s point of view is too much discrimination).

**Proof.** It is trivial to derive:
\[ W''(r) = \frac{\Pi''\Pi'' - \Pi'\Pi'''}{(\Pi' + \Pi'')} (z'(x_w(r)) - z'(x_s(r))) + (z_w - z_s) \frac{\partial}{\partial r} \frac{\Pi''\Pi''}{(\Pi' + \Pi')} . \]  
(9)  

If there exists an \( \hat{r} \) such that \( W'(\hat{r}) = 0 \), then \( W''(\hat{r}) < 0 \): the derivatives of \( z \) are positive by GIRC, the derivatives of \( x_w \) and \( x_s \) were signed above, and \( W'(\hat{r}) = 0 \) implies that \( z_w(x_w(\hat{r})) = z_s(x_s(\hat{r})) \).

Thus, if \( W'(0) < 0 \), then any discrimination reduces welfare (item 1). If \( W'(r^*) > 0 \), then any discrimination (at least up to what the firm would choose itself) increases welfare (item 2). Finally, if neither of the two is satisfied, then there is an optimal level of discrimination that is strictly between no discrimination \( (r = 0) \) and as much discrimination as the firm would want \( (r = r^*) \).
notable difference in that case is that if neither of the two conditions discussed in the Proposition are satisfied (third bullet), then social welfare first decreases and then increases, but the increasing part might never increase above the no discrimination point. In other words, if $z$ is decreasing, then the optimal policy could be, depending on the exact functional form, either to prohibit the firm from discriminating at all or to allow the firm to discriminate as much as it wants to, but intermediate solutions are suboptimal.

The assumption of $z$ being monotone allows one to apply Proposition 1 and let the two conditions in that Proposition guide the outcome of the model. For some models, the assumption on monotone $z$ might be easy to justify, see the example below. For others, this assumption might be harder to justify.

The usefulness of this approach stems from two components: (1) $W'(r)$ can be expressed as the difference of $z_w$ and $z_s$ (multiplied by a term of a known sign), where $z_i$ is a function of only the variables and functions of market $i$ and (2) the assumption of monotone $z_i$ is reasonable. While component (1) is easy to accomplish in this version of the model, I show in the Appendix that it becomes nontrivial once the model allows for interactions between the markets, effectively $\pi_w(x_w, x_s, v^*_w(x_w, x_s))$. Component (2) depends on the application in question. I discuss an application directly below.

Example

Consider a monopsonist firm hiring labor. Each worker’s marginal productivity (value to the firm) is $P$. The firm is facing two groups of workers, w and s, and can offer these groups of workers different wages (compensation) across groups: $c_w$ and $c_s$. Each group of workers has an upward-sloping supply curve, where the supply could be upward sloping for any reason: heterogeneous matches, travel costs, search costs, or any other reason. Moreover, based on the compensation offered to the group of the worker, workers might choose to engage in other occupations instead of entering this market: self-employment, leisure, or other industries. Thus, the supply curves are $q_w(c_w, v_w(c_w))$ and $q_s(c_s, v_s(c_s))$, where $v_i$ is a function of sorting out of the market for group $i$. I analyze social welfare of this market, thus I am assuming that $v$ is not costly over and above changing supply in this market. In this case, the firm’s profit function is below, and a simple renaming of variables makes it clear that this is exactly the profit function in (2).

$$\Pi(c_w, c_s) = q_w(c_w, v_w(c_w))(P - c_w) + q_s(c_s, v_s(c_s))(P - c_s).$$ (10)

Further, denote $Q_i(c_i) \equiv q_i(c_i, v_i(c_i))$ – firm’s supply taking into account the workers’ choices to engage in outside activities. Denote $L_i(c_i) \equiv \frac{P - c_i}{c_i}$ – the equivalent of the Lerner index, the markup the firm charges over the workers’ compensation. Denote $\alpha_i(c_i) \equiv \frac{c_i Q''_i(c_i)}{Q'_i(c_i)}$ – the curvature of the supply function, analogous to relative risk aversion. Then,

$$z_i = \frac{W'_i}{\pi'_i} = \frac{P - c_i}{2 - L_i \alpha_i}. \quad (11)$$
In an analogous setting, Aguirre, Cowan, and Vickers (2010) show that \( z_i \) increasing is equivalent to the slope of supply \( Q' \) being log-concave, see Bagnoli and Bergstrom (2005) for more on log-concavity. The list of functional forms of supply functions for which this assumption is satisfied includes linear, probit (normal distribution), and supply derived from extreme value and logistic distributions. I assume that \( z \) is increasing for each market for the rest of this subsection.

The following two corollaries are Propositions 1 and 2 of Aguirre, Cowan, and Vickers (2010) applied to this setting and follow directly from Proposition 1 above. These corollaries present sufficient conditions for any amount of discrimination to reduce or to increase welfare.

**Corollary 1** If the supply function in the strong market is at least as convex as that in the weak market at the nondiscriminatory price \( (\alpha_s(r = 0) \geq \alpha_w(r = 0)) \) then discrimination reduces welfare.

**Corollary 2** If \( z_w(c^*_w) \geq z_s(c^*_s) \) then discrimination increases welfare.

As Aguirre, Cowan, and Vickers (2010) note, if the profit-maximizing discriminatory worker compensations are relatively close, then condition \( z_w(c^*_w) \geq z_s(c^*_s) \) is equivalent to \( L_w \alpha_w = \frac{Q_w Q''_w}{(Q'_w)^2} \geq \frac{Q_s Q''_s}{(Q'_s)^2} = L_s \alpha_s \). That’s another condition on curvature, except that here the weak market has to be sufficiently convex.

Weyl and Fabinger (2013) reinterpret these results in terms of the pass-through rate instead of convexity of demand since the two are intricately related. In this case, the relevant pass-through rate is of changes in marginal productivity of a worker onto the worker’s compensation, and it can also be used as a sufficient statistic to analyze effects of discrimination in wages.

**APPLYING LABOR RESULTS TO INDUSTRIAL ORGANIZATION**

**Labor search model where some firms practice taste-based discrimination**

Less work is required to adopt at least some of the models from the labor literature to industrial organization questions. Here I utilize the model of Black (1995), see also the model’s discussion in Section 5.1 of Lang and Lehmann (2012). To make the example more concrete, I use car dealers as the market, according to Ayres (1991) and Ayres and Siegelman (1995). I am not aware of whether the phenomena described in these two papers still hold, and thus this example is purely for illustration.

Consider a market of car dealers, selling cars to black and white customers. Some car dealers \( (\theta) \) engage in taste-based discrimination as in Becker (1957). In this simplified version of the model, I assume that these car dealers simply won’t sell a car to black customers or, equivalently, charge such a high price that black customers do not want to purchase from these dealers. The model can be extended for taste-based discriminators simply charging black customers more, to be more in line with Becker (1957).

All other car dealers \( (1 - \theta) \) do not engage in taste-based discrimination. However, they are aware that there are car dealers that do. The \( (1 - \theta) \) car dealers set two prices, one for black customers the other for white customers. Dealers are symmetric, and all cars cost all dealers \( C \).
Customers pay search cost $\kappa$ to visit a dealership and learn about the match quality of the car, $\alpha$. While at the dealership, customers also find out the price that they will be charged, depending on their race, or black customers might find out that they’re visiting one of the dealerships engaged in taste-based discrimination. Customers’ utility from buying at a given dealership is $\alpha - p^i$, where $i$ is customer’s race. Customers’ outside utility is normalized to 0 and they can either buy one car or not buy at all (unit demand). In equilibrium, each consumer ends up buying a car. The match quality is random and symmetric across dealers, with the distribution of $F(\alpha)$ for a particular customer-dealer combination. Let $F$ be strictly log-concave.

Customers engage in non-directed sequential search. Black (1995) shows that a customer’s reservation value $V^i$ – the utility from a match that the customer is indifferent between searching one more dealership or buying a car at this one – can be expressed as

$$\kappa = \int_{V_w + p^w}^{\infty} (\alpha - p^w - V^w) f(\alpha) d\alpha \quad (12a)$$

$$\frac{\kappa}{1 - \theta} = \int_{V_b + p^b}^{\infty} (\alpha - p^b - V^b) f(\alpha) d\alpha. \quad (12b)$$

Effectively, search is more costly for black customers since with probability $\theta$ they encounter a dealer practicing taste-based discrimination.

**Result 2** For a given price, black customers accept offers with lower match value ($\alpha$). In other words, black customers’ reservation value is lower $V^b < V^w$.

On the dealer side, it is easy to derive that white customers receive the same price offer regardless of whether the dealer engages in taste-based discrimination. Dealers not engaging in taste-based discrimination maximize their profit – the probability of acceptance multiplied by the markup:

$$\pi^i = (1 - F(V^i + p^i))(p^i - C), \quad (13)$$

for $i = b, w$.

**Result 3** Black customers pay a higher price and experience worse match quality on average.

**Result 4** If the number of dealers in the market is endogenous, then the higher is the proportion of black customers, the fewer dealers that engage in taste-based discrimination enter the market.

**Labor search model without taste-based discrimination**

The same model as above can be applied in the case without taste-based discrimination. Suppose that no firm practices taste-based discrimination. However, suppose that two groups of consumers have different search costs, $\kappa_b > \kappa_w$. In this case, a customer’s reservation value is again

$$\kappa_i = \int_{V_i + p^i}^{\infty} (\alpha - p^i - V^i) f(\alpha) d\alpha, \quad (14)$$
for $i = b, w$.

Firms maximize their profit as above

$$\pi^i = (1 - F(V^i + p^i))(p^i - C), \quad (15)$$

for $i = b, w$.

Implicitly differentiating equation (14), it is clear that higher search cost implies lower reservation value, $V$. Lower reservation value, in turn, implies higher equilibrium price.

**Lemma 1** The group with higher search cost has lower reservation value, faces higher equilibrium prices, and searches less.

Two considerations affect social welfare. First, every customer buys in equilibrium, thus the effect of prices is only relevant to the extent that it leads to more search. Second, more search is costly, but it also improves customers’ matches.

**Proposition 2** Higher equilibrium prices for the group with higher search costs are welfare-reducing.

**Proof.** The efficient level of customers’ reservation utility, given that every customer makes a purchase, is such that customers do not consider price: consumers simply search for the best match. In other words,

$$\kappa = \int_{V^*}^{\infty} (\alpha - V^*) f(\alpha) d\alpha. \quad (16)$$

Introduction of a positive price, say

$$\kappa = \int_{V_\delta + \delta}^{\infty} (\alpha - (V_\delta + \delta)) f(\alpha) d\alpha, \quad (17)$$

where $\delta > 0$, leads to a lower and suboptimal level of search: $V_\delta < V^*$ or $\frac{\partial V_\delta}{\partial \delta} < 0$. Therefore, an even higher price for black customers due to their higher search cost lowers these customers’ reservation utility even more, and is thus inefficient. ■

This model can be adjusted for other asymmetries. For example, the two groups of customers could believe that they face different distributions $F_i(\bullet)$, regardless of whether this perception is correct.

This result might no longer hold once the model is closer to the real world with some customers not making any purchases in equilibrium. In order to model the effects of inelastic demand and the effects of discrimination on social welfare in that case, the models in this section can be connected to the model in Section 2. Profit in (13) can be expressed as

$$\pi^i = d_i(p^i, V^i(p^i))(p^i - C), \quad (18)$$

where $d_i(p^i, V^i(p^i)) = 1 - F(V^i(p^i) + p^i)$. 

11
Other findings from search literature that apply both to prices and wages suggest that results of the introduction of a no discrimination requirement or a softer cap on the allowed discrimination might be nontrivial. See, for example, Fershtman and Fishman (1994) and Janssen and Moraga-González (2004), both using the workhorse model of Burdett and Judd (1983).

CONCLUSION

I outlined how results on discrimination from industrial organization can be applied to issues in the labor literature, and vice versa. I hope that this insight will allow researchers from either field to import formal models, results, and intuition from the other field resulting in new insights in both fields. The applications in the text also highlight the importance of curvature of supply curves for labor and the importance of search costs for price discrimination.

References


APPENDIX: EXTENSIONS OF APPLYING INDUSTRIAL ORGANIZATION RESULTS TO LABOR

I analyze and discuss various extensions to the model used in applying industrial organization results to labor in Section 2. I analyze two extensions in detail and then I discuss two more extensions less formally.

The first extension that I discuss in detail is analyzing a firm that does not internalize consumer and worker investment, for example due to being a part of a competitive market. However, I assume that the firm’s optimal choice still affects the actual investment (as the competitive equilibrium would, despite no firm having an effect by itself).

The second extension that I discuss in detail is analyzing interactions between the two groups of consumers (or workers). For example, there could be network effects between different groups of consumers or complementarities between different groups of workers.

I then discuss two more extensions less formally: (1) presence of another endogenous variable that the firm cannot discriminate on (for example, the airplane will arrive at the same time for both business and economy class customers) and (2) strategic competition between firms.

**Firm not internalizing consumer/worker investment**

One of the assumptions in the base model was that the firm internalizes the fact that consumers’/workers’ investment, $v(x)$, might change due to the firm’s choices of $x_w$ and $x_s$. However, this does not have to be the case: the firm might be myopic or, if the choices are wages or promotion policies, and the firm is far from being the only employer in town, then it is rational for the manager to assume that the firm’s choices are not going to influence workers’ ex-ante investments in, say, education. Many of the earlier works in the labor literature make this assumption either explicitly or implicitly.

In this subsection I assume that the firm does not internalize consumers’/workers’ investment choice $v(x)$. However, I continue to assume that the firm’s profit depends on $v$: thus, while the firm does not internalize the connection between $v(x)$ and the firm’s choice of $x$, the firm does optimally adjust to the level of $v(x)$. This could be thought of as a reduced-form approach to modeling many symmetric firms: none of their policies matter unilaterally, but the optimal choice in the symmetric equilibrium does affect everyone. Alternatively, this could be self-fulfilling expectations phenomenon: a firm chooses its policies without realizing that the workers will optimally react with their investment choices, and after workers react the firm’s choice is still optimal.

Keeping the notation from the base model, the firm’s FOC in this case is

$$\frac{\partial \pi_w(x_w, v^*_w(x_w))}{\partial x_w} + \frac{\partial \pi_s(x_w + r, v^*_w(x_w + r))}{\partial x_w} = 0.$$  

(19)
The difference with the FOC in the base model is the firm not accounting for $\frac{\partial \pi_w}{\partial v_i} \frac{\partial \pi_w}{\partial x}$ in this case. Similarly as in the base model, we can totally differentiate the FOC to derive the optimal responses of the firm’s choices of $x$ to changes in the amount of discrimination allowed ($r$).

$$
\left(\frac{\partial^2 \pi_w}{\partial x_w^2} + \frac{\partial^2 \pi_s}{\partial x_w^2}\right) \frac{\partial x_w}{\partial r} + \frac{\partial \pi_w}{\partial x_w} \frac{\partial v_w}{\partial r} + \frac{\partial \pi_s}{\partial x_w} \frac{\partial v_s}{\partial r} + \frac{\partial^2 \pi_s}{\partial x_s^2} \frac{\partial x_s}{\partial r} + \frac{\partial \pi_s}{\partial x_w} \frac{\partial v_s}{\partial r} = 0. \quad (20)
$$

Note that the firm takes into account the fact that the investments $v_i$ change (the cross-partial terms), but the firm believes that its own policies have no impact on the investments (the lack of the cross-partial terms in the FOC). To close the loop, I am assuming that the firm rationally expects that in equilibrium the investment adjustment is according to the equilibrium $x$ values, in other words that $\frac{\partial \pi_s}{\partial r} = \frac{\partial \pi_s}{\partial x_i} \frac{\partial x_i}{\partial r}$. Again, this is consistent with either a symmetric equilibrium where no one firm’s policy matters for investment purposes or a steady-state equilibrium of a myopic firm whose policies do affect the investments, but the firm does not realize that. This assumption leads to

$$
\frac{\partial x_w}{\partial r} = \frac{-\frac{\partial^2 \pi_s}{\partial x_w^2} - \frac{\partial \pi_s}{\partial x_w} \frac{\partial v_s}{\partial r} - \frac{\partial \pi_s}{\partial x_w} \frac{\partial v_s}{\partial r}}{\frac{\partial^2 \pi_s}{\partial x_w^2} + \frac{\partial \pi_s}{\partial x_w} \frac{\partial v_s}{\partial r} + \frac{\partial \pi_s}{\partial x_w} \frac{\partial v_s}{\partial r}} < 0. \quad (21)
$$

Note that the condition on profit concavity has to be adjusted for this case: $\frac{\partial^2 \pi_s}{\partial x_i^2} + \frac{\partial \pi_s}{\partial x_i} \frac{\partial v_s}{\partial x_i} < 0$ instead of $\Pi'_i(x_i) < 0$. Similarly to the derivation above,

$$
\frac{\partial x_s}{\partial r} = \frac{\partial x_w}{\partial r} + 1 = \frac{-\frac{\partial^2 \pi_s}{\partial x_s^2} - \frac{\partial \pi_s}{\partial x_s} \frac{\partial v_s}{\partial r}}{\frac{\partial^2 \pi_s}{\partial x_s^2} + \frac{\partial \pi_s}{\partial x_s} \frac{\partial v_s}{\partial r}} > 0. \quad (22)
$$

Again, as in the base model,

$$
W'(r) = W'_w = \frac{-\frac{\partial^2 \pi_s}{\partial x_w^2} - \frac{\partial \pi_s}{\partial x_w} \frac{\partial v_s}{\partial r}}{\frac{\partial^2 \pi_s}{\partial x_w^2} + \frac{\partial \pi_s}{\partial x_w} \frac{\partial v_s}{\partial r}} + W'_s = \frac{-\frac{\partial^2 \pi_s}{\partial x_s^2} - \frac{\partial \pi_s}{\partial x_s} \frac{\partial v_s}{\partial r}}{\frac{\partial^2 \pi_s}{\partial x_s^2} + \frac{\partial \pi_s}{\partial x_s} \frac{\partial v_s}{\partial r}}. \quad (23)
$$

We can define a slightly modified ratio function:

$$
z_{I-NotInternalized} = \frac{W'_i}{\frac{\partial^2 \pi_i}{\partial x_i^2} + \frac{\partial \pi_i}{\partial x_i} \frac{\partial v_i}{\partial x_i}}. \quad (24)
$$

Using this modified ratio function,

$$
W'(r) = \left(\frac{\frac{\partial^2 \pi_w}{\partial x_w^2} + \frac{\partial \pi_w}{\partial x_w} \frac{\partial v_w}{\partial x_w}}{\frac{\partial^2 \pi_w}{\partial x_w^2} + \frac{\partial \pi_w}{\partial x_w} \frac{\partial v_w}{\partial x_w}} + \frac{\frac{\partial^2 \pi_s}{\partial x_s^2} + \frac{\partial \pi_s}{\partial x_s} \frac{\partial v_s}{\partial x_s}}{\frac{\partial^2 \pi_s}{\partial x_s^2} + \frac{\partial \pi_s}{\partial x_s} \frac{\partial v_s}{\partial x_s}}\right) \left(z_{I-NotInternalized}(x_w(r)) - z_{I-NotInternalized}(x_s(r))\right). \quad (25)
$$

**Corollary 3** Proposition 1 applies in this setting when assumptions on $z(*)$ are substituted by
same assumptions on $z_{\text{NotInternalized}}(\bullet)$.

The setup of the base model is advantageous because it allows to transform $W'(r)$ into a product of two factors: a factor that is always positive and a difference of $z$, with each $z$ being a function of only variables in its market.

**Interaction between the markets**

Earlier statistical discrimination literature assumed away any complementarity or substitutability between the two groups, for example Coate and Loury (1993). Later work includes these interaction, for example Moro and Norman (2004) who discuss complementarities between the two groups (and the lack of modeling these interactions in virtually all previous work). These interaction effects can come through various sources, including workers competing for the same position, in contrast to the earlier literature where there is an unlimited number of positions.\(^8\)

Similarly, in the industrial organization literature, there is significant literature on third-degree price discrimination in the input markets: discrimination done by manufacturers/wholesalers when establishing wholesale prices to groups of retailers who ultimately sell to consumers, see for example Katz (1987) and DeGraba (1990). One of the major differences in this literature is that the retailers that are being discriminated against compete for consumers in the downstream market, and thus a particular retailer cares not only about its wholesale price, but also about the wholesale price that the retailers in the other group pay. This literature also includes discussions on the ‘waterbed effects,’ where a preferred retailer getting a better deal implies that the other retailers will have to pay more due to that, see Inderst and Valletti (2011), as well as the literature on incentives to invest based on whether price discrimination is allowed in contexts such as net neutrality, see Choi and Kim (2010). See, for example, O’Brien and Shaffer (1994) for how nonlinear pricing might affect these dynamics.

In this subsection discuss potential additional assumptions on cross-effects between the markets that could allow using similar techniques as above in this setting. I assume that the characteristic in one market, $x_i$, has a **direct** effect on the profit from the other market, $\pi_j$, resulting in the following profit function:

$$\pi(x_w, x_s) = \pi_w(x_w, x_s, v_w^*(x_w, x_s)) + \pi_s(x_s, x_w, v_s^*(x_s, x_w)).$$  

(26)

As in the base model above, we can redefine the firm’s profit function to directly analyze the firm’s profit after consumers’ investment choice, also plugging in $x_s = x_w + r$:

$$\Pi(x_w, x_s) = \Pi_w(x_w, x_w + r) + \Pi_s(x_w + r, x_w),$$

(27)

\(^8\)In Norman (2003) complementarities arise due to the social planner’s budget constraint. In Lang, Manove, and Dickens (2005) the interaction between groups comes from the firms hiring only one candidate, and thus the members of the group who are discriminated against do not want to apply for the same job postings as the members of the non-discriminated group.
where \( \Pi_i(x_i) = \pi_i(x_i, x_j, v^i(x_i, x_j)) \). At this point we can analyze how the firm changes its optimal choice of \( x_w \) in response to changes in \( r \) (effectively the extent of discrimination). Analogously to the derivations above, by totally differentiating the firm’s first-order condition,

\[
\frac{\partial x_w}{\partial r} = -\left(\frac{\Pi_{ij}^1 + \Pi_{ij}^2 + \Pi_{ij}^{22}}{\Pi_{ij}^1 + 2\Pi_{ij}^2 + \Pi_{ij}^{22} + \Pi_{ij}^1 + 2\Pi_{ij}^{22} + \Pi_{ij}^{22}}\right) < 0, \tag{28}
\]

where \( \Pi^{ij} \) denotes the derivative of \( \Pi \) with respect to \( x_i \) and \( x_j \) (so cross-partial if \( i \) and \( j \) are different, and second derivative if they are the same). As in the base model, I assume that the firm’s profit maximization second-order conditions are satisfied both for the firm as a whole, as well as in any given market. These assumptions are \( \Pi_{ij}^1 + \Pi_{ij}^2 + \Pi_{ij}^{22} < 0 \) and \( SOC_i \equiv \Pi_{ij}^1 + 2\Pi_{ij}^2 + \Pi_{ij}^{22} < 0 \). These assumptions imply that \( \frac{\partial^2 x_w}{\partial r^2} < 0 \). We can also derive the change in other characteristic with respect to the level of discrimination allowed:

\[
\frac{\partial x_s}{\partial r} = \frac{\partial x_w}{\partial r} + 1 = \frac{\Pi_{ij}^1 + \Pi_{ij}^2 + \Pi_{ij}^{22}}{SOC_i + SOC_w} > 0. \tag{29}
\]

Denote \( \Delta \equiv \Pi_{ij}^2 + \Pi_{ij}^{22} - (\Pi_{ij}^1 + \Pi_{ij}^{22}) \). Then,

\[
\frac{\partial x_w}{\partial r} = -\left(\frac{\Pi_{ij}^1 + \Pi_{ij}^2 + \Pi_{ij}^{22}}{\Pi_{ij}^1 + 2\Pi_{ij}^2 + \Pi_{ij}^{22} + \Pi_{ij}^1 + 2\Pi_{ij}^{22} + \Pi_{ij}^{22}}\right) \frac{(SOC_s + \Delta)}{SOC_w + SOC_s} < 0, \tag{30}
\]

and

\[
\frac{\partial x_s}{\partial r} = \frac{\Pi_{ij}^1 + \Pi_{ij}^2 + \Pi_{ij}^{22}}{SOC_i + SOC_w} = \frac{SOC_w - \Delta}{SOC_w + SOC_s} > 0. \tag{31}
\]

We can then analyze the response of \( W \) to changes in \( r \) (effectively the extent of discrimination):

\[
W'(r) = W_1^w(x_w, x_s) \frac{\partial x_w}{\partial r} + W_2^w(x_w, x_s) \frac{\partial x_s}{\partial r} + W_1^s(x_s, x_w) \frac{\partial x_s}{\partial r} + W_2^s(x_s, x_w) \frac{\partial x_w}{\partial r} =
\]

\[
= W_1^w \frac{(SOC_s + \Delta)}{SOC_w + SOC_s} + W_2^w \frac{(SOC_w - \Delta)}{SOC_w + SOC_s} + W_1^s \frac{(SOC_w - \Delta)}{SOC_w + SOC_s} + W_2^s \frac{(SOC_s + \Delta)}{SOC_w + SOC_s} =
\]

\[
= \frac{(SOC_s + \Delta)(SOC_w - \Delta)}{SOC_w + SOC_s} \left( \frac{W_1^w}{SOC_w - \Delta} - \frac{W_2^w}{SOC_w + \Delta} - \frac{W_1^s}{SOC_s + \Delta} + \frac{W_2^s}{SOC_s - \Delta} \right). \tag{32}
\]

We can define a slightly modified ratio function (re-arrange the second and the fourth terms in the parentheses):

\[
z_{w-MarketInteraction} = \frac{W_1^w + W_2^s}{SOC_w - \Delta}, \tag{33}
\]

with a similar definition for the strong market – the only difference being the change in the sign in front of \( \Delta \). This results in

\[
W'(r) = \frac{(SOC_s + \Delta)(SOC_w - \Delta)}{SOC_w + SOC_s} \left( z_{w-MarketInteraction}(x_w(r)) - z_{s-MarketInteraction}(x_s(r)) \right). \tag{34}
\]
Corollary 4  Proposition 1 applies in this setting when assumptions on \( z(\cdot) \) are substituted by same assumptions on \( z_{MarketInteraction}(\cdot) \).

As noted above, the advantage of the method proposed by Aguirre, Cowan, and Vickers (2010) is breaking up expression \( W'(r) \) into a difference of two functions, where each function is a function of variables of only one of the markets. Of course, in this example each \( z_i \) is also a function of \( x_j \), however, that’s not a concern since \( x_s = x_w + r \), thus a substitution can always be made. However, there are still two issues in \( z_{w-MarketInteraction} \) (and corresponding issues for the strong market): \( \Delta \) is a function of derivatives of \( \Pi_s \) and the numerator has \( W^2_s \).

Given the definition of \( \Delta \equiv \Pi^{12}_w + \Pi^{22}_w - (\Pi^{12}_s + \Pi^{22}_s) \), as second-order cross-effects, it is plausible that they are either second-order in comparison to \( SOC_w \) or that with particular functional forms they can be expressed as a tractable function of \( x_w \) without the derivatives of \( \Pi_s \). A particularly convenient assumption would be that these effects are symmetric: \( \Pi^{12}_w + \Pi^{22}_w = \Pi^{12}_s + \Pi^{22}_s \), and thus \( \Delta = 0 \), completely resolving this problem. This is likely to be a bad approximation for some problems, but this can be a useful starting point of analysis.

The other issue is \( W^2_s \) in the numerator. Of course, a similar symmetric effects assumption can be made: \( W^2_s = W^2_w \). Alternatively, we could define \( \delta \equiv W^2_w - W^2_s \), in which case the discussion regarding \( \delta \) is the same as the one in the previous paragraph regarding \( \Delta \).

In the end, the usefulness of the approach taken by Aguirre, Cowan, and Vickers (2010) hinges on \( z \) being monotone. Depending on the exact application, assuming that \( z \) is monotone might not be too binding in this context, in which case Proposition 1 provides a useful characterization of welfare effects.

Discussion of two other extensions

Suppose that there is another endogenous variable that the firm can choose, but the firm cannot discriminate on that variable, so the profit function is \( \pi(x_w, x_s, k) = \pi_w(x_w, k, v^*_w(x_w, k)) + \pi_s(x_s, k, v^*_s(x_s, k)) \). The most stark example is whether the firm operates at all (sinks in the fixed cost of being open), but there are many other examples, see Hausman and MacKie-Mason (1988), Layson (1994), and Alexandrov and Deb (2012). It is clear that in such a setting \( r \) (the exogenous constraint on discrimination in \( x \)) also affects the firm’s choice of \( k \). In particular, the conditions in this setup for \( k \) to either increase or decrease with \( r \) are straight-forward (but nonetheless tedious) to derive. One of the results of this setup is that allowing more discrimination might be Pareto improving: even the disadvantaged (strong) market that ends up paying a higher price/receiving a lower wage might be better off if the firm chooses a more appropriate \( k \) because the firm is allowed to price discriminate.

The model can also be expanded to incorporate strategic competition; however, this feature is not easy to incorporate and derive straight-forward results, see Holmes (1989), Corts (1998), Armstrong and Vickers (2001), and see Stole (2007) for a review of the literature. Nonetheless, recent advances show that this intuition can be expanded on. Weyl and Fabinger (2013) use their
pass-through analysis to expand the results of Aguirre, Cowan, and Vickers (2010) to imperfect symmetric competition. Instead of comparing pass-through rates (equivalently, curvatures) in the two markets, these pass-through rates should be weighted by the conduct parameter in the market that effectively measures the firms’ market power. Thus, if the disadvantaged market is less competitive, then it is more likely that discrimination reduces welfare.