PRODUCT MARKET AND THE SIZE–WAGE DIFFERENTIAL*

By   Shouyong Shi[1]

Indiana University, U.S.A. and Center of Research on Economic Fluctuations and Employment (CREFE), Quebec, Canada

Using directed search to model the product market and the labor market, I show that large plants can pay higher wages to homogeneous workers and earn higher expected profit per worker than small plants, although plants are identical except size. A large plant charges a higher price for its product and compensates buyers with a higher service probability. To capture this size-related benefit, large plants try to become larger by recruiting at high wages. This size–wage differential survives labor market competition because a high wage is harder to get than a low wage. Moreover, the size–wage differential increases with the product demand when demand is initially low and falls when demand is already high.

1. INTRODUCTION

The *size–wage differential* refers to the fact that employers with more workers pay higher wages than smaller employers do to workers with the same observable skills (see Brown and Medoff, 1989). It is an important component of overall wage inequality. Changes in the size–wage differential alone account for 40 percent of the increase in the ninetieth-tenth percentile wage differential from 1963 to 1986 among U.S. manufacturing workers (see Davis and Haltiwanger, 1991). In contrast, changes in workers' observable characteristics like age, experience, and educational attainment account for 30 percent of this increase (see Juhn et al., 1993). Thus, it is important to explain why the size–wage differential exists and how it responds to market conditions.

What makes the size–wage differential difficult to explain is that large employers also earn higher profit per worker (Katz and Summers, 1989). Two explanations

are popular in the literature.[2] One is the sorting story, which argues that unobservable skills are complementary with capital and so large plants with high capital intensities sort out workers with high unobservable skills (Hamermesh, 1980). The second is the efficiency wage story (Stigler, 1962; Shapiro and Stiglitz, 1984). That is, it is more difficult for a large plant than for a small plant to monitor workers' effort and so a large plant must pay a higher wage to prevent shirking (see Katz and Summers, 1989). Both explanations are plausible, but they rely on unexplained additional differences between plants and leave most of the size–wage differential unaccounted for. In particular, Troske (1999) shows that the capital–labor ratio accounts for little of the establishment-size wage premium among manufacturing workers.[3]

In this article I abstract from all such differences among plants and focus on the product market as a source of the size–wage differential. The story is simple. Sellers (employers) post prices (wages) and agents cannot coordinate their decisions. With the coordination failure, buyers face the risk of being left out, and so they trade off price with the service probability. By providing a higher service probability, a larger seller can post a higher price and obtain a higher expected revenue per good (or per worker). With a realistic assumption that plants grow only gradually, this positive *size–revenue differential* makes a large employer more eager to fill the vacancy by posting a higher wage than a small employer. Labor market competition does not eliminate this size–wage differential because job applicants are indifferent between a high wage that is hard to get and a low wage that is easy to get. Thus, in equilibrium a larger plant can be more profitable and pay a higher wage than a smaller plant.

The size–revenue differential is necessary but not sufficient for a positive size–wage differential. There are two forces that reduce the size–wage differential. One is the difference between large and small employers' outside options. A larger employer has a higher outside option than a smaller employer because it can resort to its higher current employment when it fails to recruit additional workers. As a result, a larger employer gets a larger fraction of the match surplus than a smaller employer. The second force is the size distribution of plants. When the product demand is very low, there is very little entry by plants and so the size distribution is very thin on small plants. But there is a relatively large flow from large to small plants due to job separation. To keep their mass small in the steady-state distribution, small plants must recruit more quickly by posting a higher wage than large plants, yielding a negative size–wage differential.

[2] For these explanations and others, see Brown and Medoff (1989), Katz and Summers (1989), and Davis and Haltiwanger (1991). Note that explanations such as unionization cannot explain simultaneously a size–wage differential and a size-profit differential.
[3] Brown and Medoff (1989) have found that more than a half of the size–wage differential remains after taking into account the fixed effect of unobservable skills. Also, the size–wage differential is significant among piece-rate workers whose employers do not seem to face different monitoring costs depending on size. The size–wage premium is even larger for piece-rate workers than for standard-rate workers.

The size–wage differential is positive when the size–revenue differential dominates these negative forces. This happens when the goods demand is moderate. In contrast, when the goods market demand is sufficiently high, an additional capacity does not give a seller much advantage in attracting buyers and so the size–revenue differential is too small to dominate the difference in plants' outside options. When the goods market demand is sufficiently low, there are very few small plants and so the size distribution of plants requires a negative size–wage differential.

The size–wage differential depends on the goods demand in a hump-shaped pattern, although the size–revenue differential always falls when the goods demand increases. Starting from a very low goods demand, an increase in demand increases plants' entry sufficiently and increases the fraction of small plants in the plant distribution. This eases the negative pressure on the size–wage differential exerted by the size distribution of plants and increases the size–wage differential. When the goods demand is high, a further increase in demand reduces the size–revenue differential sufficiently to compress the size–wage differential.

A critical assumption is that plants grow gradually by recruiting one worker at a time. This assumption is valid when the costs of recruiting several workers at once, including the borrowing cost, are higher than those of recruiting the workers sequentially (see Section 6 for more discussions). With this assumption, the size–revenue differential enables a large plant to capture a larger increase in the expected revenue by hiring an additional worker than a small plant.

The main contribution of this article is to show how sellers' capacity differences affect the size–wage differential. Although others have proposed the product market power as an explanation for the size–wage differential (e.g., Weiss, 1966), they have not formally shown how plants that differ only in size can differ significantly in the revenue per worker when there are many plants in the industry. Nor have they shown how the size–wage differential can survive the competition among workers and plants. I provide answers to these questions and suggest new measures for plants' market power (see Section 7). To focus on the capacity difference, I will abstract from any other difference between plants such as product quality, capital intensity, and workers' skills.

It is useful to contrast the article with Montgomery (1991). Using a wage-posting model, Montgomery shows that identical workers obtain different wages if workers' value of marginal product differs exogenously across employers. For the issue of the size–wage differential, Montgomery's model has three major shortcomings. First, the wage differential therein is not the size–wage differential documented in the empirical literature, because all firms hire one worker each in Montgomery's model. Second, Montgomery does not derive firms' revenue difference from their size difference. Instead, he justifies the exogenous revenue difference by firms' different capital–labor ratios, which empirically account for little of the size–wage differential at the establishment level (see Troske, 1999). Third, since the distribution of firms is exogenous in Montgomery's model, it is not clear whether the wage differential therein is consistent with firms' entry. I overcome these shortcomings by modeling

plant size explicitly, deriving the size–revenue differential endogenously, and allowing for entry.[4]

Another contribution of this article is to use directed search to model both the goods market and the labor market. In contrast, previous price/wage-posting models have focused on only one of these markets. For example, Peters (1991) and Burdett et al. (1997) have focused on a general market, while Montgomery (1991), Moen (1997), and Cao and Shi (2000) have focused on the labor market. It is necessary to integrate the two markets to address the main issue in this article, namely the coexistence of a size–profit differential and a size–wage differential. The two markets also interact to generate the novel result that the size–wage differential depends on the goods market demand in a hump-shaped pattern.

More generally, my model is related to the search models of unemployment (e.g., Pissarides, 1990). Another related literature is the traditional search literature surveyed by McMillan and Rothschild (1994), where buyers only know the distribution of prices before search and learn about a particular seller's price after they visit the seller. Neither model generates a positive capacity–price relation that is central to my analysis.[5]

2. A ONE-PERIOD ECONOMY

I first analyze an economy with one period, leaving the economy with an infinite horizon to Section 4. I also leave the discussions on various modeling assumptions to Section 6.

2.1. *Agents, Markets, and Actions.*   The economy begins with large numbers of buyers, workers, and plants, all being risk neutral. There are B number of identical buyers, each wanting to consume one unit of goods, which yields utility 1. There are I number of workers (whose identities may overlap with the buyers). A fraction u of them are unemployed and the rest are employed at the beginning of the period. Except the employment status, all workers are identical and each worker produces one unit of output when employed. An unemployed worker obtains zero utility from leisure.

The total number of potential plants at the beginning of the period is $(1 - u)I/a$, where $a \in (0, 1)$. A fraction a of them are low-capacity plants, each having one worker already and wanting to hire a second worker. Others are entrants, with a

---

[4] Burdett et al. (1997) have also analyzed how capacity differences affect prices in a one-period model but obtained only numerical results. For this environment I obtain analytical results (see Sections 2 and 3). Also, I analyze the two markets jointly and extend the analysis to an infinite horizon.

[5] Adapting a traditional search model to allow for on-the-job search, Burdett and Mortensen (1998) obtain a size–wage differential but their model implies a lower profit per worker for large firms than for small firms. For the trade-off between price and service probability, Carlton (1978) seems to be the first one to analyze its importance. In contrast to price-posting models, he exogenously assumes that each buyer has a smooth preference ordering over the pair of price and service probability.

number $N \equiv (1 - u)I(1 - a)/a$, each wanting to hire a first worker. As a capacity constraint, each plant can have at most two workers in total and can recruit one worker in the period. Except employment, all plants are the same, using the same technology that produces one unit of output per worker. Other than wages, the cost of production is normalized to zero.

The economy proceeds as in Figure 1. First, the labor market opens, with $(1 - u)I/a$ recruiting plants and uI job seekers (employed workers do not search). Plants post wages simultaneously to attract applicants and, after observing these wages, unemployed workers apply. Each plant awards the job randomly to those who applied. After successfully hiring a new worker, a low-capacity plant becomes a high-capacity plant and an entrant becomes a low-capacity plant. Then plants produce and pay the posted wages. Plants have access to a competitive insurance market that enables them to pay wages even when the goods are not sold. Second, the goods market opens. Sellers post prices for their goods simultaneously and, after observing all prices, each buyer chooses a seller to buy from. A seller awards the goods randomly to those who visit it. Then buyers consume and the economy ends.

The two markets are clearly interdependent. Because the labor market opens before the goods market, plants' labor market decisions depend on the expected payoff in the goods market. Of course, the equilibrium in the goods market also depends on the size distribution of plants generated by the recruiting game in the labor market.

There is no Walrasian auctioneer in the markets. Instead, agents actively organize their transactions without coordination. There is also a search cost, implicit in the assumption that a buyer in the goods market (or an unemployed worker in the labor market) can contact only one agent on the other side of the market in each period. The search cost and the absence of coordination generate match failures and persistent unemployment. The coordination failure in the large markets also makes it appealing to focus on a symmetric equilibrium with mixed strategies. That is, buyers randomize over sellers in the goods market, job applicants randomize over jobs in the labor market, and identical agents use identical strategies.

2.2. *The Goods Market.*   I will analyze the equilibrium in the economy recursively, first the goods market and then the labor market. In the goods market, there are S number of sellers, a fraction $H \in (0, 1)$ of which are high-capacity sellers (each having two units of goods) and the rest are low-capacity sellers (each having one unit of goods). Both S and H are outcomes of the labor market equilibrium but are given for the goods market. I assume HS and $(1 - H)S$ to be integers. The number of buyers is B. Both B and S are large numbers with a finite and positive ratio $b \equiv B/S$.

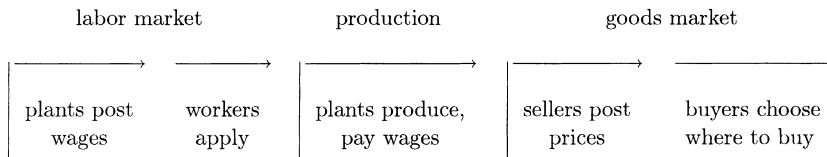| labor market		production	goods market	
plants post wages	workers apply	plants produce, pay wages	sellers post prices	buyers choose where to buy

FIGURE 1

THE SEQUENCE OF ACTIVITIES

The game in the goods market starts with sellers posting prices and then, after observing all posted prices, each buyer chooses a seller to buy from. A high-capacity seller posts a price p_H and a low-capacity seller posts a price p_L. If a low-capacity seller gets k buyers, it chooses one randomly with probability $1/k$ to trade with. Let α_H be the probability that a buyer visits a particular high-capacity seller and α_L be the corresponding probability of visiting a particular low-capacity seller. These probabilities must add up to one:

(2.1) $$HS\alpha_H + (1 - H)S\alpha_L = 1$$

To characterize an equilibrium, let me start with a supposed equilibrium price vector $P \equiv (p_H, \ldots, p_H; p_L, \ldots, p_L)$ and examine a seller's possible deviation.

Consider first a deviation by an arbitrary high-capacity seller to a price p_H^d. For convenience, let the deviator be the first high-capacity seller and name it seller D. The deviation changes the price vector to $P^d = (p_H^d, p_H, \ldots, p_H; p_L, \ldots, p_L)$. Since buyers move after observing sellers' decisions, the deviation induces buyers to modify their strategies. Each buyer revises the probability to $\alpha_H^d(P^d)$ with which he visits seller D. Let q_H^d be the service probability that he gets a good upon visiting seller D. Then his expected utility is $(1 - p_H^d)q_H^d$. For a mixed strategy to be the best response, the buyer must be indifferent between seller D and other sellers. That is,

(2.2) $$(1 - p_H^d)q_H^d = V$$

where V is the expected utility a buyer gets in the market. In a large goods market, each seller has only a negligible influence on V so I treat V as exogenous for each seller.[6] Also, $V \leq 1$, since the best outcome a buyer can achieve is to obtain a good with probability one.

Equation (2.2) states that buyers' strategies, implicit in q_H^d, depend on seller D's price. To compute q_H^d, note that if a buyer visits seller D, he gets a good for sure if the seller has no other buyer or has exactly one other buyer, which occurs with the following probability:

$$(1 - \alpha_H^d)^{B-1} + (B - 1)\alpha_H^d(1 - \alpha_H^d)^{B-2}$$

When the seller has $k \geq 2$ other buyers, the specific buyer gets a good with probability $2/(k+1)$. With the notation $C_B^k \equiv B!/[k!(B-k)!]$, the service probability provided by seller D is[7]

---

[6] See Burdett et al. (1997) for a proof of the convergence of the equilibrium in a finite economy to the limit equilibrium. Cao and Shi (2000) also examined a finite market. Peters (1984) provided a more general proof of the continuity of sellers' payoff function in prices.

[7] The summation in the expression of q_H^d is equal to $g(1)$, where

$$g(y) \equiv \sum_{k=2}^{B-1} \frac{2}{k+1} C_{B-1}^k (y\alpha_H^d)^k (1 - \alpha_H^d)^{B-1-k}$$

To calculate $g(1)$, integrate $g(y)$ with respect to y, sum over k, differentiate with respect to y, and set $y = 1$.

$$q_H^d = (1 - \alpha_H^d)^{B-1} + (B-1)\alpha_H^d(1 - \alpha_H^d)^{B-2} + \sum_{k=2}^{B-1} \frac{2}{k+1} C_{B-1}^k (\alpha_H^d)^k (1 - \alpha_H^d)^{B-1-k}$$

$$= \frac{2}{B\alpha_H^d} \left[1 - (1 - \alpha_H^d)^B \right] - (1 - \alpha_H^d)^{B-1}$$

It is useful to express the service probability as a function of the expected number, i.e., the *queue length*, of buyers who visit the seller. Denoted x_H^d, the queue length for seller D is

$$x_H^d = \sum_{k=1}^{B} k C_B^k (\alpha_H^d)^k (1 - \alpha_H^d)^{B-k} = B\alpha_H^d$$

Noting $\lim_{B \to \infty} (1 - x/B)^B = e^{-x}$, the service probability is

$$(2.3) \qquad q_H^d \to \frac{2}{x_H^d} \left(1 - e^{-x_H^d} \right) - e^{-x_H^d} \equiv q_H(x_H^d) \qquad \text{as } B \to \infty$$

Seller D sets a price to maximize its expected revenue per worker, anticipating the above influence of its price on buyers' strategies but taking other sellers' prices as given. Note that the wage cost does not appear in the seller's pricing problem because it is sunk at the time when sellers post prices. Denoted R_H^d, seller D's expected revenue *per worker* is

$$R_H^d \equiv \frac{1}{2} p_H^d \left[B\alpha_H^d \left(1 - \alpha_H^d \right)^{B-1} + 2 \sum_{k=2}^{B} C_B^k (\alpha_H^d)^k \left(1 - \alpha_H^d \right)^{B-k} \right]$$

The first term in the brackets is the probability that seller D receives exactly one buyer and the summation is the probability that seller D receives two or more buyers. Simplifying, I have $R_H^d = p_H^d q_H^d x_H^d / 2$. Therefore, seller D's best deviation solves the following problem:

$$(2.4) \qquad (PH) \qquad \max_{p_H^d} R_H^d = \frac{1}{2} p_H^d \cdot q_H^d \cdot x_H(q_H^d) \qquad \text{subject to (2.2)}$$

where $x_H(\cdot)$ is the inverse function of $q_H(x)$ in (2.3) such that $x_H(q_H^d) = x_H^d$.

This problem has a unique solution, depicted by point H in Figure 2, where the curve $p_H^d = IND(q_H^d, V)$ is the buyer's indifference relation (2.2) and the curve $p_H^d = ISR_H(q_H^d, R)$ is seller D's iso-revenue curve (depicting combinations of p_H^d and q_H^d that generate the same expected revenue R). The buyer's indifference curve is upward-sloping because a higher service probability must accompany a higher price in order to generate the same expected utility V for a buyer. The iso-revenue curve is upward-sloping because, to provide a higher service probability for buyers and yet maintain the same expected revenue, the seller must charge a higher price. Also, one can verify that the buyer's indifference curve is concave and the seller's iso-revenue curve is convex, producing a unique tangency at point H.[8]

---

[8] The solution is interior because $IND(0, V) = -\infty$, $IND(1, V) = 1 - V \geq 0$, $ISR_H(0, R) = R > 0$ and $ISR_H(1, R) = \infty$.
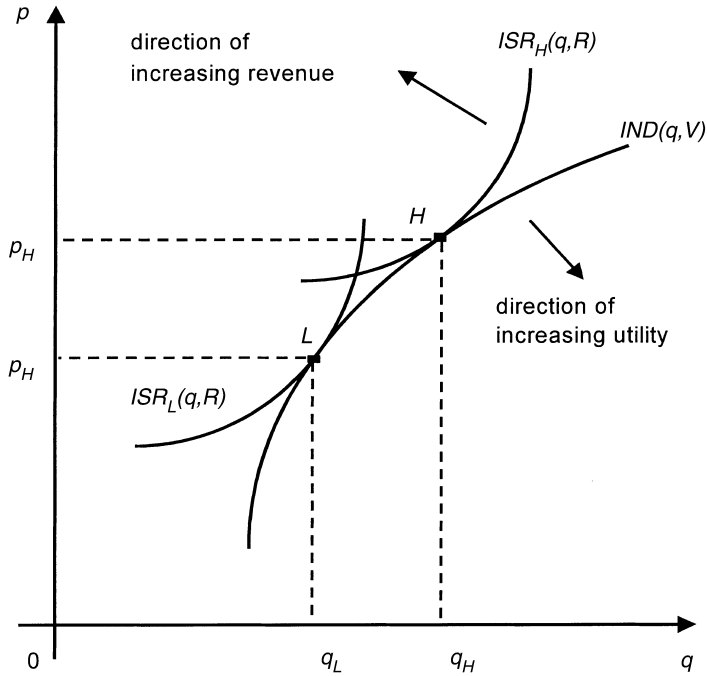
FIGURE 2

THE GOODS MARKET

An important property of the seller's problem is that the service probability $q_H(x)$ is a decreasing function of the expected queue length, with $q_H(0) = 1$ and $q_H(\infty) = 0$. That is, when the expected number of buyers for a seller increases, each buyer gets served with a smaller probability. Also, the service probability is a smooth function. By offering a price slightly lower than everyone else's, seller D cannot expect to attract significantly more buyers. If buyers chose to visit seller D with probability one, each would get served with a very low probability and hence the strategy would not maximize the buyers' expected utility. Instead, each buyer only increases the visiting probability slightly, which results in a slight increase in the queue length that just makes buyers indifferent between seller D and other sellers.

Similarly, a single low-capacity seller's decision, depicted by point L in Figure 2, solves:

(2.5) (PL) $\max\limits_{p_L^d} R_L^d = p_L^d \cdot x_L(q_L^d) \cdot q_L^d$ subject to $p_L^d = IND(q_L^d, V)$

where

(2.6) $x_L^d = B\alpha_L^d$

(2.7) $q_L^d = q_L(x_L^d) \equiv \dfrac{1 - e^{-x_L^d}}{x_L^d}$

A (mixed-strategy) *equilibrium in the goods market* is a quintuple $(p_H, p_L; q_H, q_L; V)$, with $q_H, q_L \in (0, 1)$, such that (i) given V, (p_H, q_H) solve (PH) and (p_L, q_L) solve (PL); (ii) (2.1) is satisfied. The equilibrium queue lengths are $x_H = B\alpha_H$ and $x_L = B\alpha_L$, where α_H and α_L are probabilities for a buyer to visit each high-capacity seller and each low-capacity seller.

To characterize the equilibrium, note that (q_H, q_L) obey (2.3) and (2.7) with (x_H, x_L) replacing (x_H^d, x_L^d). The solutions to (PH) and (PL) obey:

(2.8) $$x_L(q_L) = x_H(q_H) - \ln(1 + x_H) = -\ln V$$

(2.9) $$p_L = 1 - \frac{V}{q_L}, \qquad p_H = 1 - \frac{V}{q_H}$$

Together with (2.3) and (2.7), these conditions determine $(p_H, p_L; x_H, x_L)$ as decreasing functions of V and (q_H, q_L) as increasing functions of V. That is, to achieve a higher expected utility for buyers, prices should be lower, the queue lengths should be shorter, and the service probabilities should be higher.[9] Finally, to determine V, rewrite (2.1) as

(2.10) $$Hx_H(q_H) + (1 - H)x_L(q_L) = b$$

Since $x_L(\cdot)$ and $x_H(\cdot)$ are decreasing functions and since the solutions for (q_L, q_H) are increasing functions of V, the left-hand side of (2.10) is a decreasing function of V and so the equation uniquely determines the equilibrium value of V. A formal statement is as follows:

PROPOSITION 1.   *For any $b \in (0, \infty)$ and $H \in (0, 1)$ there is a unique equilibrium in the goods market that satisfies $x_H \in (b, b/H)$.*

PROOF.   Substitute x_L from (2.10) into (2.8):

(2.11) $$\ln(1 + x_H) = \frac{x_H - b}{1 - H}$$

It is easy to show that there is a unique positive solution for x_H to this equation and the solution satisfies $x_H \in (b, b/H)$. Substituting this solution into (2.3) and (2.7)–(2.9), one determines $(q_H, q_L; x_H, x_L; p_H, p_L)$ and V in equilibrium.   ∎

The property $x_H > b$ means that a high-capacity seller gets an above-average number of buyers and hence gets more buyers than a low-capacity seller does. Equivalently, a buyer visits each high-capacity seller with a higher probability than visiting each low-capacity seller. The property $x_H < b/H$ means that high-capacity sellers do not get all the buyers; buyers also go to low-capacity sellers with a positive probability.

---

[9] To show these properties with Figure 2, note that an increase in V shifts down the indifference curve $IND(q, V)$ southeast but does not affect the iso-revenue curve. Thus, the tangency points L and H move southeast.

2.3. *Properties of the Goods Market Equilibrium.* First, I want to show that a high-capacity seller charges a higher price and obtains a higher revenue per worker than a low-capacity seller. To begin, substituting (q_L, q_H) from (2.3) and (2.7), (x_L, V) from (2.8), and (p_L, p_H) from (2.9), all as functions of x_H, I have

$$(2.12) \qquad \frac{R_H}{R_L} = F2(x_H) \equiv \left. \frac{e^x - 1 - x - x^2/2}{e^x - (1+x)[1 + x - \ln(1+x)]} \right|_{x=x_H}$$

Similarly, substituting (q_L, q_H) and (x_L, V) into (2.9), I have

$$(2.13) \qquad \frac{p_H}{p_L} = \phi(x_H) \equiv \left. \frac{(e^x - 1 - x)[e^x - 1 - x - x^2/2]}{(e^x - 1 - x/2)\{e^x - (1+x)[1 + x - \ln(1+x)]\}} \right|_{x=x_H}$$

Appendix A.1 establishes the following lemma.

LEMMA 1. *$F2(x)$ and $\phi(x)$ are decreasing functions of x.*

Now, since $F2(\infty) = \phi(\infty) = 1$, Lemma 1 immediately implies a positive price differential and a positive revenue differential between high-capacity and low-capacity sellers for all $x_H < \infty$. Equation (2.9) then implies a positive differential in the service probability.

PROPOSITION 2. *In the goods market equilibrium, $R_H > R_L$, $p_H > p_L$, $q_H > q_L$. Moreover, $R_H/R_L > p_H/p_L$.*

A simple explanation for this result is that buyers get served with a higher probability by a high-capacity seller than by a low-capacity seller. To offer the same expected utility to buyers, a low-capacity seller must cut its price below a high-capacity seller's and hence must obtain a lower expected revenue per worker. With the capacity advantage, a large seller not only maintains a higher price but also attracts more buyers per good than a low-capacity seller; i.e., $x_H/2 > x_L$. Thus, the relative revenue per worker exceeds the relative price.

How can a high-capacity seller serve buyers with a higher probability than a low-capacity seller? It is because the buyer has a chance of getting *either* of a high-capacity seller's two goods. To see this, consider the alternative setting where a high-capacity seller sells the two goods separately. That is, such a seller marks the two goods differently and each buyer must specify exactly which good to buy before visiting the seller. Since a buyer cannot choose between the two goods after visiting a high-capacity seller, a high-capacity seller is equivalent to two low-capacity sellers and hence gets exactly twice as many buyers as a low-capacity seller. In this case there is no price differential or revenue differential per worker.

Now I analyze how the size–revenue differential responds to the goods market conditions.

PROPOSITION 3. *The relative price, p_H/p_L, and the size–revenue differential, $\ln(R_H/R_L)$, are decreasing functions of the buyer–seller ratio, b, and increasing functions of the fraction of high-capacity sellers in the market, H.*

PROOF. Under Lemma 1, the proposition is valid if the solution for x_H increases in b and decreases in H. One can directly verify these properties through (2.11). ∎

To illustrate these effects of b and H in Figure 2, note that b and H affect Figure 2 exclusively through the expected utility index V and that their effects are opposite. An increase in b increases the number of buyers per seller, reduces buyers' expected utility, and shifts the indifference curve $IND(q, V)$ up northwest. This increases prices and reduces service probabilities provided by both types of sellers. On the opposite, an increase in H increases the total goods supply, increases buyers' expected utility, and hence shifts the indifference curve down southeast. This reduces prices and increases service probabilities provided by both types of sellers.

Focus on the effect of b. Since an increase in b reduces buyers' expected utility V, it shifts the buyers' indifference curve up northwest, but the shift is not uniform for all levels of q. Since $|\partial p/\partial V| = 1/q$ along the indifference curve, the indifference curve shifts up by more for low levels of q than for high levels of q. Thus the tangency point L moves up by more than the tangency point H in Figure 2, resulting in a proportionally larger increase in p_L than in p_H. This reduces the relative price and the relative revenue.

Intuitively, an increase in the buyer–seller ratio reduces the relative price and the relative revenue because it reduces the seller's benefit of an additional capacity. The additional capacity of a high-capacity seller is highly effective in attracting buyers only when there are few buyers per seller. When there are many buyers per seller, a seller finds it easy to sell the goods. By adding an additional capacity, a seller does not provide a much better service probability to buyers (each of them still finds it hard to get served) and so a high-capacity seller cannot charge a much higher price than a low-capacity seller. In the extreme case $b \to \infty$, the relative price and the relative revenue approach 1.

In Proposition 3 I have treated b and H as parameters. The labor market equilibrium described below determines (b, H).

2.4. *The Labor Market.* In the labor market, plants take the expected revenues (R_L, R_H) as given. The reason is that these revenues depend on the labor market entirely through the summary statistics (b, H); when the labor market is large, a single plant's recruiting decision has no influence on these statistics.[10] Thus, (R_L, R_H) are exogenous in this section.

---

[10] That is, given a plant's employment after recruiting, the plant cannot change the *maximum* expected revenue that it expects from the goods market equilibrium. This is quite different from saying that a plant in the labor market cannot use its recruiting strategy to alter its feasibility set in the goods market. I do not assume the latter.

In the labor market, there are uI number of unemployed workers and $(1 - u)I/a$ number of recruiting plants. A number $(1 - u)I$ of the recruiters are low capacity, each having one worker already, and N of them are entrants, where $N = (1 - u)I(1 - a)/a$. To recruit, each entrant posts a wage w_L and each low-capacity plant posts a wage w_H. If a low-capacity plant successfully recruits the second worker, it pays the wage w_H to both the new worker and the existing one; if it fails to recruit the second worker, it pays the wage w_L to the existing worker.[11] Thus, all plants that produce with one worker pay a wage w_L, regardless of their past histories of employment. All plants that produce with two workers pay a wage w_H.

The aim of the analysis is to find the conditions for $w_H > w_L$. Anticipating this, I call a low-capacity plant a *high-wage recruiting plant* and an entrant a *low-wage recruiting plant*. As in a standard search model, an employed worker must quit first before applying for another job. Since an employed worker obtains a higher utility than an unemployed worker, no worker quits and hence only unemployed workers apply for jobs.

The recruiting game is similar to the pricing game in the goods market. A recruiting plant posts a wage, taking other plants' wages as given. Observing the posted wages, unemployed workers decide which job to apply for. If there are at least two workers applying for the same job, one is chosen by the plant with equal probability. Let γ_H be the probability with which an unemployed worker applies for each high-wage job opening and γ_L be the probability with which he applies for each low-wage job opening. Denote z_H as the expected number of applicants, i.e., the queue length, for each high-wage job and z_L as the queue length for each low-wage job. Then $z_H = uI\gamma_H$ and $z_L = uI\gamma_L$. When $I \to \infty$, these probabilities are finite and positive. In this limit, an unemployed worker gets a job with probability $(1 - e^{-z_H})/z_H$ when applying for a high-wage job and gets a job with probability $(1 - e^{-z_L})/z_L$ when applying for a low-wage job. With probability $1 - e^{-z_L}$ each entrant hires a worker and becomes a low-capacity plant; with probability $1 - e^{-z_H}$ each low-capacity plant hires a second worker and becomes a high-capacity plant. Moreover, since $N\gamma_L + Na/(1 - a)\gamma_H = 1$, I have

$$(2.14) \qquad z_L + \frac{a}{1 - a} z_H = \frac{uI}{N} \qquad \left(= \frac{au}{(1 - a)(1 - u)} \right)$$

Each job applicant maximizes the expected wage and, in equilibrium, obtains the same expected wage V_u from all job openings. Each recruiting firm maximizes the expected profit, anticipating the influence of its wage decisions on the applicants' choices. For an entrant, the recruiting problem is

$$(PN) \qquad \max_{w_L} (1 - e^{-z_L})(R_L - w_L)$$

$$\text{s.t.} \qquad \frac{1 - e^{-z_L}}{z_L} w_L = V_u$$

---

[11] If the existing worker's wage were not updated, there would be a negative experience wage premium in the firm. It is possible to rule out this unrealistic result by eliminating the assumption that employed workers do not search (see Section 6).

where the plant takes (V_u, R_L) as given. Similarly, a low-capacity plant's recruiting problem is

$$(P1) \qquad \max_{w_H}(1 - e^{-z_H})2(R_H - w_H) + e^{-z_H}(R_L - w_L)$$

$$\text{s.t.} \qquad \frac{1 - e^{-z_H}}{z_H} w_H = V_u$$

Again, the firm takes (V_u, R_H) as given. The second term in the objective function is a low-capacity plant's outside option when it fails to recruit the second worker.

Solving the above maximization problems, I obtain the following equations:

$$(2.15) \qquad\qquad\qquad V_u = R_L e^{-z_L}$$

$$(2.16) \qquad\qquad\qquad w_L = 2V_u e^{z_H} - 2R_H + R_L$$

$$(2.17) \qquad\qquad \frac{1 - e^{-z_L}}{z_L} w_L = \frac{1 - e^{-z_H}}{z_H} w_H = V_u$$

A *labor market equilibrium* is a collection $(V_u, z_L, z_H, w_L, w_H)$ such that the following conditions hold: (i) given (V_u, R_L, R_H) and given that other recruiting plants post w_L (if they are entrants) or w_H (if they are low-capacity plants), w_L solves an entrant's recruiting problem (PN) and w_H solves a low-capacity firm's recruiting problem $(P1)$; (ii) the expected lengths of applicants, (z_L, z_H), satisfy the adding-up constraint (2.14) and make each applicant indifferent between a job opening with w_L and a job opening with w_H; and (iii) an unemployed worker's expected utility, V_u, obeys (2.17).

The solutions to (2.14)–(2.17) characterize a labor market equilibrium. From (2.17), it is apparent that a low-capacity firm posts a higher wage than an entrant if and only if it has a longer queue of applicants:

REMARK 1. $w_H > w_L$ if and only if $z_H > z_L$.

3.   ECONOMY-WIDE EQUILIBRIUM

3.1. *Characterization.*   An *economy-wide equilibrium* is a goods market equilibrium and a labor market equilibrium with the following additional restrictions: (i) the expected revenues (R_L, R_H) taken as given in the labor market equilibrium are equal to those generated by the goods market equilibrium; and (ii) the statistics (H, b) taken as given in the goods market equilibrium are equal to those generated by the recruiting activities in the labor market.

An economy-wide equilibrium determines the variables $(R_H, R_L, p_H, p_L, x_H, x_L)$, $(w_H, w_L, z_H, z_L, V_u)$, and (H, b). The parameters are (a, u, N, I, B), with $N = (1 - u)I(1 - a)/a$. Since N, I, and B all approach infinity, only their relative values are finite. Choose the total number of workers I as the denominator and denote $\theta = B/I$ as the buyer–worker ratio. Then

$$\frac{N}{I} = \frac{1 - a}{a}(1 - u), \qquad z_H = \frac{u}{1 - u} - \frac{1 - a}{a} z_L$$

The relation between z_H and z_L comes from (2.14). In the rest of this article, I shorten the notation x_H to x and z_L to z.

The restrictions on (H, b) imposed by the recruiting additivities are

$$SH = \frac{a}{1-a}N(1 - e^{-z_H}), \qquad S = \frac{a}{1-a}N + (1 - e^{-z})N$$

The first equation states that high-capacity sellers in the goods market are those plants that have low capacity at the beginning of the period but have recruited successfully for the second worker. The second equation states that the total number of sellers in the goods market is the sum of low-capacity plants at the beginning of the period and those entrants that have recruited successfully. Substituting N/I, z_H, and $S(= B/b)$, I have

$$(3.1) \qquad H = \frac{a\left[1 - e^{-\left(\frac{u}{1-u} - \frac{1-a}{a}z\right)}\right]}{1 - (1-a)e^{-z}}$$

$$(3.2) \qquad b = \frac{B}{S} = \frac{a\theta/(1-u)}{1 - (1-a)e^{-z}}$$

To determine the economy-wide equilibrium, I determine z by deriving two relationships between the relative revenue R_H/R_L and z from the two markets. For the labor market, substituting z_H and eliminating (V_u, w_L, w_H) from (2.15)–(2.17) I have

$$(3.3) \qquad \frac{R_H}{R_L} = F1(z) \equiv \frac{1}{2}\left\{1 + e^{-z}\left[2e^{\left(\frac{u}{1-u} - \frac{1-a}{a}z\right)} - \frac{z}{1 - e^{-z}}\right]\right\}$$

For the goods market, substituting (3.1) and (3.2) into (2.11) I have

$$(3.4) \qquad \ln(1 + x) = \frac{[1 - (1-a)e^{-z}]x - \frac{a\theta}{1-u}}{(1-a)(1 - e^{-z}) + ae^{-\left(\frac{u}{1-u} - \frac{1-a}{a}z\right)}}$$

Proposition 1 shows that there is a unique solution for x to this equation. Denote this solution as $x(z)$. Then (2.12) implies

$$(3.5) \qquad \frac{R_H}{R_L} = F2(x(z))$$

Equations (3.3) and (3.5) determine $(R_H/R_L, z)$ jointly. Then I can recover other variables.

The following lemma documents the features of $F1(z)$ and the proposition shows that the economy-wide equilibrium exists. The proofs appear in Appendix A.2.

LEMMA 2.   *Define z_1 as the positive solution to the following equation*:

$$(3.6) \qquad 2e^{\left(\frac{u}{1-u} - \frac{1-a}{a}z\right)} - \frac{z}{1 - e^{-z}} = e^z$$

Then, (i) z_1 *is well defined and* $z_1 \in (0, au/[(1-a)(1-u)])$; (ii) $F1'(z) < 0$ *for* $z \leq z_1$; *and* (iii) $F1(z) < 1$ *iff* $z > z_1$.
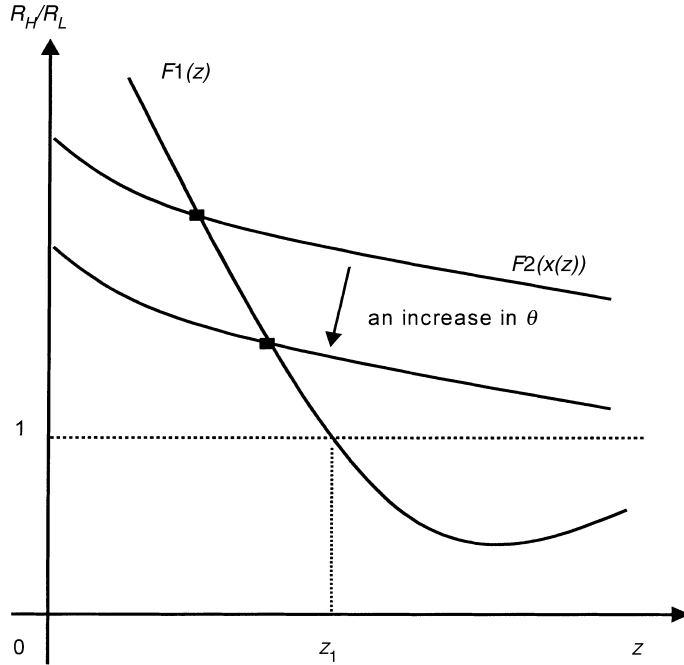
FIGURE 3

THE ECONOMY-WIDE EQUILIBRIUM

PROPOSITION 4. *There exists $\theta_0 > 0$ such that an economy-wide equilibrium exists if $\theta > \theta_0$. The solution for z in the equilibrium lies in $(0, z_1)$.*

Figure 3 depicts the functions $F1(z)$ and $F2(x(z))$. Although $F2(x(z))$ need not be a decreasing function of z, it is drawn so in the figure. The condition $\theta > \theta_0$ is necessary and sufficient for the curve $F1(z)$ to be above the curve $F2(x(z))$ when $z = 0$. When $\theta > \theta_0$, there is at least one solution for z and the solution lies in $(0, z_1)$. Since $F2(x(z))$ may be an increasing or a decreasing function of z, the solution may not be unique but is assumed so in the analysis below.[12]

For the economy-wide equilibrium to exist, the number of buyers per worker cannot be too small. If there are only a small number of buyers per worker, then there are only a small number of buyers per seller, since the relative number of plants to workers is finite. In this case a large capacity gives the seller an important advantage in attracting buyers (see Section 2.3). Small sellers must drastically cut prices in order to attract buyers and so the size–revenue differential is very large. Anticipating this large revenue differential, high-wage recruiting plants will post a

---

[12] The number of solutions is odd. If there is more than one solution, the smallest and largest solutions for z both have the analytical properties analyzed here.

very high wage in order to hire a second worker, attracting all applicants. The restriction $\theta > \theta_0$ ensures that large and small plants coexist in equilibrium.

3.2. *Properties.* The most important property of the economy-wide equilibrium is a positive size–wage differential, documented in the following proposition and proven in Appendix A.3.

PROPOSITION 5. *There exists $\theta_1 > 0$ such that $w_H > w_L$ if and only if $\theta < \theta_1$. Moreover, $\theta_1 > \theta_0$ if a is either close to 1 or close to 0.*

For the size–wage differential to be positive, the number of buyers per worker cannot be too large. To explain, note that a low-capacity plant has a higher outside option than an entrant: When no applicant shows up, a low-capacity plant gets some rent from the existing worker but an entrant gets nothing. This higher "outside option" enables a low-capacity plant to extract a larger fraction of the match surplus from a worker than an entrant does, hence tending to produce a negative size–wage differential. For the differential to be positive, a low-capacity plant must have a significantly larger match surplus ("pie") to be shared with a new recruit than an entrant does. This requires the buyer–worker ratio to be not too large. If, instead, there are many buyers per worker, then there are many buyers per seller in the goods market. Since every seller can sell the goods easily, a larger capacity does not increase a seller's expected revenue per worker much and so the pie (per worker) is almost the same for large and small sellers (see Proposition 3).

For a positive size–wage differential to be consistent with equilibrium, θ_1 must be greater than θ_0. Proposition 5 shows $\theta_1 > \theta_0$ when the fraction of low-capacity (high-wage recruiting) plants in the labor market is either large or small. I believe that $\theta_1 > \theta_0$ holds for all $a \in (0, 1)$, although I have not been able to prove this general result analytically.[13]

The size–wage differential responds to changes in parameters (θ, u, a). The effects of θ are as follows (the proof is straightforward and omitted):

COROLLARY 1. *In the economy-wide equilibrium, z increases in θ and z_H decreases in θ. Thus, R_H/R_L and w_H/w_L are both lower when θ is larger.*

The explanation for this result is the same as the above one for why a size–wage differential is positive only when $\theta < \theta_1$. That is, by reducing the advantage of a large capacity, an increase in the number of buyers reduces the size–revenue differential and hence reduces the size–wage differential. One can trace the effects of θ through Figure 3. Since $x(z)$ is an increasing function of θ for given z (see (3.4)) and $F2(x)$ is a decreasing function of x, an increase in θ shifts down the curve $F2(x(z))$. The curve

---

[13] Numerical exercises support this general result. It is difficult to prove the result analytically because θ_0 and θ_1 are both defined implicitly through $F2(x(z)) = F1(z)$ by setting z to some special values; the dependence of $F2(x(z))$ on θ is complicated.

$F1(z)$ remains intact. Thus, the solution for z increases, z_H falls, and the relative revenue R_H/R_L falls. By (2.17), the size–wage differential falls.

The effects of a and u are more complicated, since they affect the equilibrium value of z ambiguously and also change the size–wage differential directly. In an infinite-horizon economy, a and u are endogenous. I turn to this infinite-horizon economy in the following section.

4. AN INFINITE HORIZON

In an infinite-horizon economy, the number of entrants must be consistent with free entry rather than being exogenous as in the one-period economy. Also, the distribution of plants/workers must be consistent with a stationary equilibrium. By focusing on a stationary equilibrium, I underscore the fact that a positive size–wage differential can persist and survive plants' entry. I also obtain new features of the size–wage differential through the size distribution of plants.

4.1. *Flows of Plants and Workers.*   The economy lasts forever, with discrete time. In each period, actions follow the sequence in Figure 1, with an additional phase—job separation—that takes place at the end of the period after the goods market closes. Goods are perishable across periods. The goods market functions as before and so all the results in Sections 2.2 and 2.3 hold for given (H, b). The recruiting game in the labor market is also similar to the one before. Only entrants and low-capacity plants recruit, each recruiting one worker at a time.

Anyone can pay a recruiting cost $c > 0$ per period and enter the labor market as an entrant at the beginning of a period. Some matches separate after the goods market closes. Each plant receives a shock to the match-specific productivity that causes only *one* worker to separate. The shock occurs to a high-capacity plant with probability $2\sigma > 0$ and to a low-capacity plant with probability σ. For a high-capacity plant that experiences the shock, either worker is chosen randomly with probability $1/2$ to be the one who separates. Thus, ex ante, each worker faces the same separation probability σ in both large and small plants. Job separation turns a high-capacity plant into a low-capacity plant and a low-capacity plant into a potential entrant.[14]

When a high-capacity plant experiences job separation, it cuts the remaining worker's wage to w_L next period unless it successfully recruits the second worker in the next period. Also, as before, if a low-capacity plant succeeds in recruiting the second worker, it pays both the new worker and the existing worker the high wage w_H. These assumptions ensure that all low-capacity plants pay the same wage w_L and so simplify the analysis.

---

[14] Although it is more reasonable to assume that the job separation rate is lower in larger plants, the assumption of equal separation rates here strengthens the results. Also, note that the specific way of modeling separation is different from the standard one (e.g., Pissarides, 1990), which assumes each worker to experience the shock independently with probability σ. With this standard modeling approach, a high-capacity plant can lose both workers with probability σ^2. In contrast, the approach here ensures that a high-capacity plant never loses two workers in one period. This makes the flows of plants more manageable.

With the same notation $(z, z_H, a, N, S, H, B, I)$ as before, the job market activities generate a size distribution of plants summarized in Table 1. The following restrictions are necessary and sufficient for the size distribution to be stationary:

(4.1) $$2\sigma SH = \frac{a}{1-a}N(1 - e^{-z_H})$$

(4.2) $$\sigma S(1 - H) = N(1 - e^{-z})$$

(4.3) $$\frac{a}{1-a}N = 2\sigma SH + (1 - \sigma)S(1 - H)$$

Equation (4.1) requires that the number of high-capacity plants that switch into low-capacity plants through separation be equal to the number of low-capacity plants that successfully recruit the second worker. Equation (4.2) requires that the number of low-capacity plants that experience job separation be equal to the number of entrants that successfully recruit a worker. Equation (4.3) requires that the number of low-capacity plants at the beginning of each period be equal to the number of those low-capacity plants that survived job separation in the last period, $(1 - \sigma)S(1 - H)$, plus the flow of high-capacity plants that experienced job separation last period, $2\sigma SH$. The distribution of workers is also stationary under (4.1)–(4.3).

The number of workers (I) and the number of buyers (B) are exogenous but other distribution variables (N, S, H, a, u, b) are all endogenous. I take the limit $I \to \infty$, fix $\theta \equiv B/I$ at an exogenous (finite) level, and determine ($H, a, u, b, N/I, S/I$). It is convenient to write (z_H, z) and (N, S, u, b) as functions of (H, a), leaving (H, a) to be determined later. With (4.1)–(4.3), (2.14), and the definitions of (u, b), I can solve

(4.4)
$$
\begin{cases}
z_H = \ln\left[1 + \frac{2\sigma H}{(1-\sigma)(1-H)}\right] \\
z = -\ln\left[1 - \frac{a}{1-a} \cdot \frac{\sigma(1-H)}{(1-\sigma)(1-H)+2\sigma H}\right]
\end{cases}
$$

(4.5)
$$
\begin{cases}
\frac{N}{I} = \left[z + \frac{a}{1-a}\left(z_H + \frac{1-\sigma(1-H)}{2\sigma H+(1-\sigma)(1-H)}\right)\right]^{-1} \\
\frac{S}{I} = \frac{a}{1-a} \cdot \frac{N/I}{2\sigma H+(1-\sigma)(1-H)} \\
u = \left(z + \frac{a}{1-a}z_H\right)\frac{N}{I} \\
b = \frac{\theta(1-a)}{a} \cdot \frac{2\sigma H+(1-\sigma)(1-H)}{N/I}
\end{cases}
$$

TABLE 1

DISTRIBUTION OF PLANTS

	Plants		
Time	High-Capacity Plants	Low-Capacity Plants	Entrants
Before recruiting	$(1 - 2\sigma)SH$	$\frac{a}{1-a}N$	N
After recruiting	SH	$S(1 - H)$	Ne^{-z}

4.2. Wages and Value Functions. The wage-posting game determines the wages, queue lengths, and value functions. The value functions are as follows:

V_H (J_H): the present value of a worker (owner) in a high-capacity plant after production;

V_L (J_L): the present value of a worker (owner) in a low-capacity plant after production;

V_R (J_R): the present value of a worker (owner) in a low-capacity plant before recruiting;

V_u: the present value of unemployment before recruiting;

J_N: the present value of an entrant before recruiting.

Let $\beta \in (0, 1)$ be the common discount factor of workers and plants. Workers' value functions obey the following Bellman equations:

$$(4.6) \qquad \begin{cases} V_u &= \frac{1-e^{-z_H}}{z_H}(V_H - \beta V_u) + \beta V_u \\[2mm] &= \frac{1-e^{-z}}{z}(V_L - \beta V_u) + \beta V_u \end{cases}$$

$$(4.7) \qquad V_H = w_H + (1 - 2\sigma)\beta V_H + \sigma\beta(V_u + V_R)$$

$$(4.8) \qquad V_L = w_L + (1 - \sigma)\beta V_R + \sigma\beta V_u$$

$$(4.9) \qquad V_R = (1 - e^{-z_H})V_H + e^{-z_H}V_L$$

The explanations from (4.6)–(4.8) are similar. Consider (4.7) for example. It states that the present value of a worker in a high-capacity plant equals the sum of the current wage and the expected future value. The expected future value consists of two terms. With probability $(1 - 2\sigma)$ the plant does not experience job separation, in which case the future value for the worker is βV_H. With probability 2σ the plant experiences job separation and with probability $1/2$ the specific worker is the one who separates. The corresponding future value for the worker is $\beta(V_u + V_R)/2$.

The Bellman equation (4.6) has two lines, the first line calculating the expected value to an unemployed worker when he applies to a high-wage job opening and the second line calculating the expected value when he applies to a low-wage job opening. The expected gains to an unemployed worker from applying to the two types of jobs must be the same and so

$$(4.10) \qquad \frac{1 - e^{-z_H}}{z_H}(V_H - \beta V_u) = \frac{1 - e^{-z}}{z}(V_L - \beta V_u)$$

When an applicant gains more from a high-wage job than from a low-wage job, he also faces a longer queue when applying for a high-wage job.

Similarly, the Bellman equations for plants' value functions are as follows:

$$(4.11) \qquad J_N = (1 - e^{-z})J_L + e^{-z}\beta J_N - c$$

$$(4.12) \qquad J_H = 2(R_H - w_H) + (1 - 2\sigma)\beta J_H + 2\sigma\beta J_R$$

(4.13) $$J_L = (R_L - w_L) + (1 - \sigma)\beta J_R + \sigma\beta J_N$$

(4.14) $$J_R = (1 - e^{-z_H})J_H + e^{-z_H}J_L - c$$

A recruiting plant posts the wage to maximize the expected value at the time of recruiting. Consider first a one-period deviation by a single low-capacity plant to a wage w_H^d. Observing this deviation, each job applicant revises his application probability that results in a new queue length for the deviator, z_H^d. The deviator's value function is

(4.15) $$J_R^d = (1 - e^{-z_H^d})J_H^d(w_H^d) + e^{-z_H^d}J_L - c$$

where $J_H^d(w_H^d)$ obeys (4.12) with J_H^d replacing J_H and w_H^d replacing w_H.[15] Taking (J_R, J_L, V_u, V_R) as given, the deviator maximizes J_R^d subject to applicants' indifference condition:

(4.16) $$\frac{1 - e^{-z_H^d}}{z_H^d}\left[V_H^d(w_H^d) - \beta V_u\right] = (1 - \beta)V_u$$

Equation (4.16) comes from (4.6) and incorporates the recruiter's deviation. The value function $V_H^d(w_H^d)$ is the present value of a worker working for the wage w_H^d, which obeys (4.7) with V_H^d replacing V_H and w_H^d replacing w_H.

The first-order condition of the deviator's problem is

(4.17) $$\frac{J_H^d(w_H^d) - J_L}{2} = (1 - \beta)V_u\left[e^{z_H^d} - \frac{z_H^d}{1 - e^{-z_H^d}}\right]$$

Since the right-hand side is an increasing function of z_H^d, this condition states intuitively that the fraction of the match surplus a low-capacity plant gets from hiring the second worker increases with the queue length. Similarly, by analyzing an entrant's deviation I obtain[16]

(4.18) $$J_L^d(w_L^d) - \beta J_N = (1 - \beta)V_u\left[e^{z^d} - \frac{z^d}{1 - e^{-z^d}}\right]$$

For the wage distribution (w_H, w_L) and queue lengths (z_H, z) to be in equilibrium, these deviations cannot be profitable and so the above conditions must be satisfied by $(w_H^d, z_H^d) = (w_H, z_H)$ and $(w_L^d, z^d) = (w_L, z)$. Retrieving (J_H, J_L) from (4.14) and (4.11) and substituting into the above equations, I have

(4.19) $$J_R - J_L + c = 2(1 - \beta)V_u(e^{z_H} - 1 - z_H)$$

(4.20) $$(1 - \beta)J_N + c = (1 - \beta)V_u(e^z - 1 - z)$$

---

[15] To compute J_H^d I treat the term J_R in (4.12) as being unaffected by the deviation. This is because the deviation is a one-period deviation, which does not affect the plant's future recruiting behavior after job separation.

[16] A one-period deviation by an entrant does not change the entrant's future value if he fails to recruit in the current period. Thus, the deviator's present value is $J_N^d = [1 - e^{-z^d}]J_L^d(w_L^d) - c + (e^{-z^d})\beta J_N$, where $J_L^d(w_L^d)$ is calculated by replacing w_L by w_L^d and J_R by J_R^d in (4.13). In turn, J_R^d is calculated through (4.14) by replacing J_L with J_L^d.

4.3. *Steady-State Equilibrium.* A *steady-state equilibrium in the labor market* consists of value functions (V_u, V_H, V_L, V_R) given by (4.6)–(4.9), (J_N, J_H, J_L, J_R) given by (4.11)–(4.14), queue lengths $z_H, z \in (0, \infty)$, and distribution variables $(H, a, u, b, N/I, S/I)$ such that the following conditions hold:

(i) Each applicant is indifferent between the two wages; i.e., (4.10) and (2.14) hold.

(ii) A recruiting plant has no incentive to deviate from the wages; i.e., (4.19) and (4.20) hold.

(iii) There is a stationary distribution of plants; i.e., (4.4) and (4.5) hold.

(iv) Free entry: the net profit of an entrant is zero; i.e., $J_N = 0$.

(v) A low-capacity plant is willing to recruit a second worker; i.e., $J_R \geq J_L$.

(vi) Employed workers do not have the incentive to quit; i.e., $V_H \geq V_u$ and $V_L \geq V_u$.

Condition (vi) is always satisfied in a mixed equilibrium and (v) is satisfied as long as a higher wage attracts more applicants than a lower wage does (see Appendix A.4 for a proof).

LEMMA 3. $V_H, V_L \geq V_u$ whenever $z_H, z \in (0, \infty)$. Also, $J_R \geq J_L$ if $z_H \geq z$.

An *economy-wide equilibrium* is a joint equilibrium in the goods market and the labor market. I determine this equilibrium following the steps below. First, set $J_N = 0$ in (4.20) and (4.11) to obtain (V_u, J_L) as functions of z:

$$(4.21) \qquad V_u = \frac{c/(1 - \beta)}{e^z - 1 - z}, \qquad J_L = \frac{c}{1 - e^{-z}}$$

Since (z, z_H) are functions of (H, a) by (4.4), V_u and J_L are functions of (H, a). Second, substitute V_u into (4.10), (4.6), and (4.9) to obtain (V_H, V_L, V_R) as functions of (H, a); substitute J_L into (4.19) and (4.14) to obtain (J_R, J_H) as functions of (H, a). Third, substitute the functions (V_H, V_L, V_R, V_u) into (4.7) and (4.8) to express (w_H, w_L) as functions of (H, a). Substituting these functions and (J_H, J_L, J_R) into (4.12) and (4.13), I obtain two equations that involve $(H, a; R_H, R_L)$. Since (R_H, R_L) are functions of (H, b) as calculated in the goods market equilibrium, and b is a function of (H, a), these two equations solve (H, a).

PROPOSITION 6. *In the equilibrium, $w_H > w_L$ only if $R_H > R_L$. The following four statements are equivalent*: $z_H > z$, $V_H > V_L$, $w_H > w_L$, $J_H > 3J_L$.

The proof of this proposition appears in Appendix A.4. The proposition extends two features of the one-period economy into an infinite-horizon economy. First, as in Remark 1, a size–wage differential is consistent with the equilibrium if and only if a low-capacity plant attracts more applicants than an entrant. An applicant gets the same expected present value from applying for the two jobs. Second, a positive size–revenue differential between plants is necessary but may not be sufficient for a positive size–wage differential.

Suppose that the queue is longer for a low-capacity plant than for an entrant. Then Proposition 6 states that a large plant and its workers both get higher values than a

small plant and its worker. When $z_H > z$, a high-capacity plant's present value per worker, $J_H/2$, is more than 50 percent higher than a low-capacity plant's present value per worker, J_L. This positive correlation between the size–wage and size–profit differentials is consistent with the empirical observation by Katz and Summers (1989).[17]

For all these desirable features to arise from the model, I need to find the parameter region in which $z_H > z$. Doing so analytically is difficult so I turn to numerical exercises.

5. NUMERICAL EXERCISES

There are four parameters in the model: the job separation rate (σ), the recruiting cost (c), the discount factor (β), and the buyer–worker ratio (θ). I allow θ to vary in [0.2, 5.2] and interpret an increase in θ as an increase in the product demand.[18] I set the discount factor at $\beta = 0.99$ to give a quarterly real interest rate 0.01, the job separation rate at $\sigma = 0.06$ to match the quarterly transition rate from employment to unemployment (Blanchard and Diamond, 1989), and the recruiting cost at $c = 0.02$ (see Hamermesh, 1993).

Figure 4(a) depicts the price differential and size–revenue differential as functions of the buyer–worker ratio θ, where $RP = \ln(p_H/p_L)$ and $RR = \ln(R_H/R_L)$. Both differentials are positive. As shown in Proposition 2, the size–revenue differential is larger than the price differential, indicating that competition through the service probability is an important source of the revenue differential. The two differentials fall when the buyer–worker ratio increases, as shown in Proposition 3. As the buyer–worker ratio becomes sufficiently large, the price differential and the revenue differential between large and small plants approach zero.

Since the expected revenue per worker increases in both low-capacity and high-capacity plants when the product demand increases, all plants are able to offer higher wages when θ is higher. Figure 4(b) depicts the size–wage differential, defined as $RW = \ln(w_H/w_L)$, and the variance of logarithmic wages, VLW.[19] The two measures RW and VLW have qualitatively the same dependence on the demand for the industry's goods, except when $RW < 0$ (since VLW is always nonnegative by definition). The size–wage differential is positive (i.e., $RW > 0$) only when the product demand is moderate. When the product demand is sufficiently low ($\theta < 0.4$) or sufficiently high ($\theta > 5$), the size–wage differential is negative.

Moreover, the size–wage differential depends on θ in a hump-shaped pattern. An increase in the product demand increases the size–wage differential when the demand is initially low and reduces the size–wage differential when the demand is high. As the product demand becomes sufficiently high, the size–wage differential

---

[17] To explain why $z_H > z$ implies $J_H > 3J_L$, note that $(J_L - \beta J_N)$ is an entrant's gain from recruiting the first worker and $(J_H - J_L)/2$ is a low-capacity plant's gain per worker from recruiting the second worker. When $z_H > z$, a low-capacity plant must obtain a larger gain per worker from recruiting than an entrant; i.e., $(J_H - J_L)/2 > J_L - \beta J_N$. Since $J_N = 0$ in equilibrium, this implies $J_H > 3J_L$. To support this argument, compare (4.17) and (4.18).

[18] A fall in the number of workers has the same effects.

[19] With RW as the definition of the size–wage differential, changes in wages that have the same proportion across plants do not affect the wage differential.

approaches zero. The size–wage differential peaks at $\theta^* \approx 1.1$. I explain the segments $\theta > \theta^*$ and $\theta < \theta^*$ of Figure 4(b) separately below.

When $\theta > \theta^*$, the intuition for Proposition 5 explains the behavior of the size–wage differential. That is, when the number of buyers per seller increases, every seller finds it easier than before to sell the goods successfully and so the additional revenue per worker generated by a larger capacity falls. A diminishing size–revenue differential compresses the size–wage differential. When the product demand is sufficiently high ($\theta > 5$), a large seller and a small seller obtain almost the same expected revenue per worker, in which case the higher outside option of a low-capacity plant enables it to post a lower wage than an entrant (see the discussion on Proposition 5).

When $\theta < \theta^*$, the size–wage differential rises with the product demand, in contrast to the falling size–revenue differential. This is a unique feature of the infinite-horizon economy. Notice that, for small θ, the log difference between the queues to the two job openings, denoted $RZ(\theta) = \ln(z_H/z)$, increases with θ (Figure 4(c)). So does the present value differential between hiring and not hiring the second worker for a low-capacity plant, denoted $RJ(\theta) = \ln(J_R/J_L)$.

What is the unique feature of the infinite-horizon economy that makes the size–wage differential rise with the product demand when $\theta < \theta^*$? It is the size distribution of plants. To see this, it is necessary to explain first why the size–wage differential is negative when θ is very low (e.g., $\theta < 0.4$). When θ is sufficiently small, the number of buyers per seller is small and so sellers' expected revenues are low. The incentive to enter the industry is low and so, as depicted in Figure 4(d), entrants are a negligible fraction of the recruiters (i.e., $a \approx 1$). There are many more low-capacity plants than entrants, although both are small fractions of the total number of plants. This implies that, for every entrant, there is a large flow from low-capacity plants to entrants generated by job separation. To balance this flow, entrants must recruit very quickly and the only way to do so is to post a high wage (see (4.2)). This is why the size–wage differential is negative and why there are more applicants for each entrant than for each low-capacity plant when θ is sufficiently small.[20]

Now I can explain why the size–wage differential rises with θ when $\theta < \theta^*$. An increase in θ, by increasing the product demand, increases the expected revenue for all plants and stimulates entry. The relative number of entrants to low-capacity plants increases; i.e., a falls as depicted in Figure 4(d). If the relative wage between the two plants were unchanged, the flow from entrants to low-capacity plants generated by recruiting would exceed the reverse flow generated by separation. To maintain a stationary size distribution, entrants' recruiting rate increases by less than

---

[20] Since low-capacity plants are also a small fraction of the total number of plants (i.e., $H \approx 1$ as depicted in Figure 4(d)), job separation generates a large flow from high-capacity plants to low-capacity plants and hence the size distribution requires a high recruiting rate for each low-capacity plant. However, the requirement on entrants' recruiting is stronger. When $\theta = 0.2$, for example, the inflow per entrant from low-capacity plants to entrants is $\sigma S(1 - H)/N \approx 0.952$. This is larger than the inflow per low-capacity plant from high-capacity plants to low-capacity plants, which is $2\sigma SH(1 - a)/(aN) \approx 0.926$. To maintain a stationary distribution, (4.1) and (4.2) require that entrants recruit more quickly than low-capacity plants; i.e. $z > z_H$.
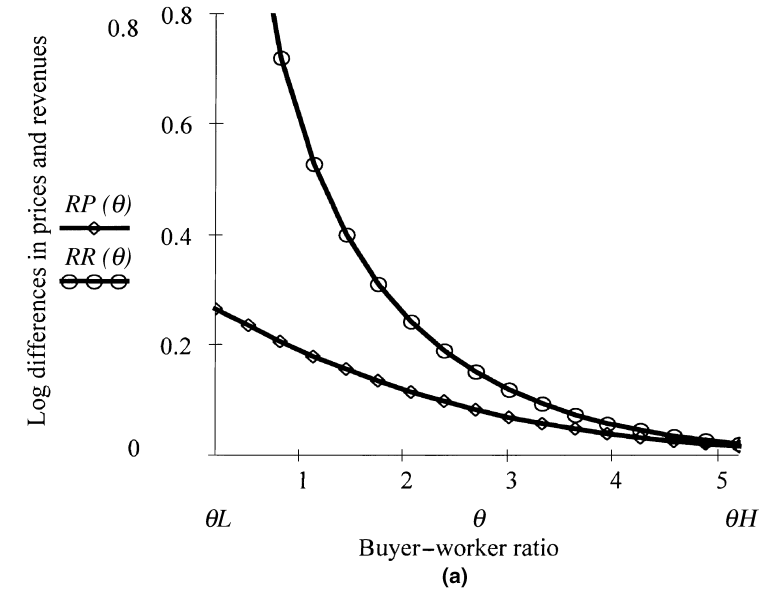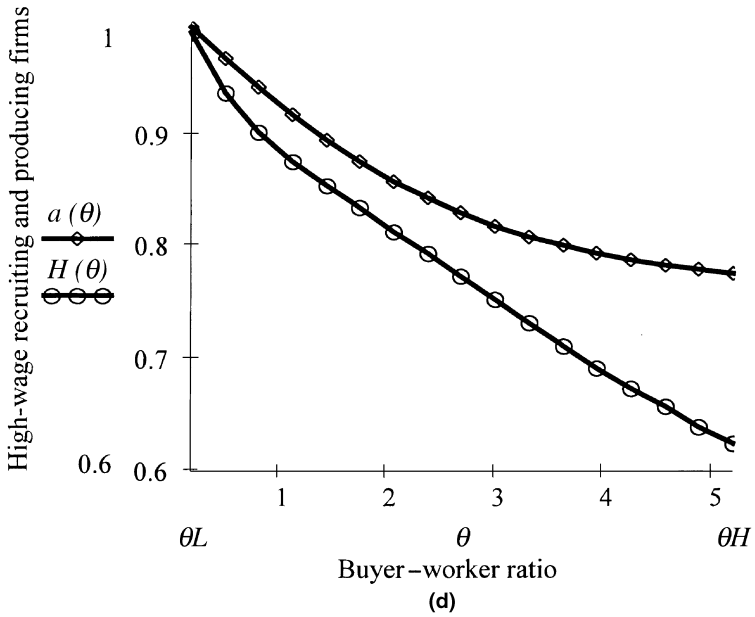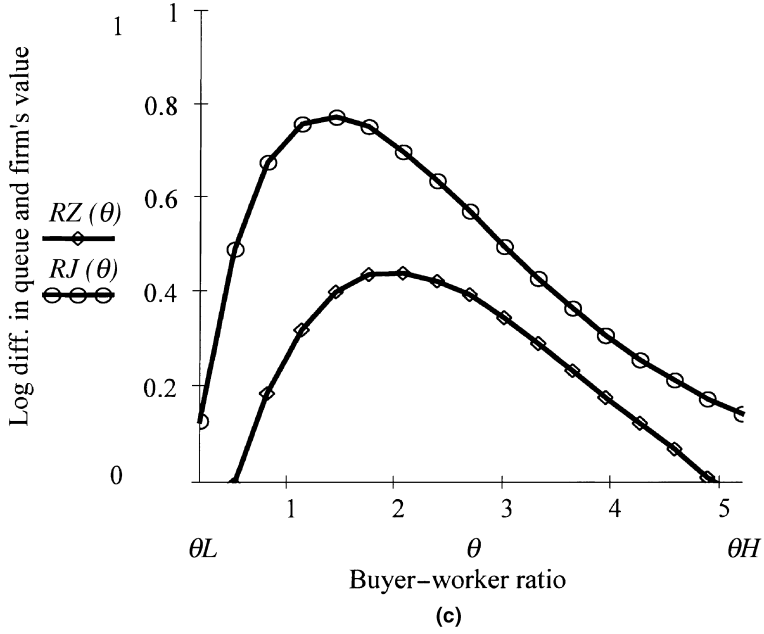
FIGURE 4a AND 4b (CONTINUED)

(c)

(d)

FIGURE 4c AND 4d

(a) SIZE–PRICE AND SIZE–REVENUE DIFFERENTIALS, (b) SIZE–WAGE DIFFERENTIAL,
(c) RELATIVE QUEUE LENGTHS AND PLANTS' VALUES, AND (d) SIZE DISTRIBUTION OF PLANTS

low-capacity plants'. Thus, the queue for a high-wage job increases relative to that for a low-wage job (Figure 4(c)) and so the size–wage differential increases (Figure 4(b)).[21]

The increase in the size–wage differential is feasible only when the number of buyers per worker is small. When the number of buyers continues to increase, the size–revenue differential falls. After $\theta > \theta^*$, this falling size–revenue differential dominates the effect of the size distribution and reduces the size–wage differential.

6. MODELING ASSUMPTIONS AND ROBUSTNESS

A coherent theme of my analysis is that agents make a trade-off between prices/wages and the probability of obtaining them. Holzer et al. (1991) present some evidence for this trade-off in the labor market. They found that minimum wage jobs attract more applicants than do jobs paying less than the minimum wage and that there is positive correlation between industry job application differentials and industry wage differential.[22] For the goods market, Carlton (1988) found that delivery lags are as important as prices in equating demand and supply and that demand elasticities with respect to prices and delivery time are both significantly negative.[23]

The main assumptions for my results are as follows:

(i) Agents cannot coordinate and there is a search cost. The coordination failure creates the possibility that buyers (job applicants) may fail to get the good (job) and so they can tolerate a high price (a low wage) if it is easy to obtain. The search cost, implicit in the assumption that a buyer (job applicant) can visit only one seller (employer) in a period, prevents agents from making instantaneous arbitrage between all matches. Both assumptions are standard in the search literature of unemployment (e.g., Pissarides, 1990) and are realistic for large markets.

More specifically, in a large market it might be very costly to communicate between all participants and between all matches. The coordination failure makes it reasonable to focus on the completely mixed-strategy equilibrium, as I did, where agents adopt symmetric strategies and buyers (job applicants) randomize over the offers. If agents could perfectly coordinate their decisions, there would be pure-strategy equilibria in which there is no apparent relationship between capacity and price.[24] Likewise, if buyers could visit all sellers at once, prices would have to be the same for all sellers.

---

[21] Another way to see this is to resort to (4.4), which indicates that the queue length for a low-wage job, z, is an increasing function of the fraction of high-wage recruiting plants in labor market, a, while the queue length for a high-wage job z_H is independent of a for given H. As a falls with the product demand, the relative queue length z_H/z can increase.

[22] Holzer et al. (1991) also found that minimum wage jobs attract more applicants than do jobs paying more than the minimum wage. This is likely because of the large difference between the applicants to the two types of jobs.

[23] Casual observations suggest that large retailers like Wal-Mart charge lower prices than smaller retailers. These are exceptions of the model since Wal-Mart also provides less service to customers and pays lower wages.

[24] See Burdett et al. (1997) for a characterization of pure-strategy equilibria in a similar environment.

Of course, the extreme version of these assumptions is not necessary for the results. In particular, I can allow a buyer to visit several sellers at once in a period, as long as the number of sellers visited by a buyer is a negligible fraction of the total number of sellers. In this case, a high capacity will still be useful for a seller to attract buyers and to support a positive size–revenue differential. But the equilibrium will be more difficult to characterize, since there will be price dispersion even among sellers of the same capacity, as illustrated by Lang (1991).[25] Similarly, by assuming that a job applicant can apply to only one employer in a period, I have simplified the exercise on the size–wage differential.

(ii) Another critical assumption is that plants grow slowly by recruiting one worker at a time. Given the size–revenue differential, this assumption ensures that a low-capacity plant has a higher marginal benefit from filling a vacancy than an entrant. If, instead, all recruiters can recruit two workers at the same cost, all plants will try to reach size two in one period (see Deneckere and Peck, 1995), although not all will be successful. Since an entrant is at least as eager to fill the vacancies as a low-capacity plant, there is no apparent relation between the plant's size and the wage in this case.

The above assumption is realistic. For example, it might be much more costly to recruit two workers at once than to recruit two workers sequentially. Alternatively, one may think that each worker uses one unit of capital and plants have to borrow to finance the purchase of capital. If the borrowing cost is higher for small plants than for large plants, then an entrant may find it attractive to start small and build up the collateral gradually in order to increase capital stocks.

One can even allow some but not all entrants to recruit two workers at a time. For example, one can model the opportunity to recruit two workers to be a random event obeying a Poisson process. When the arrival rate of this opportunity is small, a majority of entrants can recruit only one worker and so they will recruit at a lower wage than a low-capacity plant (when θ is moderate). In this extension, workers in different high-capacity plants get different wages depending on whether the plants are low-capacity plants that have succeeded in recruiting a second worker or entrants that have succeeded in recruiting two workers at one time.

Maintaining the above assumptions (i) and (ii), I now turn to auxiliary assumptions. First, I have deliberately abstracted from any other difference between plants such as product quality, capital intensity, and workers' skills. These differences introduce additional reasons why plants may have different profits and wages. By abstracting from them I underscore the importance of sellers' capacity differences for the size–wage differential.[26]

Second, the labor market opens before the goods market in each period. This timing sequence is natural since most plants first produce and then sell the product.

---

[25] Lang (1991) focuses on employers' ex post concerns on how to induce applicants to *accept* the offer. In particular, he does not model applicants' decisions on which employer to contact but, instead, assumes that each applicant receives two offers. In contrast, I focus on employers' ex ante concerns on how to attract workers to *apply* to the job in the first place, where workers' decisions on which jobs to apply to are essential.

[26] See Shi (1997) for an analysis of how skill differences may interact with the within-group wage differential in a wage-posting environment.

An alternative sequence is to have the goods market open first, where buyers and potential sellers trade forward contracts on goods. A low-capacity seller can contract two units and an entrant can contract one, both having a probability of default. Other than making the analysis complicated, this alternative sequence does not seem to overturn the analytical result. In particular, since a low-capacity seller can guarantee at least one good, its overall default probability is smaller than an entrant's. In this case, a low-capacity seller will charge a higher price on the forward contracts than an entrant and, to fulfill the contracts, it will post a higher wage to recruit than an entrant.

Third, the capacity constraint is two. For any arbitrary upper bound, the intuition for the size–revenue differential and the size–wage differential is valid, but analytical results like those in Sections 2 and 3 are possible only for an upper bound of two. Furthermore, one may want to endogenize the upper bound by assuming, for example, that the marginal cost of capital increases sufficiently or the marginal productivity of labor diminishes sufficiently when a plant increases its employment. In this extension, the size–revenue differential may fall below zero for a sufficiently high level of production and so plants will never want to extend employment beyond this level. Fixing θ at a moderate value, the size–wage differential first increases and then decreases with the size, as in the model here.

Fourth, a plant automatically increases the existing worker's wage to match the new recruit's wage. I used this assumption to prevent a negative experience premium. Although one could eliminate the assumption and examine how a negative experience premium interacts with the size–wage differential, such an exercise rewards little. The reason is that the model also assumed unrealistically that employed workers do not search. If employed workers can search as well, as in Burdett and Mortensen (1998), an employed worker will get a wage that is at least as high as the new recruit's. It is interesting but difficult to analyze on-the-job search in the current model.

Finally, there is no long-term relationship between agents. The model's predictions would be robust to long-term relationships if a significant fraction of these relationships are destroyed each period and replaced by new ones. For example, the large job separation in the U.S. labor market destroys the relationships between employers and employees. With large separation, a significant fraction of plants do not have any contractual relations with buyers or workers and so behave in the way modeled here.

7. CONCLUSION

I have integrated the product market and the labor market into a directed search framework and shown that although plants are identical except size, larger plants can pay higher wages to homogeneous workers and earn higher expected profit per worker. A large plant charges a higher price and uses the larger capacity to compensate buyers with a higher service probability. This strategy yields a higher expected revenue per worker for a seller and, to capture this size-related benefit, large plants try to become larger by posting higher wages to recruit than small plants. Labor market competition does not eliminate this size–wage differential because job applicants are indifferent between a high wage that is hard to get and a low wage that is easy to get. The coexistence of the size–revenue differential and the size–wage differential is consistent with the empirical finding in Katz and Summers (1989).

My analysis indicates that two seller-specific variables are important for explaining the size–wage differential—the plant's delivery time in the product market and the queue for the plant's job openings. The size–wage differential depends negatively on the delivery time and positively on the queue length of job applicants. For the size–wage differential, these variables are more relevant measures of the seller's market power than traditional ones like a high concentration ratio and a small price elasticity. By traditional measures, the concentration ratio is zero in my model and each seller has zero market power, since there are infinitely many sellers. But this does not imply that sellers cannot adjust prices to affect their sales, since there are also infinitely many buyers. For price elasticity, both large and small sellers face elastic demand. There is not much difference between these sellers' price elasticities, since the price differential between large and small sellers is relatively flat in response to an increase in the product demand (Figure 4(a)). In contrast, the difference in the queue lengths (or the delivery lags) between large and small sellers responds to the product demand in a large magnitude, generating a large response in the size–revenue differential. Not surprisingly, Brown and Medoff (1989) found that traditional measures of the product market power explain little of the size–wage differential.

An important conclusion of my analysis is that the size distribution of plants in the industry interacts with the size–revenue differential to determine the size–wage differential. An increase in the product demand increases the size–wage differential only when it increases the fraction of small plants sufficiently, which occurs when the product demand is initially low. When the product demand is already high, the size–revenue differential falls sufficiently in response to the higher demand and dominates the effect of the size distribution of plants. This result sheds new light on how trade liberalization affects wage inequality, since the size–wage differential is a significant component of overall wage inequality. For example, when trade agreements give a country the access to a foreign product market, wage inequality increases only when there is a large shift of employment to small plants. If one fixes the employment distribution across plants of different sizes (e.g., Davis and Haltiwanger, 1991), one may be omitting an important effect of trade liberalization on wage inequality.

APPENDIX

A.1. *Proof of Lemma* 1. Differentiating $F2(x)$ in (2.12), I can show that $F2' < 0$ if and only if

$$\ln(1 + x) - \frac{x(e^x - 1)}{2(e^x - 1) - x} < 0$$

Temporarily denote the left-hand side of the above inequality as $f(x)$. Applying L'Hôpital's rule, I can show $f(0) = 0$. If $f'(x) < 0$ for all $x > 0$, then indeed $f(x) < f(0) = 0$ for all $x > 0$. Computing f' and noting $e^x - 1 - x - x^2/2 > 0$, I find $f' < 0$ if and only if $1 - x - e^{-x} < 0$. Since the function $(1 - x - e^{-x})$ has a value 0 when $x = 0$ and a negative derivative for all $x > 0$, it is indeed negative for all $x > 0$. Thus, $f' < 0$ for all $x > 0$, as desired.

To show $\phi'(x) < 0$, verify that $\phi' < 0$ iff $A1(x) + A2(x)\ln(1+x) < 0$, where

$$A1(x) \equiv [(x-1)e^x - 1][e^x - (1+x)^2][e^x - 1 - x - x^2/2]$$
$$- x^2(e^x - 1)(e^x - 1 - x)(e^x - 1 - x/2)$$
$$A2(x) \equiv (1+x)(e^x - 1 - x)[(x^2 - x - 1)e^x + 1 + 2x + x^2/2]$$
$$+ 2x(e^x - 1 - x/2)(e^x - 1 - x - x^2/2)$$

First, I show $A2(x) > 0$ for all $x > 0$, which amounts to $[(x^2 - x - 1)e^x + 1 + 2x + x^2/2] > 0$ for all $x > 0$. The latter expression is zero when $x = 0$. Its derivative has the same sign as that of $[x - 2 + x^2 + (2+x)e^{-x}]$, which is positive since $e^{-x} > 1 - x$. Thus, the expression is positive for all $x > 0$, implying $A2(x) > 0$.

With $A2(x) > 0$ and $\ln(1+x) > 0$, $\phi' < 0$ if and only if $\ln(1+x) + A1(x)/A2(x) < 0$. Notice that $A1^{(j)}(0) = A2^{(j)}(0) = 0$ for $j = 0, 1, \ldots, 4$, where (j) stands for the order of derivatives. Also, $A1^{(5)}(0) = 0$ but $A2^{(5)}(0) > 0$. Applying L'Hôpital's rule five times, I can show $A1(x)/A2(x) \to 0$ when $x \to 0$. Since $\ln(1+x) \to 0$ as well when $x \to 0$, the function $\ln(1+x) + A1(x)/A2(x)$ has a value 0 when $x \to 0$. Then, $\phi' < 0$ if this function has a negative derivative for all $x > 0$; i.e., if

$$(1+x)[A1'(x)A2(x) - A2'(x)A1(x)] + A2^2(x) < 0$$

Temporarily denote the left-hand side of the above inequality as $f_0(x)$. Through the following steps, I can establish $f_0(x) < 0$ for all $x > 0$:

   (i) Start with $i = 0$.
   (ii) Verify $f_i(0) \le 0$.
   (iii) Arrange $f_i(x)$ by separating terms that are multiplied by e^x (or power terms of e^x) and terms that are not multiplied by e^x. The terms that are not multiplied by e^x form a polynomial of x. Let the highest order of x in this polynomial be I_i.
   (iv) Verify $f_i^{(j)}(0) \le 0$ for all $j = 0, 1, 2, \ldots, (I_i + 1)$, where (j) is the order of derivatives.
   (v) Define $f_{(i+1)}(x) = e^{-x} f_i^{(I_i + 1)}(x)$.
   (vi) Replace i by $i + 1$, repeat steps (i)–(v) to obtain $f_5(x)$.
   (vii) $f_5(x) = f_{5m}(x) - f_{5n}(x)e^x$, where $f_{5m}(x)$ and $f_{5n}(x)$ are polynomials of x. Also $f_{5n}(x) > 0$ for all $x > 0$. Substituting $e^x > 1 + x + x^2/2$, I can then show $f_5(x) < 0$ and so $f_4^{(I_4+1)}(x) < 0$ by step (v). Since $f_4^{(I_4)}(0) \le 0$ by step (iv), then $f_4^{(I_4)}(x) < f_4^{(I_4)}(0) = 0$. Recursively, I can show $f_4(x) < 0$. Repeating this process, I can show $f_0(x) < 0$ for all $x > 0$. ∎

A.2. *Proofs of Lemma* 2 *and Proposition* 4. To prove Lemma 2, note that the left-hand side of (3.6) is a decreasing function of z and the right-hand side is an increasing function. When $z \to 0$, the left-hand side is greater than 1 and the right-hand side is 1; when $z \to au/[(1-a)(1-u)]$, the left-hand side is less than 1 and the right-hand side is greater than 1. Thus, there is a unique positive solution, z_1, to (3.6) and z_1 lies in $(0, au/(1-a)(1-u))$. The properties (i)–(iii) in the lemma are evident.

To prove Proposition 4, it suffices to show that there is a solution for z to the equation $F1(z) = F2(x(z))$ and that the solution lies in $(0, z_1)$. Since $F1(z_1) = 1$ and

$F2(x(z_1)) > F2(\infty) = 1$ ($F2'(x) < 0$ by Lemma 1), in turn it suffices to show that $F2(x(0)) < F1(0)$, where $x(0)$ is defined by (3.4) with $z = 0$. Lemma A.1 below shows that $x(0)$ is an increasing function of θ, with $x(0) = 0$ when $\theta = 0$ and $x(0) = \infty$ when $\theta = \infty$. Since $F2(x)$ is a decreasing function of x (Lemma 1), $F2(x(0))$ is a decreasing function of θ. Moreover,

$$F2(x(0))|_{\theta=0} = F2(0) = \infty > F1(0)$$

$$F2(x(0))|_{\theta=\infty} = F2(\infty) = 1 < F1(0)$$

Thus, there exists $\theta_0 > 0$ such that $F2(x(0)) < F1(0)$ if and only if $\theta > \theta_0$. That is, if $\theta > \theta_0$, there exists at least one solution for z to $F2(x(z)) = F1(z)$. ∎

LEMMA A.1. *The function $x(z)$, defined by (3.4), has the following features*: (i) *For any given $z \in (0, au/[(1 - a)(1 - u)])$, $x(z)$ is an increasing function of θ*; (ii) *$x(z) = 0$ when $\theta = 0$ and $x(z) = \infty$ when $\theta = \infty$.*

PROOF. By Proposition 1, the solution to (3.4), $x(z)$, exists and is unique, provided $H \in (0, 1)$. The requirement $H \in (0, 1)$ is equivalent to z, $z_H > 0$ and hence to $z \in (0, au/[(1 - a)(1 - u)])$. Note that the right-hand side of (3.4) crosses the left-hand side from below when x increases from below $x(z)$ to above $x(z)$. Since, for given x, the right-hand side of (3.4) is a decreasing function of θ, $x(z)$ is an increasing function of θ. With (3.4), it is also easy to verify that $x(z) = 0$ when $\theta = 0$ and $x(z) = \infty$ when $\theta = \infty$. This completes the proof of Lemma A.1. ∎

A.3. *Proof of Proposition 5.* From Proposition 1, $w_H > w_L$ if and only if $z_H > z$; i.e., if and only if $z < \xi \equiv au/(1 - u)$. In light of Figure 3, this is equivalent to $F2(x(\xi)) > F1(\xi)$, where $x(\xi)$ is defined by setting $z = \xi$ in (3.4):

(A.1) $$\ln(1 + x) = \frac{(e^\xi - 1 + a)x - \frac{a\theta}{1-u}e^\xi}{(1 - a)e^\xi + 2a - 1}$$

Lemma A.1 above shows that $x(\xi)$ is an increasing function of θ, with $x(\xi) = 0$ when $\theta = 0$ and $x(\xi) = \infty$ when $\theta = \infty$. Then, $F2(x(\xi))$ is a decreasing function of θ, with the following features:

$$F2(x(\xi))|_{\theta=0} = F2(0) = \infty > F1(\xi)$$

$$F2(x(\xi))|_{\theta=\infty} = F2(\infty) = 1 < F1(\xi)$$

Thus, there exists $\theta_1 > 0$ such that $F2(x(\xi)) > F1(\xi)$ (i.e., $w_H > w_L$) if and only if $\theta < \theta_1$.

I now show that $\theta_1 > \theta_0$ when a is either close to 1 or close to 0. When $a = 1$, (A.1) becomes identical to (3.4) with $z = 0$. That is, $x(\xi)|_{a=1} = x(0)$. Since $F1(\xi)|_{a=1} < F1(0)$, then

$$[F2(x(\xi)) - F1(\xi)]_{a=1} = F2(x(0)) - F1(\xi)|_{a=1} > F2(x(0)) - F1(0)$$

By the definitions of θ_0 and θ_1, I have

$$[F2(x(\xi)) - F1(\xi)]_{a=1,\theta=\theta_0} > [F2(x(0)) - F1(0)]_{\theta=\theta_0}$$
$$= 0$$
$$= [F2(x(\xi)) - F1(\xi)]_{a=1,\theta=\theta_1}$$

Since $[F2(x(\xi)) - F1(\xi)]_{a=1}$ is a decreasing function of θ, this implies $\theta_1 > \theta_0$ when $a = 1$. By continuity, $\theta_1 > \theta_0$ when a is close to 1.

When $a \to 0$, $\xi \to 0$ and so $(e^\xi - 1)/a \to u/(1-u)$. Equation (A.1) becomes $\ln(1+x) = x - \theta$ and so $x(\xi)|_{a\to 0}$ is a continuous function of θ, which has a value 0 when $\theta = 0$ and a value ∞ when $\theta = \infty$. Note that $F1(\xi)|_{a\to 0} = 1 = F2(\infty)$. Since θ_1 solves $F2(x(\xi)) = F1(\xi)$, when $a \to 0$ the value θ_1 must be such that $x(\xi)|_{a\to 0} = \infty$; i.e., $\theta_1 = \infty$. In contrast, $F1(0)$ and $F2(x(0))$ do not depend on a and so $\theta_0 < \infty$ when $a \to 0$. Thus, $\theta_1 > \theta_0$ when a is close to 0. ∎

A.4. *Proofs of Lemma* 3 *and Proposition* 6. For Lemma 3, use (4.6) to show that $V_H > V_u$ if and only if $z_H \in (0, \infty)$. Similarly, $V_L > V_u$ if and only if $z \in (0, \infty)$. Thus, (vi) is always satisfied in a mixed-strategy equilibrium. To show that $J_R > J_L$ whenever $z_H \geq z$, set $J_N = 0$ in (4.20) and use (4.19) to eliminate c. I get

$$J_R - J_L = (1 - \beta)V_u[2(e^{z_H} - 1 - z_H) - (e^z - 1 - z)]$$

Since $e^z - 1 - z$ is an increasing function of z for $z > 0$, clearly $z_H \geq z$ implies $J_R > J_L$.

For Proposition 6, I first show that $z_H > z \Leftrightarrow V_H > V_L$. Use (4.6) to write

$$V_H = \left[\beta + \frac{(1-\beta)z_H}{1 - e^{-z_H}}\right]V_u, \qquad V_L = \left[\beta + \frac{(1-\beta)z}{1 - e^{-z}}\right]V_u$$

Since $V_u > 0$ by (4.21) and since $z/(1 - e^{-z})$ is an increasing function of z for $z > 0$, $V_H > V_L$ if and only if $z_H > z$. To show $V_H > V_L \Leftrightarrow w_H > w_L$, retrieve the wage rates (w_H, w_L) from the value functions in (4.7) and (4.8) and substitute V_R from (4.9). I have

$$w_H - w_L = [1 - \beta(1 - 2\sigma)e^{-z_H}](V_H - V_L)$$

Since $z_H \in (0, \infty)$, $\beta(1 - 2\sigma)e^{-z_H} < 1$ and hence $w_H > w_L$ if and only if $V_H > V_L$.

To show that $z_H > z \Leftrightarrow J_H > 3J_L$, set $(w_H^d, z_H^d) = (w_H, z_H)$ in (4.17) and $(w_L^d, z^d) = (w_L, z)$ in (4.18). Subtracting the two equations and noting $J_N = 0$, I have

$$\frac{J_H - J_L}{2} - J_L = (1 - \beta)V_u\left[\left(e^{z_H} - \frac{z_H}{1 - e^{-z_H}}\right) - \left(e^z - \frac{z}{1 - e^{-z}}\right)\right]$$

Since $[e^z - z/(1 - e^{-z})]$ is an increasing function of z, $z_H > z$ iff $J_H > 3J_L$.

Finally, I show that $w_H > w_L$ only if $R_H > R_L$; i.e., $w_H > w_L \Rightarrow R_H > R_L$. To do so, retrieve the wage rates (w_H, w_L) from (4.12) and (4.13) and subtract them to obtain

$$R_H - R_L = (w_H - w_L) + \left\{[1 - \beta(1 - 2\sigma)]\frac{J_H}{2} + \beta(1 - 2\sigma)J_R - J_L\right\}$$

Clearly, the desired result follows if

$$w_H > w_L \Rightarrow [1 - \beta(1 - 2\sigma)]\frac{J_H}{2} + \beta(1 - 2\sigma)J_R > J_L$$

Recall that $w_H > w_L \Leftrightarrow z_H > z \Leftrightarrow J_H > 3J_L$. Also, Lemma 3 shows that $z_H > z \Rightarrow J_R > J_L$. Thus, $w_H > w_L$ implies

$$[1 - \beta(1 - 2\sigma)]\frac{J_H}{2} + \beta(1 - 2\sigma)J_R > [1 - \beta(1 - 2\sigma)]\frac{3J_L}{2} + \beta(1 - 2\sigma)J_L > J_L$$

This completes the proof of the proposition. ∎

REFERENCES

BLANCHARD, O., AND P. DIAMOND, "The Beveridge Curve," *Brookings Papers on Economic Activity* 1 (1989), 1–60.

BROWN, C., AND A. MEDOFF, "The Employer Size–Wage Effect," *Journal of Political Economy* 97 (1989), 1027–59.

BURDETT, K., AND D. T. MORTENSEN, "Wage Differential, Employer Size, and Unemployment," *International Economic Review* 39 (1998), 257–73.

———, S. SHI, AND R. WRIGHT, "Pricing with Frictions," unpublished manuscript, University of Pennsylvania, 1997.

CAO, M., AND S. SHI, "Coordination, Matching, and Wages," *Canadian Journal of Economics* 33 (2000), 1009–33.

CARLTON, D., "Market Behavior with Demand Uncertainty and Price Flexibility," *American Economic Review* 68 (1978), 571–88.

———, "The Theory and Facts of How Markets Clear: Is Industrial Organization Valuable for Understanding Macroeconomics?" in R. Schmalensee and R. Willig, eds., *Handbook of Industrial Organization*, Vol. 1 (Amsterdam: North-Holland, 1988), 909–46.

DAVIS, S., AND J. HALTIWANGER, "Wage Dispersion between and within U.S. Manufacturing Plants, 1963–86," *Brookings Papers: Microeconomics* (1991), 115–200.

DENECKERE, R., AND J. PECK, "Competition over Price and Service Rate When Demand Is Stochastic: A Strategic Analysis," *RAND Journal of Economics* 26 (1995), 148–62.

HAMERMESH, D. S., "Commentary," in J. J. Siegfried, ed., *The Economics of Firm Size, Market Structure, and Social Performance* (Washington, DC: Federal Trade Commission, 1980).

———, *Labor Demand* (New Jersey: Princeton University Press, 1993).

HOLZER, H. J., L. F. KATZ, AND A. B. KRUEGER, "Job Queues and Wages," *Quarterly Journal of Economics* 106 (1991), 739–68.

JUHN, C., K. MURPHY, AND B. PIERCE, "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy* 101 (1993), 410–42.

KATZ, L. F., AND L. H. SUMMERS, "Industry Rents: Evidence and Implications," *Brookings Papers: Microeconomics* (1989), 209–90.

LANG, K., "Persistent Wage Dispersion and Involuntary Unemployment," *Quarterly Journal of Economics* 106 (1991), 181–202.

MCMILLAN, J., AND M. ROTHSCHILD, "Search," in R. Aumann and S. Hart, eds., *Handbook of Game Theory*, Vol. 2 (Amsterdam: North-Holland, 1994), 905–27.

MOEN, E. R., "Competitive Search Equilibrium," *Journal of Political Economy* 105 (1997), 385–411.

MONTGOMERY, J. D., "Equilibrium Wage Dispersion and Interindustry Wage Differentials," *Quarterly Journal of Economics* 106 (1991), 163–79.

PETERS, M., "Bertrand Equilibrium with Capacity Constraints and Restricted Mobility," *Econometrica* 52 (1984), 1117–27.

———, "Ex Ante Price Offers in Matching Games: Non-Steady State," *Econometrica* 59 (1991), 1425–54.

PISSARIDES, C., *Equilibrium Unemployment Theory* (Cambridge, MA: Basil Blackwell, 1990).

SHAPIRO, C., AND J. STIGLITZ, "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review* 74 (1984), 433–44.

SHI, S., "Unskilled Workers in an Economy with Skill-Biased Technology," unpublished manuscript, Queen's University (at http://qed.econ.queensu.ca/pub/faculty/shi/), 1997.

Stigler, G. J., ''Information in the Labor Market,'' *Journal of Political Economy* 70 (1962), 94–105.

Troske, K. R., ''Evidence on the Employer Size–Wage Premium from Worker-Establishment Matched Data,'' *Review of Economics and Statistics* 81 (1999), 1–12.

Weiss, L. W., ''Concentration and Labor Earnings,'' *American Economic Review* 56 (1966), 96–117.