University of Toronto
Department of Economics

Working Paper 429

Customer Relationship and Sales

By Shouyong Shi

May 11, 2011

Customer Relationship and Sales

Shouyong Shi*
Department of Economics
University of Toronto
(shouyong@chass.utoronto.ca)

This version: May 10, 2011

Abstract

I construct a search model to formalize the intuitive idea that sellers hold sales to attract buyers and build customer relationships. The market consists of a large number of buyers and sellers. All sellers sell a homogeneous good and all buyers have the same publicly known valuation of the good. Buyers know the terms of trade offered by sellers before choosing which seller to visit. A buyer is related to a seller if the buyer just bought a good from the seller and the relationship is broken if the buyer fails to continue to buy from the seller. Sellers are restricted to offer the same price to all buyers, but they are allowed to give priority to their related buyers. I prove that there is an equilibrium in which a seller gives priority to the related buyer and a buyer makes repeat purchases from the related seller. In the equilibrium, a seller who does not have a related buyer posts a low (sale) price to attract the buyers who are unrelated to any seller and, once the seller is related to a buyer after a trade, the seller will post a high (regular) price to sell only to the related buyer. The fraction of related sellers is endogenous in the equilibrium. I calibrate the steady state of the model to the data and find that the sale price represents a sizable markdown, and the regular price a sizable markup, on the marginal cost. With the calibrated model, I examine comparative statics and dynamics of the equilibrium with respect to changes in the cost of and the demand for the good.

JEL classifications: D83; D40; E30.
Keywords: Sales; Customer relationship; Directed search.

---

* Address: 150 St. George Street, Toronto, Ontario, Canada, M5S 3G7. I thank Oleksiy Kryvtsov for useful conversations and for providing some references. I would like to acknowledge financial support from Canada Research Chair, the Bank of Canada Fellowship and the Social Sciences and Humanities Research Council of Canada. The opinion expressed here is my own and does not reflect the view of the Bank of Canada.

1. Introduction

Sales are common in the retail market. In the microdata collected monthly by the US Bureau of Labor Statistics on goods and services that cover 70% of consumer expenditure, about 11% of all price quotes are sales (see Klenow and Kryvtsov, 2008, and Nakamura and Steinsson, 2008). Sales represent a significant fraction of the size and the frequency of price changes. The average sale price in the BLS microdata is about 25% to 30% off the regular price, depending on the estimation method. Excluding sales increases the median duration of prices from 3.7 months to 7.2 months (Klenow and Kryvtsov, 2008). Moreover, sale prices seem to respond to shocks differently from regular prices. These empirical regularities of sales pose a challenge for macro models that rely on price rigidity to explain aggregate propagation, because most of those models do not incorporate sales. On the other hand, there is a sizable literature on sales in industrial organization (see a brief review later) that is not designed to explain the regularities of sales in large markets. In this paper, I formulate a theory of sales in a large market and calibrate the model to the data to investigate the model's comparative statics and dynamics.

The theory formalizes the intuitive idea that sellers hold sales to attract buyers and build customer relationships. Consider a market with a large number of buyers and sellers, where sellers sell one indivisible unit of a homogeneous good and buyers have the same publicly known valuation of the good. In each period, buyers observe the terms of trade offered by sellers before choosing which seller to visit. Suppose that buyers cannot coordinate on their visiting decisions and they cannot switch sellers within the same period. Because of these frictions, a seller may fail to get a visitor in a period and a buyer may fail to be selected by the seller he visits. In this environment, there is a tradeoff between the price of the good and the probability of trade. Individuals on the two sides of the market may want to form relationships in order to increase the probability of trade. Suppose that a seller treats a buyer who bought from him in the previous period as the related buyer and can reward the relationship with the priority of trade over unrelated buyers. However, to conform with the definition (e.g., in the BLS data) that a sale is a price cut available to all buyers, suppose that a seller is not allowed to offer one price to the related buyer and another price to unrelated buyers.[1]

[1] In reality, sellers may offer price discounts through frequent customer memberships. However, ab-

I characterize the equilibrium in which it is optimal for a seller to give priority to the related buyer and optimal for a buyer to make repeat purchases from the related seller. Like most relationships, the informal relationship described in this model generates benefits to the partners intertemporally and has implications on prices. For a buyer, forming the relationship with a seller enables him to have a higher probability of trade in the future with the related seller. Recognizing this benefit to the buyer, a seller who has a related buyer will charge a high (regular) price on the good that will make the related buyer just indifferent between visiting him and visiting an unrelated seller. Anticipating this high price as the return to a relationship, a seller who does not have a related buyer posts a low (sale) price in order to attract customers and acquire a relationship. The equilibrium has complete separation between related and unrelated individuals: while related buyers only visit their related sellers and pay the regular price, unrelated buyers only visit sellers who do not have related buyers and pay the sale price.

The fraction of related sellers is endogenous. A relationship is broken if the buyer fails to visit the related seller. To avoid the trivial case where all individuals on the shorter side of the market are related, I introduce exogenous separation from a relationship in the form of a taste shock that makes a buyer inactive in a period. Thus, a buyer may fail to show up at the related seller, in which case the relationship is destroyed. This flow of sellers out of relationships and the opposite flow of unrelated sellers who acquire relationships through trade determine the dynamics of the fraction of related sellers.

Because the fraction of sellers who are unrelated is always positive, the fraction of sellers who hold sales in any given period is positive. Similarly, because a seller has a positive probability of losing a relationship, each seller holds sales regularly over time. The model has precise predictions on the frequency, the duration and the price discount of sales. These features of sales depend differently on the extensive and intensive margins of the market. The extensive margin of the demand has two dimensions: the buyer/seller ratio and the probability that a buyer is active in a period. The intensive margin of the market consists of the utility of consuming a good and the marginal cost of a good. In the steady state, the

stracting from price discrimination enables me not only to focus on sales as price cuts available to all buyers, but also to make the results on sales more robust. If a seller is allowed to use price to discriminate buyers, it is easy to generate a differential between the price posted by a seller with a related buyer and the price posted by a seller without a related buyer. For search models of the labor market that allow for firms to post wages contingent on the worker's type, see Shi (2002, 2006).

frequency of sales, the duration of a sale and the fraction of related sellers in the market depend only on the extensive margin and not on the intensive margin of the market. In contrast, prices, markups and price discounts of sales depend on both margins.

To investigate the predictions of the model more specifically, I calibrate the steady state of the model to the data. The identified model reveals that the sale price is below the marginal cost, i.e., a markdown on the marginal cost, while the regular price is a positive markup on the marginal cost. The markdown is 18.9%, the markup is 12.7%, and the average markup weighted by the transaction frequency is 11.3%. Thus, customer relationships induce large variations in retail prices.

With the identified model, I examine first comparative statics and then dynamics of the equilibrium with respect to changes in the parameters. Relegating the details to subsections 4.2 and 4.3, I mention some of the results here. First, the duration of a sale reflects changes in market conditions more accurately than do price-related variables such as markups and price discounts of sales. While markups and price discounts can respond to market conditions non-monotonically, the duration of a sale always responds monotonically. Second, the two dimensions of the extensive margin of the demand affect prices differently. An increase in the demand arising from a higher buyer/seller ratio increases the average price of the good, but an increases in the demand arising from a higher probability that a buyer is active in the market can reduce the price, which may explain the paradoxical finding in the data by Chevalier et al. (2003). Third, an increase in the marginal cost of the good induces relatively large and non-linear reductions in the markdown and the markup, thus narrowing the gap between the regular price and the sale price. The reduction in the markdown is larger in percentage terms than in the markup, indicating that the regular price is more stable than the sale price. Fourth, in the short run, a shock to the demand typically induces prices and the duration of a sale to overshoot the steady state because the fraction of related sellers (as an aggregate state variable) is temporarily outside the steady state. However, this fraction adjusts relatively quickly toward the steady state. Finally, I examine how free entry of sellers affects the equilibrium by affecting the buyer/seller ratio.

In this model, search is directed in the sense that buyers know the terms of trade before choosing which seller to visit. Thus, the model is related to the growing literature on directed search. This literature is divided into two approaches. One assumes that a market is organized in many submarkets each offering specific terms of trade and that an exogenous matching function determines the number of matches inside each submarket (e.g., Montgomery, 1991, Moen, 1997, Acemoglu and Shimer, 1999, and Shi, 2009). The other approach derives the matching function endogenously from the equilibrium of individuals' strategies (e.g., Peters, 1991, and Burdett et al., 2001). In this paper I follow the second approach in order to capture explicitly how a seller's decision on whether to give priority to the related buyer affects the seller's trading probability. To my knowledge, it is new to use a directed search model to explain customer relationship and sales.

The literature on sales has a few strands. The first strand relies on the result that when sellers face a discontinuous demand curve, there is an equilibrium in which each seller follows a mixed strategy to determine the posted price (Shilony, 1977). Extending this result, some authors interpret price reductions in the mixed strategy as sales. For example, Salop and Stiglitz (1982) introduce durability of the product and Varian (1980) introduces heterogeneity in buyers' valuation. Interesting as it is, the interpretation of price variations in a mixed strategy as sales seems too dubious to match the regularity and the duration of sales in the data. Moreover, the mixed strategy may be purified when there are a large number of sellers and buyers, as in Burdett and Judd (1983). The second strand of the literature is based on the elegant model by Sobel (1984). In Sobel's model, there is a constant flow of buyers entering the market in each period who are either high-valuation buyers local to particular sellers or low-valuation shoppers who can patiently wait for sales. As the stock of patient shoppers builds up due to buyers' entry, some sellers cut prices to clear this stock of patient shoppers. This model has precise predictions on the frequency and the price discount of sales, but the duration of a sale is one period. Related to but different from this model is the one by Lazear (1986), where sellers hold sales to clear the inventory of goods rather than a stock of patient shoppers.[2] Broadening this category of models even further, one may include the so-called sS models in which sellers use sales to manage inventory and/or prices (e.g., Slade, 1998, Aguirregabiria, 1999). The third strand of the literature on sales contains signaling models in which a seller uses promotional sales to signal either the quality of the product (e.g., Milgrom and Roberts, 1986) or the cost of

[2]In Lazear's model, a seller does not know the value of the good to the buyers and learns about the distribution of such valuation from whether a good was sold at the previously posted price. Sales are an outcome of downward updating of the seller's beliefs on buyers' valuation of the good. Gonzalez and Shi (2010) integrate a similar learning process into a search equilibrium of a large labor market.

the product (e.g., Bagwell, 1987). Although inventory management and promotional sales are important for some sales, many other sales are not motivated by these considerations. In particular, promotional sales are likely to dissipate over time once buyers have learned about the product, but sales in the data (e.g., Klenow and Kryvtsov, 2008) can occur on the same product regularly over time. Finally, most of these models cannot generate markdowns observed in the data (e.g., Dutta et al., 2002).

In contrast to this literature, I emphasize customer relationships as a motive for sales. There are two main reasons for constructing such a theory. First, it is common sense that sellers use sales to attract buyers and build customer relationships. The marketing literature (e.g., Blattberg and Sen, 1974) emphasizes the importance of customer loyalty but takes customer loyalty as a primitive of the model. In my model, there is nothing hard-wired about customer loyalty; instead, the relationship is endogenously generated through trade. A buyer values the relationship only if the related seller offers priority in trade, while a seller values the relationship only if he can charge a higher price than an unrelated seller does. Second, the above models on sales typically have a few sellers and/or buyers rather than a large retail market. Some of the models are also difficult to be made dynamic. In contrast, my model has an infinite horizon and many (in fact, infinitely many) sellers and buyers. These features make the model promising for macro analyses, although more work is needed to enrich the model. To focus on the link between customer relationships and sales, I deliberately abstract from some elements that are important in the above literature on sales, such as durability of the good, heterogeneity in buyers' preferences, and private information in buyers' valuation or the quality/cost of the product.

2. A Model of Directed Search with Customer Relationship

2.1. The model environment

Time is discrete and lasts forever. There are M sellers and N buyers, where M and N are large numbers which I will take to the limit ∞. Denote $b = N/M$ as the buyer-seller ratio and fix $b \in (0, \infty)$ until section 5 where I will allow for free entry of sellers to determine b. All individuals discount future at a rate $r > 0$. In each period, a seller can produce one indivisible unit of good, and the cost of producing a unit is $c \geq 0$. Because goods are perishable, however, a seller produces a good only after meeting a buyer. In any given period, a buyer needs to consume a good with probability $\lambda \in (0, 1)$ and does not have such a need with probability $(1 - \lambda)$. These taste shocks occur at the beginning of the period and are independently and identically distributed among the buyers and across time. Call a buyer who has the need to consume an active buyer, and a buyer who does not have the need an inactive buyer. In each period, the number of active buyers is λN. The utility of consumption is U $(> c)$ to an active buyer. Because the number of active buyers per seller is $b\lambda$, I refer to $b\lambda$ as the *extensive margin* of the demand, as opposed to the *intensive margin* of the market which consists of U and c.

Although all sellers are identical in their production capacity and cost, they might have had different outcomes of selling in the previous period. If a seller sold a good in the previous period to a buyer, the two individuals are related to each other. If a seller did not sell a good in the previous period, the seller is unrelated. I focus on the *equilibrium with priority* where sellers give priority to their related buyers; that is, if the related buyer and some unrelated buyers both visit a seller, the seller sells the good to the related buyer first. I will find a condition under which this choice of priority is optimal for a seller.

The fraction of sellers who have related buyers is denoted ρ, which is an aggregate state variable. Because each related seller has one and only one related buyer, the number of buyers who have related sellers is equal to ρM, and the number of active buyers who have a related seller is $\lambda \rho M$. The number of buyers who do not have related sellers is $N - \rho M$ and the number of active buyers who do not have related sellers is $\lambda(N - \rho M)$. I will verify later that $\rho < \min\{1, b\}$ in the equilibrium; that is, some sellers are unrelated to any buyer and some buyers are unrelated to any seller.

In each period, sellers simultaneously post the terms of trade. After observing the terms of trade, active buyers choose which seller to visit. Sellers must post the terms of trade before knowing which buyer is active in the period, and each buyer must make the visiting decision without knowing other buyers' choices. Moreover, a seller must sell a good for the same price independently of which type of buyer comes to him; that is, a seller is not allowed to use price to discriminate different buyers. As explained in the introduction, this assumption is intended to capture the fact that a sale is available to all buyers. However, a seller is allowed to give priority to the related buyer. If all visitors are unrelated to the seller, the seller randomly selects one to trade with. Because a seller can give priority to the related buyer, a seller's choice of price may depend on whether or not the seller has a related buyer, i.e., whether or not the seller sold a good in the previous period.

A relationship between a buyer and a seller is informal, and the buyers are free to shop at any seller. Moreover, a relationship ends when the buyer fails to visit the seller, either because the buyer does not have the need to consume in the period or because the buyer chooses to visit another seller. This assumption keeps the analysis tractable by reducing an individual's history from a potentially infinite sequence to one of two numbers.[3]

2.2. A buyer's decision and payoff

In any arbitrary period, a buyer can be one of two "types", denoted $i \in \{0, 1\}$. If $i = 1$, the buyer has a related seller, and if $i = 0$, the buyer does not have a related seller. Let $V_i^a(\rho)$ denote a type-i active buyer's value function, which is the maximum value that the buyer can obtain in the market. Let $V_i(\rho)$ denote a type-i buyer's value at the *end of the previous period*.[4] The functions V_i and V_i^a are related as follows:

$$V_i(\rho) = \frac{1}{1+r} \left[\lambda V_i^a(\rho) + (1-\lambda) V_0(\rho_{+1})\right], \quad i \in \{0, 1\}. \tag{2.1}$$

The subscripts "+1" indicate next period. This equation is intuitive. When a type-i buyer at the end of the previous period looked forward, he expected to be active in the current period with probability λ, in which case his value function is $V_i^a(\rho)$, and he expected to be inactive in the current period with probability $1 - \lambda$, in which case he becomes unrelated in the market and his value function is equal to $V_0(\rho_{+1})$. Discounting the expected value of these two cases yields the value at the end of the previous period.

Consider the decision of a particular buyer after he becomes active in the period. If the buyer does not have a related seller, he can visit one of two types of sellers, $j \in \{0, 1\}$. If

[3]Because the realization of a buyer's taste shock is likely to be private information, it may be rational for a seller to terminate a relationship when the related buyer does not show up. However, it is complicated to explicitly model this decision under private information. Suppose that a buyer's taste shock is publicly observed. In the equilibrium I will establish later, a buyer is indifferent between visiting the related seller and visiting an unrelated seller. For a seller to keep a relationship with a buyer when the buyer is inactive in the period, the seller must pay a positive amount to the buyer. Such a sidepayment amounts effectively to price discrimination, which is not allowed in this model.

[4]The fraction of sellers who are related to some buyers at the end of the previous period is the same as that at the beginning of the current period, and there is no uncertainty about this aggregate state. Also, since an individual's state variable has only a finite number of values, I simplify the notation by putting it as a subscript instead of an argument of the value function.

$j = 1$, the seller has a related buyer, and if $j = 0$, the seller does not have a related buyer. If the buyer has a related seller, he can visit one of three types of sellers, $j \in \{0, 1, s\}$. If $j = 0$, the seller is not related to any buyer; if $j = 1$, the seller is related to some buyer but not to the specific buyer in the discussion; and if $j = s$, the seller is related to the specific buyer. Let $v_{ij}(\rho, p)$ denote the value to an active type-i buyer from visiting an individual type-j seller who posts price p, where $ij \in \{00, 01, 10, 11, 1s\}$. The buyer's choice is a probability $\theta_{ij}(\rho, p) \in [0, 1]$ with which the buyer visits the seller. Note that $v_{ij}(\rho, p) \leq V_i^a(\rho)$ for all i and j by the definition of V_i^a. Thus, an active type-i buyer's optimal decision satisfies:

$$\theta_{ij}(\rho, p) \begin{cases} = 0, & \text{if } v_{ij}(\rho, p) < V_1^a(\rho) \\ \in [0, 1], & \text{if } v_{ij}(\rho, p) = V_0^a(\rho), \end{cases} \tag{2.2}$$

where $j \in \{0, 1, s\}$ if $i = 1$ and $j \in \{0, 1\}$ if $i = 0$.

Note that θ_{ij} is the probability that a type-i buyer visits an individual seller. When the number of sellers goes to infinity, θ_{ij} is likely to approach zero, except for $j = s$. To characterize a buyer's strategy in the limit, I define the (expected) *queue length* of buyers for a seller. Recall that the number of active buyers who are related to some sellers is $\lambda\rho M$ and the number of active buyers who are unrelated to any seller is $\lambda(N - \rho M)$. For $j \in \{0, 1\}$, define the (expected) queue length of type-1 buyers visiting a type-j seller posting price p as $q_{1j}(\rho, p) \equiv \lambda\rho M \theta_{1j}(\rho, p)$, and the (expected) queue length of type-0 buyers visiting a type-0 seller as $q_{0j}(\rho, p) \equiv \lambda(N - \rho M)\theta_{0j}(\rho, p)$.[5]

Let me calculate the value v_{ij}. Consider first an active type-1 buyer. If the buyer visits the seller he is related to, the buyer will be chosen by the seller with certainty. In this case, the buyer will obtain net utility from consumption in the period, $U - p_1$, and will have an ex ante value $V_1(\rho_{+1})$ at the end of the period. If p is the price posted by the related seller, then the buyer's value of visiting the related seller is

$$v_{1s}(\rho, p) = U - p + V_1(\rho_{+1}). \tag{2.3}$$

If the buyer visits a seller who is not related to the buyer, then the buyer's value is the same as the value of a buyer who visits the same seller but is not related to any seller.

[5] As it is clear from this definition of the queue length, I follow the literature on directed search to focus on the equilibrium where all buyers of the same type respond to a seller's price (including deviations from an equilibrium) in the same way (see Peters, 1991).

That is, the following equalities hold for all p:

$$v_{11}(\rho,p) = v_{01}(\rho,p) \quad \text{and} \quad v_{10}(\rho,p) = v_{00}(\rho,p). \tag{2.4}$$

Next, consider an active type-0 buyer and calculate v_{0j}. I need to compute the probability with which an active type-0 buyer will be chosen by a seller. Consider first a particular type-0 seller who posts price p and label him seller A. In the limit N, $M \to \infty$, the particular buyer visiting seller A will be chosen to trade with by the seller with probability $\frac{1-e^{-(q_{10}+q_{00})}}{q_{10}+q_{00}}$, where $q_{i0} = q_{i0}(\rho,p)$ for $i \in \{1,2\}$ (see Appendix A for the derivation). Let me explain this probability intuitively. The particular buyer is chosen by seller A if and only if seller A receives at least one visitor and if the particular buyer is the chosen one. Since the queue length of type-i buyers visiting seller A is $q_{i0}(\rho,p)$, where $i \in \{0,1\}$, the total queue length of buyers visiting seller A is $(q_{10}+q_{00})$. This queue length is the expected number of buyers visiting seller A. The actual number of visitors is a random variable generated by buyers' visiting probabilities, $\theta_{10}(\rho,p)$ and $\theta_{00}(\rho,p)$. The probability that seller A has at least one visitor is $1-e^{-(q_{10}+q_{00})}$. Conditional on having at least one visitor, seller A randomly chooses one visitor to trade with, and each visitor is chosen with the same probability. The unconditional probability that the particular buyer in the discussion is the chosen one is $\frac{1-e^{-(q_{10}+q_{00})}}{q_{10}+q_{00}}$.

If a buyer is chosen by seller A, the buyer obtains utility $U-p$ in the period. In addition, the buyer becomes related to the seller and, relative to not trading, the relationship changes the value for the buyer by $V_1(\rho_{+1}) - V_0(\rho_{+1})$. Thus, the buyer's surplus from the trade is $[U - p + V_1(\rho_{+1}) - V_0(\rho_{+1})]$ and the buyer's value of visiting seller A is:

$$v_{00}(\rho,p) = \frac{1-e^{-(q_{10}+q_{00})}}{q_{10}+q_{00}}[U - p + V_1(\rho_{+1}) - V_0(\rho_{+1})] + V_0(\rho_{+1}), \tag{2.5}$$

where $q_{i0} = q_{i0}(\rho,p)$ for $i \in \{0,1\}$.

Now consider an active type-0 buyer visiting a particular type-1 seller who posts price p, i.e., a seller who has a related buyer. The type-0 buyer will have a chance to be chosen by the seller only if the seller's related buyer does not visit the seller, which occurs with probability $(1-\lambda\theta_{1s})$, where $\theta_{1s} = \theta_{1s}(\rho,p)$. If all visitors to the seller are unrelated to the seller, the particular type-0 buyer will be chosen to trade with by the seller with a probability that can be calculated similarly to the above, with q_{i0} being replaced with q_{i1}.

In the limit $M, N \to \infty$, this probability becomes:

$$(1 - \lambda\theta_{1s})\frac{1 - e^{-(q_{11}+q_{01})}}{q_{11} + q_{01}},$$

where $q_{i1} = q_{i1}(\rho, p)$ for $i \in \{0, 1\}$, and $\theta_{1s} = \theta_{1s}(\rho, p)$. A trade gives the buyer utility $U - p$ in the period and, relative to not trading, the buyer's value at the end of the period changes by $[V_1(\rho_{+1}) - V_0(\rho_{+1})]$. Thus, the buyer's value of visiting the seller is:

$$v_{01}(\rho, p) = (1 - \lambda\theta_{1s})\frac{1 - e^{-(q_{11}+q_{01})}}{q_{11} + q_{01}}[U - p + V_1(\rho_{+1}) - V_0(\rho_{+1})] + V_0(\rho_{+1}), \tag{2.6}$$

where $q_{i1} = q_{i1}(\rho, p)$ for $i \in \{0, 1\}$, and $\theta_{1s} = \theta_{1s}(\rho, p)$.

Now I can express buyers' optimal strategies in (2.2) explicitly as follows:

$$\theta_{1s}(\rho, p) \begin{cases} = 0 & \text{if } v_{1s}(\rho, p) < V_1^a(\rho) \\ \in [0, 1] & \text{if } v_{1s}(\rho, p) = V_1^a(\rho), \end{cases} \tag{2.7}$$

$$q_{i1}(\rho, p) \geq 0 \text{ and } v_{01}(\rho, p) \leq V_i^a(\rho), \quad i \in \{0, 1\}, \tag{2.8}$$

$$q_{i0}(\rho, p) \geq 0 \text{ and } v_{00}(\rho, p) \leq V_i^a(\rho), \quad i \in \{0, 1\}. \tag{2.9}$$

In (2.8) and (2.9), the two inequalities hold with complementary slackness.

Let P_0 be the set of prices posted by unrelated sellers and P_1 by related sellers. The market value for an active buyer is:

$$\begin{aligned} V_0^a(\rho) &= \max_{j \in \{0,1\}} \max_{p \in P_j} v_{0j}(\rho, p), \\ V_1^a(\rho) &= \max\{v_{1s}(\rho, p_1),\ \max_{j \in \{0,1\}} \max_{p \in P_j} v_{1j}(\rho, p)\}. \end{aligned} \tag{2.10}$$

2.3. A seller's decision and payoff

In any arbitrary period, a seller either has a related buyer (type-1) or has no related buyer (type-0). Let $J_1(\rho)$ be the value function of a seller who has a related buyer and $J_0(\rho)$ that of a seller who does not have a related buyer. Both functions are measured at the *end of the previous period.* To characterize a seller's decision, consider a type-0 seller who posts a price p_0. Recall the assumption that a seller must set a price before knowing which buyer is active. There are ρM type-1 buyers and, counting the probability that a buyer is active, each of these buyers visiting the seller with probability $\lambda\theta_{10}(\rho, p_0)$. There are $(N - \rho M)$ type-0 buyers each visiting the seller with probability $\lambda\theta_{00}(\rho, p_0)$. In the limit $M, N \to \infty$, the probability that the seller gets a buyer is:

$$1 - (1 - \lambda\theta_{10})^{\rho M}(1 - \lambda\theta_{00})^{N-\rho M} \to 1 - e^{-(q_{10}+q_{00})},$$

where $\theta_{i0} = \theta_{i0}(\rho, p_0)$ and $q_{i0} = q_{i0}(\rho, p_0)$ for $i \in \{0, 1\}$.

When the seller has a trade, the current profit is $p_0 - c$ and relative to no trading, the seller's value at the end of the period changes by $[J_1(\rho_{+1}) - J_0(\rho_{+1})]$. Thus, the surplus to the seller from the trade is $[p_0 - c + J_1(\rho_{+1}) - J_0(\rho_{+1})]$. The seller's value function obeys the following Bellman equation:

$$\begin{aligned} J_0(\rho) = \; & \tfrac{1}{1+r} \max_{p_0} \left\{ \left[1 - e^{-(q_{10}+q_{00})}\right] [p_0 - c + J_1(\rho_{+1}) - J_0(\rho_{+1})] + J_0(\rho_{+1}) \right\} \\ & \text{s.t.} \quad (2.9), \qquad \text{where } q_{i0} = q_{i0}(\rho, p_0),\ i \in \{0, 1\}. \end{aligned} \tag{2.11}$$

By incorporating (2.9) as a constraint, the seller explicitly takes into account the effect of the price choice on queue lengths of buyers visiting him. Also, as an implication of dynamic programming, a seller's current choice affects whether the seller's individual state in the future will be 0 or 1 but does not affect the form of future value functions.

Next consider a type-1 seller who posts a price p_1. The related buyer visits the seller with probability $\lambda\theta_{1s}(\rho, p_1)$. A buyer whose related seller is someone else visits the seller with probability $\lambda\theta_{11}(\rho, p_1)$, and a buyer who does not have a related seller visits the seller with probability $\lambda\theta_{01}(\rho, p_1)$. Taking the limit $M, N \to \infty$ and suppressing the arguments (ρ, p_1) of θ_{1s}, q_{11} and q_{01}, I compute the probability that the seller has a trade as

$$1 - (1 - \lambda\theta_{1s})(1 - \lambda\theta_{11})^{\rho M}(1 - \lambda\theta_{01})^{N-\rho M-1} \to 1 - (1 - \lambda\theta_{1s})e^{-(q_{11}+q_{01})}.$$

The seller's value function obeys the following Bellman equation:

$$\begin{aligned} J_1(\rho) = \; & \tfrac{1}{1+r} \max_{p_1} \left\{ \left[1 - (1 - \lambda\theta_{1s})e^{-(q_{11}+q_{01})}\right] [p_1 - c + J_1(\rho_{+1}) - J_0(\rho_{+1})] + J_0(\rho_{+1}) \right\} \\ & \text{s.t.} \quad (2.7) \text{ and } (2.8), \qquad \text{where } (\theta_{1s}, q_{i0}) = (\theta_{1s}, q_{i0})(\rho, p_1),\ i \in \{0, 1\}. \end{aligned} \tag{2.12}$$

With (2.7) and (2.8), the seller explicitly takes into account the effect of the price choice on the related buyer's visiting probability and the queue length of unrelated buyers.

2.4. Equilibrium definition

With many sellers and many buyers in the economy, the analysis is tractable only for equilibria which are symmetric in the sense that all individuals of the same type use the same strategy. In the definition of queue lengths, I have already assumed that all buyers of the same type respond to a seller's price (including a deviation) in the same way. If all sellers of the same type also post the same terms of trade, they attract the same queue length of visitors. In such a symmetric equilibrium, the set of prices posted by type-j sellers is $P_j = \{p_j\}$, where $j \in \{0, 1\}$. For now on, I will suppress the arguments (ρ, p_0) of $(\theta_{i0}, q_{i0}, v_{i0})$ and (ρ, p_1) of $(\theta_{i1}, \theta_{1s}, q_{i1}, v_{i1}, v_{1s})$, where $i \in \{0, 1\}$.

A buyer's visiting probabilities must add up to one across the sellers. Consider first a type-1 buyer. The number of sellers related to the buyer is one, the number of sellers who are related to other buyers but not to the particular buyer is $(\rho M - 1)$, and the number of sellers who do not have related buyers is $(1-\rho)M$. The buyer's visiting probabilities must satisfy: $\theta_{1s} + (\rho M - 1)\theta_{11} + (1-\rho)M\theta_{10} = 1$. Using the definition of the queue length, I can rewrite this adding-up condition in the limit $M, N \to \infty$ as follows:

$$\rho q_{11} + (1-\rho)q_{10} = (1-\theta_{1s})\rho\lambda. \tag{2.13}$$

Similarly, for a type-0 buyer, the visiting probabilities across the sellers must add up to one: $\rho M\theta_{01} + (1-\rho)M\theta_{00} = 1$, and the limit version of this condition is:

$$\rho q_{01} + (1-\rho)q_{00} = (b-\rho)\lambda. \tag{2.14}$$

In the equilibrium, the fraction of related sellers, ρ, is endogenously determined by the flows of sellers between the two types. In a period, if a seller with a related buyer fails to sell, the seller becomes unrelated. This happens with probability $(1-\lambda\theta_{1s})e^{-(q_{11}+q_{01})}$. In the reverse direction, a seller without a related buyer becomes related if the seller succeeds in selling in the period. This occurs with probability $1 - e^{-(q_{10}+q_{00})}$. The fraction of related sellers in the market changes between the current and the next period as

$$\rho_{+1} - \rho = (1-\rho)\left[1 - e^{-(q_{10}+q_{00})}\right] - \rho(1-\lambda\theta_{1s})e^{-(q_{11}+q_{01})}. \tag{2.15}$$

Note that if $\theta_{1s} = 1$, which will be proven to hold in the equilibrium, then $\lambda = 1$ implies $\rho = 1$ in the steady state. This is why the assumption $\lambda < 1$ is needed for $\rho < 1$.

An *equilibrium with priority* consists of buyers' choices θ_{ij}, buyers' value functions (V_i^a, V_i, v_{ij}) (where $j \in \{0, 1, s\}$ for $i = 1$ and $j \in \{0, 1\}$ for $i = 0$), sellers' choices (p_0, p_1), sellers' value functions (J_0, J_1), and the fraction of related sellers ρ that satisfy:

(i) Buyers' value functions (V_i^a, V_i, v_{ij}) satisfy (2.1), (2.3), (2.4), (2.5), (2.6) and (2.10);

(ii) Buyers' choices θ_{ij} and implied queue lengths, $q_{1j} = \lambda \rho M \theta_{1j}$ and $q_{0j} = \lambda(N - \rho M)\theta_{0j}$, satisfy (2.7), (2.8), (2.9), (2.13), and (2.14);

(iii) Sellers' value functions (J_0, J_1) satisfy (2.11), (2.12) and $J_1 \geq J_0$;

(iv) p_0 solves (2.11) and p_1 solves (2.12);

(v) The fraction of related sellers, ρ, satisfies (2.15).

I have explained most of the requirements in the above definition, except the inequality $J_1 \geq J_0$ in (iii). This inequality is necessary and sufficient for a seller to find it optimal to give priority to the related buyer. If $J_1 < J_0$, a seller with a related buyer can increase his value by treating all visitors in the same way.

3. Equilibrium Characterization

Under the maintained assumption $\lambda < 1$, an equilibrium must have the feature that a buyer obtains a strictly positive surplus from a trade. To see this, note that when $\lambda < 1$, there is a positive probability that a seller's related buyer will not show up. In this case, a seller may fail to sell in a period with positive probability because there is lack of coordination among the unrelated buyers. Thus, in any given period, there is a strictly positive fraction of sellers in the market who are not related to any buyer. For such a seller, the tradeoff between the price p_0 and the trading probability $1 - e^{-(q_{10}+q_{00})}$ is smooth. Competition among such unrelated sellers implies that the price p_0 gives a strictly positive surplus to a buyer. This also implies that the price charged by a seller with a related buyer, p_1, must give a strictly positive surplus to a buyer. For if it did not, the buyer related to the seller would choose to visit an unrelated seller instead. I express this result as

$$p_j < U + V_1(\rho_{+1}) - V_0(\rho_{+1}), \quad j \in \{0, 1\}.$$

3.1. Types of equilibria with priority

In this subsection I characterize individuals' optimal choices in more detail and narrow down the set of equilibria. The following lemma is proven in Appendix B:

Lemma 3.1. *An equilibrium with priority satisfies: (i) $\theta_{11} = 0 = \theta_{10}$, and so $\theta_{1s} = 1$; (ii) $\theta_{00} > 0$; (iii) if $\theta_{01} > 0$, then $v_{1s} > v_{00} = v_{10}$; (iv) if $\theta_{01} = 0$, then $v_{1s} = v_{00}$.*

The result $\theta_{11} = 0$ in (i) of Lemma 3.1 says that a buyer with a related seller does not visit a seller whose related buyer is someone else. Relative to visiting such a seller, the buyer can get a strictly higher payoff from visiting his related seller, since the two sellers post the same price and the buyer gets a good with probability one from the related seller. The result $\theta_{10} = 0$ says that it is a dominant strategy for a buyer with a related seller not to visit a seller who does not have a related buyer; that is, the buyer only visits his related seller. To explain this result, let me refer to the particular buyer as buyer B and the seller related to him as seller A. Consider first the case where other buyers do not visit seller A. For seller A to get a positive (expected) payoff in this case, the seller must attract buyer B with positive probability. If this probability were strictly less than one, seller A could increase the visiting probability to one by cutting the price only slightly, which would increase the seller's expected payoff by a discrete amount. Next, consider the case where other buyers visit seller A with positive probability. In this case, if seller A cuts the price slightly to induce buyer B to visit with probability one, it will generate a discrete reduction in the probability that an unrelated visitor is chosen, which may drive unrelated visitors away. To explain the result $\theta_{1s} = 1$ in this case, suppose that buyer B visits another seller, say, seller C. Because buyer B is unrelated to seller C, his payoff from the visit will be the same as the payoff to a buyer without a related seller who visits seller C. Because such unrelated buyers visit seller A with positive probability in the case in discussion, their payoff from visiting seller A must be at least as large as the payoff from visiting seller C. This implies that buyer B's payoff from visiting seller C cannot be higher than the payoff from visiting the related seller A without being given priority. Because buyer B does have priority from seller A, he must be strictly better off visiting seller A than visiting any other seller.

The result (ii) says that a buyer without a related seller visits a seller without a related buyer with positive probability. This result is easy to understand: Since a seller without a related buyer does not discriminate the visitors, he can always attract the buyers without related sellers. The result (iii) says that if unrelated buyers visit the sellers with related buyers with positive probability, that the equilibrium payoff to a related buyer must be strictly higher than the payoff to an unrelated buyer. This result is explained as part of the explanation for (i) above. Finally, the result (iv) says that if unrelated buyers only visit unrelated sellers, then the two types of buyers must have the same payoff in the equilibrium. This is because in the case where a related seller's only potential buyer is the related buyer, a price increase does not generate crowding-out among the seller's potential visitors. In this case, the related seller can increase the payoff by raising price until the related buyer just slightly prefers visiting the seller.

With Lemma 3.1, an equilibrium must be one of the two types:
(i) Partial mixing: An active buyer with a related seller only visits that seller, and an active buyer without a related seller mixes between the two types of sellers. In this case, $q_{01} > 0$, $q_{00} > 0$, $\theta_{1s} = 1$, and $q_{10} = q_{11} = 0$. This case has $v_{1s} > v_{01} = v_{00} = v_{10}$ and, hence, $V_1(\rho) > V_0(\rho)$. For partial mixing to be an equilibrium, it should not be profitable for a type-1 seller to raise p_1 to drive away type-0 buyers.
(ii) Complete separation: An active buyer with a related seller only visits his related seller, and an active buyer without a related seller only visits the sellers without related buyers. In this case, $q_{01} = 0$, $q_{00} > 0$, $\theta_{1s} = 1$, and $q_{10} = q_{11} = 0$. This case has $v_{1s} = v_{00} = v_{10} \geq v_{01}$ and, hence, $V_1(\rho) = V_0(\rho)$. For complete separation to be an equilibrium, it should not be profitable for a type-1 seller to reduce p_1 to attract type-0 buyers.

In both cases, the equilibrium satisfies $V_1^a(\rho) = \max_p v_{1s}(\rho, p)$ and $V_0^a(\rho) = \max_p v_{00}(\rho, p)$. Also, because $q_{10} = q_{11} = 0$, I shorten the notation q_{00} as q_0, and q_{01} as q_1.

I examine first the possibility of an equilibrium with partial mixing. For a seller without a related buyer, the problem in (2.11) with partial mixing can be rewritten as

$$\begin{aligned} (1+r)J_0(\rho) &= J_0(\rho_{+1}) + \max_{(p_0,q_0)} (1-e^{-q_0})\left[p_0 - c + J_1(\rho_{+1}) - J_0(\rho_{+1})\right] \\ \text{s.t.} \quad & \tfrac{1-e^{-q_0}}{q_0}\left[U - p_0 + V_1(\rho_{+1}) - V_0(\rho_{+1})\right] = V_0^a(\rho) - V_0(\rho_{+1}). \end{aligned} \tag{3.1}$$

Here I have expressed the dependence of the queue length q_0 on the price p_0 explicitly as a constraint and added q_0 to the list of choices accordingly. The constraint is a form of (2.9) with $q_0 > 0$, which requires that the expected surplus that a type-0 buyer gets from the seller should be equal to the buyer's expected surplus in the market.[6] The seller

[6]The actual form of the constraint is that the expected surplus to the buyer should be at least as large as the buyer's surplus in the market. But it is never possible in an equilibrium for a seller to give a buyer more than the market surplus.

takes future value functions, $(V_0, V_1, V_0^a, J_0, J_1)(\rho_{+1})$, as given. The first-order condition, the constraint and the Bellman equation in (3.1) imply:

$$\begin{aligned} &V_0^a(\rho) - V_0(\rho_{+1}) = e^{-q_0}\Delta(\rho_{+1}), \\ &p_0 = U + V_1(\rho_{+1}) - V_0(\rho_{+1}) - \tfrac{q_0}{1-e^{-q_0}}[V_0^a(\rho) - V_0(\rho_{+1})], \end{aligned} \tag{3.2}$$

$$(1+r)J_0(\rho) - J_0(\rho_{+1}) = \left[1 - (1+q_0)e^{-q_0}\right]\Delta(\rho_{+1}), \tag{3.3}$$

where Δ is the total surplus of the match defined as

$$\Delta(\rho_{+1}) \equiv U - c + V_1(\rho_{+1}) - V_0(\rho_{+1}) + J_1(\rho_{+1}) - J_0(\rho_{+1}). \tag{3.4}$$

These results have two noteworthy features, although they are standard in models of directed search. First, an individual's trading probability (i.e., matching rate) and the share of the match surplus are endogenous and are functions of only the queue length q_0. A buyer's probability of trade when visiting an unrelated seller is $\frac{1-e^{-q_0}}{q_0}$, which is a decreasing function of q_0. When the buyer gets a trade, the buyer's surplus is $U+V_1(\rho_{+1})-V_0(\rho_{+1})-p_0$. As implied by the two equations in (3.2), the buyer's surplus is a share $\frac{q_0}{e^{q_0}-1}$ of the match surplus Δ. This share depends only on q_0 and is a decreasing function of q_0. Thus, a buyer's share of the match surplus has a one-to-one positive relation to the buyer's trading probability. Second, an individual is compensated with an ex ante surplus according to the individual's social marginal contribution to the match. For example, a buyer's ex ante surplus from visiting an unrelated seller, $V_0^a(\rho) - V_0(\rho_{+1})$, is equal to $e^{-q_0}\Delta$. To see why this is equal to the buyer's social marginal contribution, suppose that an additional buyer joins the matching process with unrelated sellers. Because this buyer contributes to the match only if an unrelated seller fails to get a match, which occurs with probability e^{-q_0}, the buyer's social marginal contribution is precisely $e^{-q_0}\Delta$. Similarly, an unrelated seller's social marginal contribution is equal to the expected surplus the seller creates, $(1-e^{-q_0})\Delta$, minus the expected surplus the seller crowds out, $q_0e^{-q_0}\Delta$, as shown by (3.3).

With partial mixing, a type-1 seller's decision problem in (2.12) can be written as

$$(1+r)J_1(\rho) = J_0(\rho_{+1}) + \max_{(p_1,q_1)} \left[1 - (1-\lambda)e^{-q_1}\right][p_0 - c + J_1(\rho_{+1}) - J_0(\rho_{+1})] \tag{3.5}$$

subject to

$$(1-\lambda)\frac{1-e^{-q_1}}{q_1}[U - p_1 + V_1(\rho_{+1}) - V_0(\rho_{+1})] = V_0^a(\rho) - V_0(\rho_{+1}), \tag{3.6}$$

$$\left[1-(1-\lambda)e^{-q_1}\right]\left[p_1-c+J_1(\rho_{+1})-J_0(\rho_{+1})\right] \geq \lambda\left[\bar{p}-c+J_1(\rho_{+1})-J_0(\rho_{+1})\right], \tag{3.7}$$

where

$$\bar{p}=U+V_1(\rho_{+1})-V_0^a(\rho). \tag{3.8}$$

As a form of (2.8) with $q_1>0$, (3.6) requires that the expected surplus to a buyer who does not have a related seller from visiting the type-1 seller should be equal to the buyer's surplus in the market. Note that this constraint implies $U-p_1+V_1(\rho_{+1})>V_0^a(\rho)$. That is, if p_1 satisfies (3.6), then a buyer gets a strictly higher payoff from visiting the related seller than from visiting any other seller. The second constraint, (3.7), requires that the seller should not gain from deviating to a price that attracts only the related buyer. To see this, note that $\bar{p}$ satisfies $v_{1s}(\rho,\bar{p})=V_0^a(\rho)$. That is, $\bar{p}$ is the highest price that the type-1 seller can charge and still attract the related buyer to visit him. By posting $\bar{p}-\varepsilon$, where $\varepsilon>0$ is sufficiently small, the seller can attract the related buyer with certainty, provided that the related buyer is active in the period. Also, the price $\bar{p}-\varepsilon$ does not attract any type-0 buyer, because the expected value to an unrelated buyer from visiting the seller is strictly less than the value $V_0^a(\rho)$ that the buyer can get elsewhere. Thus, to the type-1 seller, posting $\bar{p}$ yields the expected surplus $\lambda\left[\bar{p}-c+J_1(\rho_{+1})-J_0(\rho_{+1})\right]$. Constraint (3.7) requires this expected surplus from the deviation not to exceed that from posting p_1 to attract both the related buyer and the buyers who have no related sellers.[7]

If (3.7) is binding, then the type-1 seller can gain by deviating to $\bar{p}$ that attracts only the related buyer. In this case, an equilibrium with partial mixing does not exist. Thus, to characterize an equilibrium with partial mixing, I omit (3.7) for the moment and check later whether it is satisfied. The first-order condition of the problem in (3.5) yields:

$$V_0^a(\rho)-V_0(\rho_{+1})=\frac{(1-\lambda)^2(1-e^{-q_1})^2 e^{-q_1}\Delta(\rho_{+1})}{(1-\lambda)\left[1-e^{-q_1}\right]^2+\lambda\left[1-(1+q_1)e^{-q_1}\right]}, \tag{3.9}$$

where Δ is defined in (3.4), and (3.5) yields:

$$(1+r)J_1(\rho)-J_0(\rho_{+1})=\frac{\left[1-(1+q_1)e^{-q_1}\right]\left[1-(1-\lambda)e^{-q_1}\right]^2\Delta(\rho_{+1})}{(1-\lambda)\left[1-e^{-q_1}\right]^2+\lambda\left[1-(1+q_1)e^{-q_1}\right]}. \tag{3.10}$$

[7] In technical terms, a related seller's payoff function has a discrete jump at $q_1=0$. Constraint (3.7) ensures that the maximum payoff that a related seller obtains with $q_1>0$ is greater than or equal to the payoff at this point of jump.

Similar to the case with a type-0 seller, the expressions in (3.9) and (3.10) show that an individual's share of the surplus of a match with a type-1 seller is a function of the queue length of buyers for such a seller. However, because a type-1 seller gives priority to the related buyer, this function is more complicated than the ones in (3.2) and (3.3).

The two expressions for $[V_0^a(\rho) - V_0(\rho_{+1})]$ in (3.2) and (3.9) must be consistent with each other, and so the following relation holds between q_0 and q_1:

$$q_0 = h(q_1) \equiv q_1 + \ln\left\{\frac{1}{1-\lambda}\left[1 + \frac{\lambda}{1-\lambda}\frac{[1-(1+q_1)e^{-q_1}]}{(1-e^{-q_1})^2}\right]\right\}. \tag{3.11}$$

It can be verified that $h(q)$ is an increasing function for all $q > 0$.

Now I can establish the following lemma (see Appendix B for a proof):

Lemma 3.2. *An equilibrium with partial mixing does not exist.*

The reason why an equilibrium with partial mixing does not exist lies in its requirement that a type-0 buyer should be indifferent between visiting a type-0 seller and a type-1 seller. When visiting a type-0 seller, a type-0 buyer has the same chance of being chosen to trade with as any other visitor to the seller. In contrast, when visiting a type-1 seller, a type-0 buyer has a chance of being chosen to trade with only when the seller's related buyer does not show up. In the presence of this low priority, a type-0 buyer is willing to visit a type-1 seller only when either the queue length of buyers for a type-1 seller or the price posted by a type-1 seller is sufficiently lower than that at a type-0 seller. In either case, a type-1 seller's expected surplus is lower than a type-0 seller's, and so $J_1 < J_0$. That is, a type-1 seller would rather increase the price to $\bar{p}$ to attract only the related buyer. This destroys the equilibrium with partial mixing.

3.2. The equilibrium with complete separation

The analysis so far has established that the only possible equilibrium with priority is the equilibrium with complete separation. With complete separation, a type-0 seller's maximization problem is still (3.1) and the optimality conditions are still given by (3.2) and (3.3). In contrast, a type-1 seller now attracts only the related buyer (since $q_1 = 0$) and the price $\bar{p}$ given by (3.8) is the best for doing so. Because $\bar{p}$ does not attract any type-0 buyer to visit the seller, a type-1 seller's value function is now given by

$$(1+r)J_1(\rho) = J_0(\rho_{+1}) + \lambda\left[\bar{p} - c + J_1(\rho_{+1}) - J_0(\rho_{+1})\right]. \tag{3.12}$$

For complete separation to be an equilibrium, a type-1 seller should not gain from deviating to a price that attracts the buyers who do not have related sellers. Among all prices that attract both types of buyers, the best price p_1 and the implied queue length q_1 solve the maximization problem in (3.5) subject to the constraint (3.6). This pair (p_1, q_1) satisfies (3.6) and (3.9). Let me denote the values of such (p_1, q_1) as $(\tilde{p}_1, \tilde{q}_1)$. Because $\tilde{q}_1$ satisfies (3.9), the consistency between (3.9) and (3.2) implies $q_0 = h(\tilde{q}_1)$, where h is defined in (3.11). The price $\tilde{p}_1$ can be retrieved from (3.6). Let me express this deviation as

$$\tilde{q}_1 = h^{-1}(q_0), \quad \tilde{p}_1 = U + V_1(\rho_{+1}) - V_0(\rho_{+1}) - \frac{\tilde{q}_1[V_0^a - V_0]}{(1-e^{-\tilde{q}_1})(1-\lambda)}. \tag{3.13}$$

With this deviation, the type-1 seller sells the good with probability $[1-(1-\lambda)e^{-\tilde{q}_1}]$, and the resulted surplus of trade to the seller is $[\tilde{p}_1 - c + J_1(\rho_{+1}) - J_0(\rho_{+1})]$. Thus, the deviation is not profitable for the type-1 seller if and only if

$$\left[1-(1-\lambda)e^{-\tilde{q}_1}\right]\left[\tilde{p}_1 - c + J_1(\rho_{+1}) - J_0(\rho_{+1})\right] \leq \lambda\left[\bar{p} - c + J_1(\rho_{+1}) - J_0(\rho_{+1})\right]. \tag{3.14}$$

An equilibrium with complete separation must also satisfy $J_1(\rho) \geq J_0(\rho)$ in order to ensure that it is optimal for a type-1 seller to give priority to the related buyer. Moreover, q_0 and ρ need to be determined and verified to satisfy $q_0 > 0$ and $\rho \in (0, \min\{1, b\})$. With complete separation, the adding-up constraint (2.14) yields:

$$q_0 = H(\rho) \equiv \lambda(b-\rho)/(1-\rho). \tag{3.15}$$

Then, the law of motion of ρ given in (2.15) becomes:

$$\rho_{+1} = G(\rho) \equiv \lambda\rho + (1-\rho)\left[1 - e^{-H(\rho)}\right]. \tag{3.16}$$

Note that (3.15) and (3.16) are independent of other variables than (ρ, q_0). Thus, the steady state and the dynamics of (ρ, q_0) can be solved from these two equations without the help of any other equilibrium relations.

Denote the steady state of the equilibrium by adding the superscript $*$ to the variables. I prove the following proposition in Appendix C:

Proposition 3.3. *(i) The fraction of related sellers in the market has a unique steady state ρ^*, which is locally stable. The steady state and the dynamics of (ρ, q_0) depend only on (λ, b) and not on other parameters such as (c, U). Moreover, $0 < \rho^* < \min\{1, b\lambda\}$, $q_0^* > 0$, $\frac{d\rho^*}{db} > 0$, $\frac{dq_0^*}{db} > 0$, and $\frac{d\rho^*}{d\lambda} > 0$.*
(ii) There exists $B(\lambda) > 0$, defined in Appendix C, such that the steady state with complete separation exists if and only if $b \leq B(\lambda)$. Under this condition, the steady state is unique. $B(\lambda)$ is independent of other parameters such as (c, U).
(iii) For ρ close to ρ^*, $\Delta(\rho) > 0$, and the dynamic equilibrium satisfies:

$$p_0 = U - \frac{q_0 e^{-q_0}}{1 - e^{-q_0}} \Delta(\rho_{+1}), \tag{3.17}$$

$$p_1 = \bar{p} = U - e^{-q_0} \Delta(\rho_{+1}) > \max\{p_0, c\}, \tag{3.18}$$

$$V_1(\rho) = V_0(\rho) = \frac{1}{1+r} \left[V_0(\rho_{+1}) + \lambda e^{-q_0} \Delta(\rho_{+1}) \right], \tag{3.19}$$

$$\Delta(\rho) = U - c + \frac{1}{1+r} \left[\lambda(1 - e^{-q_0}) - 1 + (1 + q_0) e^{-q_0} \right] \Delta(\rho_{+1}), \tag{3.20}$$

where $q_0 = H(\rho)$ and $\rho_{+1} = G(\rho)$ are given by (3.15) and (3.16).

The fraction of related sellers in the market has a unique steady state which is locally stable. To understand this result, it is useful to consider the special case $b = 1$, i.e., the case where the number of buyers is equal to the number of sellers. In this case, $q_0 = \lambda$ (see (3.15)), and $G(\rho)$ is linear in ρ with a slope $(\lambda - 1 + e^{-\lambda}) \in (0, 1)$. In this case, the flow of related sellers who lose relationships due to the absence of related buyers from the market is $(1 - \lambda)\rho$, and the flow of unrelated sellers who gain relationships through trade is $(1 - \rho)(1 - e^{-\lambda})$. If $\rho > \rho^*$, the flow of related sellers who lose relationships is greater, and the flow of unrelated sellers who acquire relationships is smaller, than the flow in the steady state. Because there is a positive net flow of sellers out of relationships, the fraction of related sellers falls toward the steady state. On the other hand, if $\rho < \rho^*$, the flow of related sellers who lose relationships is smaller, and the flow of unrelated sellers who acquire relationships is greater, than the flow in the steady state. The fraction of related sellers increases toward the steady state. Similar stabilizing forces are at play when $b \neq 1$, although the dependence of q_0 on ρ adds to the dynamics.

It is notable that ρ and q_0 depend only on (λ, b) and not on other parameters such as (c, U). Recall that the extensive margin of the demand is $b\lambda$. To explain why ρ and q_0 depend on the extensive margin of the demand but not on the intensive margin of the market, note that ρ and q_0 in the equilibrium are determined by the requirements on the matching rates and not by the size of the match surplus. Specifically, given the parameters (λ, b) and the composition of sellers, ρ, the queue length q_0 is uniquely pinned down by the requirement that a buyer's visiting probabilities across the sellers should add up to one. Conversely, given (λ, b) and q_0, each seller's matching rate is uniquely determined, as explained above. These matching rates in turn determine the transition of sellers between the two types and, hence, the dynamics of the composition of sellers, ρ. These two relations are given above as $q_0 = H(\rho)$ and $\rho_{+1} = G(\rho)$, respectively. Because the only parameters in these relations are (λ, b), the solutions for (ρ, q_0) to these relations depend only on (λ, b) and not on other parameters such as the cost c and the utility level U.[8]

It is intuitive that the fraction of related sellers in the steady state increases in the two dimensions of the extensive margin of the demand, b and λ. A higher b or λ leads to a larger number of trades in each period. Because a trade keeps a related seller related to a buyer and turns an unrelated seller into a related one, the increase in the number of trades increases the fraction of sellers who are related to some buyers in the equilibrium. It is also intuitive that a higher buyer/seller ratio increases the queue length of buyers for an unrelated seller in the steady state, q_0^*, because the larger number of buyers must be eventually allocated to the sellers. However, it is ambiguous whether a higher λ increases q_0^*. On the one hand, a higher λ increases the demand per seller, which has a positive effect on q_0^*. On the other hand, since sellers give priority to related buyers, a higher λ can lead to proportionally more buyers visiting related sellers, which reduces the queue length of buyers for each unrelated seller. The overall effect of λ on q_0^* depends on the values of λ and b (see subsection 4.2.3 for an example).

The steady state with complete separation exists if and only if the buyer/seller ratio is not too high, i.e., if and only if $b \leq B(\lambda)$. When the number of buyers for each seller is small, it is difficult for a seller to obtain a trade. Because increasing the probability of trade is relatively important for a seller in this case, it is optimal for a seller to give

[8]If there is free entry of sellers, then b is determined endogenously. In this case, ρ and q_0 depend on (c, U) through b. See section 5.

priority to his related buyer so as to guarantee a trade when the buyer is active. Moreover, when the buyer/seller ratio is low, the price charged by a seller without a related buyer is likely to be low. Competing against such sellers for unrelated buyers is not optimal for a seller with a related buyer, because it requires the seller to cut price sufficiently to compensate unrelated buyers for their low priority. Thus, the equilibrium in this case has complete separation between related and unrelated individuals. On the other hand, if the buyer/seller ratio is high, getting a buyer is relatively easy for a seller, in which case it is optimal for a seller to treat all buyers equally.

The critical level B depends only on λ and not on other parameters such as the cost c and the utility U. To see why, note that an *individual* seller's decisions on whether or not to give priority to the related buyer and whether to attract one type or two types of buyers change the seller's expected share of the match surplus but not the size of the surplus (which is given as Δ). As explained above, a seller's expected share of the match surplus is determined by the endogenous matching rate which, in turn, is only a function of the queue length q_0 and the parameters (λ, b). Because the queue length q_0 depends only on (λ, b), so does a seller's expected surplus share. Other parameters, such as (c, U), affect the size of the match surplus but do not affect how the surplus is shared between the two sides. Different strategies of a seller lead to different expected shares of the match surplus to the seller, all of which are only functions of (λ, b). The seller chooses the strategy that yields the highest share. Thus, the condition needed for the strategy of complete separation to be optimal is a restriction only on (λ, b), which is expressed as $b \leq B(\lambda)$.

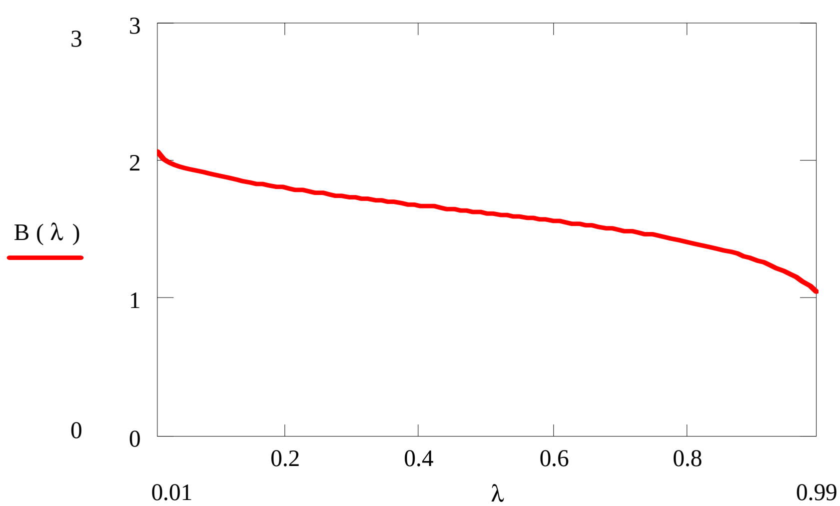

Figure 1. The critical level $B(\lambda)$ of the buyer/seller ratio

An increase in λ has two effects on the critical level B, similar to the two effects of λ on q_0. One effect is the general effect of a higher λ on the demand. This effect makes a relationship less valuable, reduces B and makes the steady state with complete separation less likely to exist. The other effect is that a higher λ tilts the demand toward related sellers because of the priority. This effect makes a relationship more valuable, increases B and makes the steady state with complete separation more likely to exist. The overall effect of λ on B is difficult to determine analytically, despite that the $B(\lambda)$ involves no other parameter. It is straightforward to compute $B(\lambda)$ numerically, which is depicted in Figure 1. It is clear that $B(\lambda) > 1$ and $B'(\lambda) < 0$ for all $\lambda \in (0, 1)$. Thus, a higher λ makes the equilibrium with complete separation less likely to exist for any given b.

3.3. Relationship, the regular price and sales

In any period, related sellers post price $p_1 = \bar{p}$ and unrelated sellers post $p_0 < \bar{p}$ (see (3.18)). Each price follows dynamics toward its own steady-state level as the fraction of related sellers adjusts to the steady state. For reasons that will become clear below, let me interpret p_1 as the *regular price* and p_0 as the *sale price*. A seller posts the regular price as long as he has a related buyer. Once a seller loses the relationship, he holds a sale at price p_0 until he has a trade after which he switches back to the regular price.

There are two noteworthy features of such sales. First, the primary consideration for holding a sale is intertemporal. By holding a sale, a seller intends to attract buyers and build a relationship, but the relationship will be paid off only in the future through a higher (regular) price. Second, the sale price can be even below the marginal cost of the good. That is, it is possible that $p_0 < c$, which can be verified with the formula of p_0 in (3.17). To explain why this possibility exists, note that the surplus of a trade to a seller without a related buyer is $[p_0 - c + J_1 - J_0]$. Even if $p_0 < c$, the surplus can still be positive if the gain in the future through the relationship, as measured by $(J_1 - J_0)$, is large enough to outweigh the temporary loss $(p_0 - c)$. In contrast, the regular price p_1 is always strictly higher than the marginal cost (see (3.18)). This contrast between the two prices and the intertemporal motivation for posting the lower price p_0 justify the interpretation of p_0 as the sale price and p_1 as the regular price.

The equilibrium has precise predictions on the frequency, the duration and the price discount of sales. To calculate the frequency of trades at the sale price, note that the number of trades at p_1 is $M\rho\lambda$ and the number of trades at p_0 is $M(1-\rho)(1-e^{-q_0})$. Denote the frequency of trades at the sale price as $dprob$. Then,

$$dprob = \frac{(1-\rho)(1-e^{-q_0})}{\rho\lambda + (1-\rho)(1-e^{-q_0})}. \tag{3.21}$$

When $\rho = \rho^*$, where ρ^* is the steady state of (3.16), it is easy to verify that $dprob^* = 1-\lambda$. The frequency of trades at the sale price in the steady state depends on λ but not on (b, c, U). This result is intuitive: because the motive for holding a sale is to gain a relationship with a buyer, how often a seller holds sales should depend only on how frequently the seller loses a relationship, which is $1-\lambda$.

Let $dlength$ denote the expected duration of a sale. If a seller holds a sale in the current period, the sale will continue next period if and only if the seller fails to trade in the current period, which occurs with probability e^{-q_0}. Thus, $dlength = 1 + e^{-q_0} \times dlength$, and so

$$dlength = \frac{1}{1-e^{-q_0}}. \tag{3.22}$$

As explained for Proposition 3.3, q_0 depends only on the extensive margin of the demand, (λ, b). Thus, the duration of a sale depends only on (λ, b) and not on the intensive margin of the market. Moreover, the two dimensions of the extensive margin can affect the duration of a sale differently. Because a higher b increases q_0, the duration of a sale always falls when the buyer/seller ratio increases. In contrast, λ affects the duration of a sale ambiguously because it affects q_0 ambiguously. The duration decreases in λ if q_0 increases in λ, and the duration increases in λ if q_0 decreases in λ.

Let $dsize$ denote the percentage price discount of a sale. Then,

$$dsize = 1 - \frac{p_0}{p_1} = \left(\frac{q_0}{1-e^{-q_0}} - 1\right)\left[\frac{Ue^{q_0}}{\Delta} - 1\right]^{-1}. \tag{3.23}$$

In contrast to the frequency and duration of a sale, the price discount depends on the factors on the intensive margin of the market, such as (c, U), as well as the factors on the extensive margin, (λ, b). It depends (λ, b) through (q_0, Δ) and on (c, U) through U/Δ. For future use, I define the markup implied by each price as

$$markup0 = \frac{p_0}{c} - 1, \quad markup1 = \frac{p_1}{c} - 1. \tag{3.24}$$

Because the regular price is strictly higher than the cost (see (3.18)), the regular price always implies a positive markup. In contrast, since it is possible to have $p_0 < c$, the sale price can imply a *markdown* instead of a markup. Weighting each markup by the frequency of trades occurring at the associated price, I obtain the average markup as

$$markavg = dprob \times markup0 + (1 - dprob) \times markup1. \tag{3.25}$$

Similarly, I can calculate the average price in the equilibrium.

4. Calibration, Comparative Statics and Dynamics

In this section I calibrate the steady state of the model to the data, compute the responses of the steady state to changes in (c, b, λ), and then examine the dynamics.

4.1. Calibration

The model has five parameters, (r, U, c, b, λ). Table 1 lists the targets used to identify these parameters and the identified values. The length of a period is chosen to be one week. By setting the annual discount rate to 4%, I determine the value of the weekly discount rate through $(1 + r)^{52} = 1.04$. The utility of consuming a good is normalized to one. I set the fraction of price quotes (not trades) at p_0^* in the steady state to $1 - \rho^* = 0.15$, which solves ρ^*. Klenow and Kryvtsov (2008) report that roughly 11% of all price quotes are sale prices in the microdata on the U.S. monthly CPI in 1988-2004 collected by the Bureau of Labor Statistics. Because sales tend to be short-lived, the monthly data are likely to under-report the frequency of sales, and so I use a larger number for the frequency of sales.[9]

Table 1. Parameters and calibration targets

parameter	value	target
r: discount rate	7.545×10^{-4}	annual discount rate = 0.04
U: utility level	1	normalization
λ: active prob.	0.956	frequency of sale price quotes = 0.15
b: buyer/seller ratio	0.895	average duration of a sale = 4 weeks
c: marginal cost	0.311	percentage price discount of a sale = 0.28

Next, I set the duration of a sale in the steady state to 4 weeks. Using the microdata on U.S. CPI in 1988-2004, Nakamura and Steinsson (2008) report that the fraction of sales

[9] To get a sense of the effect of time aggregation, note that Klenow and Kryvtsov (2008) report that 15% of all (monthly) price quotes for food items are sales. With daily data collected on a specific food item (bottled ketchup), Pesendorfer (2002) reports that 23% of prices fall in the range \$0.99 - \$1.19, while 60% of all daily prices fall in the range \$1.39 - \$1.49. If the first range is regarded as the range of sale prices, then sale prices are at least 23% of all price quotes.

that last just one month ranges between 35% and 60% in the four major groups (processed food, unprocessed food, household furnishings, and apparel), and the average length of sales is 1.8-2.3 months. Since these major groups tend to be the ones with more frequent and longer sales than other goods, and since sales shorter than one month are likely to be under-sampled in the monthly data, the length of 4 weeks is a reasonable target for the duration of a sale.[10] Because the length of a sale in the model is given by (3.22), this target determines q_0^*. Now that ρ^* and q_0^* are solved, I can retrieve λ and b from the steady-state version of (3.15) and (3.16), which are (C.1) and (C.2) in Appendix C. Finally, I match the price discount of a sale in the steady state to the target 28%. The sale price change is 25.1% on average in the U.S. CPI microdata (see Klenow and Kryvtsov, 2008) and, in eleven major groups of goods, it is 29.5% (see Nakamura and Steinsson, 2008). I take a number between these two as the target. Since the price discount of a sale in the model is given by (3.23), this target determines c.

Note that b is noticeably less than one, and so the condition $b \leq B(\lambda)$ under which the steady state with complete separation exists is satisfied. Also, since λ is the fraction of trades (rather than price quotes) that occur at the price p_1^*, an overwhelming majority of trades (about 95.6%) occur at p_1^* in steady state, which lends further support for the interpretation of p_1 as the regular price. Moreover, the average price in the steady state is 0.346 and the average markup in the steady state is $markavg^* = 0.113$.

4.2. Comparative statics

Using the calibrated model, I conduct comparative statics with respect to changes in (c, b, λ). The superscript $*$ on steady-state variables is suppressed in this subsection.

4.2.1. Steady-state responses to changes in the cost

The first experiment is to compute the responses of the steady state to changes in the marginal cost of a good. The value of c identified above is 0.311. I vary c from 0.1 to 0.9.

[10] Kehoe and Midrigan (2010) report that the duration of a temporary price (such as a sale price) is 1/0.53 *months* in the BLS data and 1/0.46 *weeks* in scanner data collected from a large supermarket chain. The difference between these two numbers reflects not only the fact that the BLS data covers more goods and services than scanner data does, but also the fact that the BLS data samples prices monthly while scanner data samples weekly. The duration of 4 weeks that I choose is close to the average of the two numbers in the two datasets.

Most variables respond to the cost in a predictable way. For example, as the cost increases, both the regular price p_1 and the sale price p_0 increase. The values to a buyer and a seller fall, regardless of whether the individual is related in the market. Because q_0 and ρ are independent of c, as established in Proposition 3.3, the matching rates, the frequency of trades at the sale price and the duration of a sale are all independent of c.

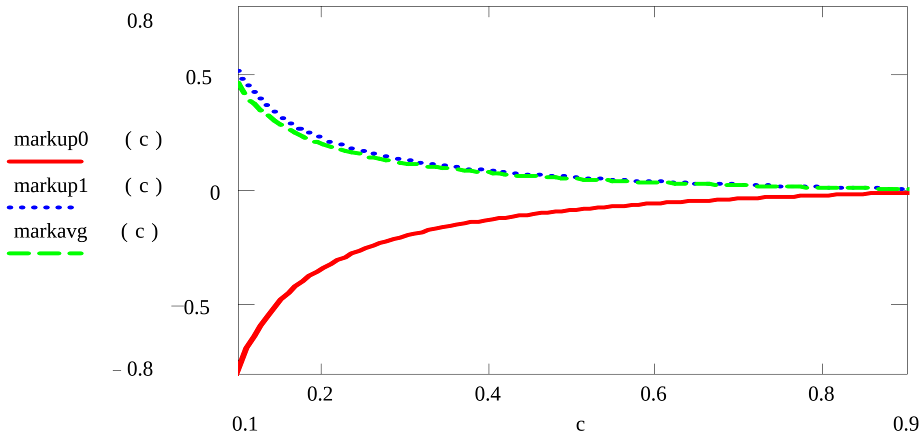

Figure 2. Responses of markups to the marginal cost c

The responses of the markups to the marginal cost are noteworthy. In Figure 2, I depict steady-state markups as functions of c. The markup implied by the sale price, $markup0$, is negative for all values of c, and so the sale price is a markdown on the marginal cost. The occurrence of this markdown is intuitive. Because the buyer/seller ratio is noticeably less than one ($b = 0.895$), a relationship is important for a seller. By offering a markdown, a seller can attract a buyer to visit and the resulting relationship enables the seller to sell a good with a higher probability in the future. The current lose from the the markdown is compensated with a markup at the regular price in the future. At the baseline value of the cost, the markdown implied by the sale price is $markup0 = -0.189$ and the markup implied by the regular price is $markup1 = 0.127$. The average markup is positive and close to the markup implied by the regular price, because a majority (95.6%) of trades take place at the regular price. The occurrence of markdowns is consistent with the finding by Dutta et al. (2002) using scanner data from a supermarket chain. Note that most of the IO models of sales cited in the introduction do not generate markdowns.

The markdown and the markup are large when the cost is very low and their absolute values decrease precipitously when the cost increases. For example, when the cost is

$c = 0.1$, the markdown is -0.765, the markup is 0.513 and the average markup is 0.457, but when the cost increases to $c = 0.5$, the markdown is -0.085, the markup is 0.057 and the average markup is 0.051. When the cost is very low, the markdown and the markup are large because the gain to a seller from making a buyer related to him is large. As this gain decreases in the cost, the room for markups is reduced, and so is the need to use large markdowns to attract a buyer to build a relationship.

These responses of the markups to the cost have useful implications for constructing macro models. First, if customer relationships are important in the retail market, then the models with monopolistic competition popularly used in macro are not suitable for capturing the responses of prices to cost shocks. Those models typically have constant markups. Second, the percentage change in the regular price in response to the cost change is less than that in the sale price. In this sense, the regular price is relatively more stable than the sale price. The source of this difference is that a related seller's trading probability is higher than an unrelated seller's, and so the same change in the price amounts to a larger change in expected profit for a related seller than for an unrelated seller. To cover the increase in the cost, a smaller change in the regular price is needed than in the sale price. Third, since the sale price responds to the cost differently from the response of the regular price, it has a life of its own. Specifically, after a cost shock, the new sale price differs from the previous sale price and it is not a fixed fraction of the new regular price. A model that ignores or simply filters out sales prices distorts the responses of prices to cost shocks. Finally, across goods that differ in the cost and utility, the higher the difference $(U - c)$, the larger the difference between the regular price and the sale price. That is, goods with a higher profit margin have a higher variability in prices.

4.2.2. Steady-state responses to changes in the buyer/seller ratio

The second experiment is to change the extensive margin of the demand by changing the buyer/seller ratio b. As demonstrated analytically in Proposition 3.3, an increase in b increases the queue length of buyers for each unrelated seller, q_0, and increases the fraction of sellers who have related buyers, ρ. As a result, the duration of a sale decreases. For b ranging from 0.45 to 1.1, Figure 3.1 depicts steady-state markups and Figure 3.2 depicts the duration and the price discount of a sale in the steady state.[11]

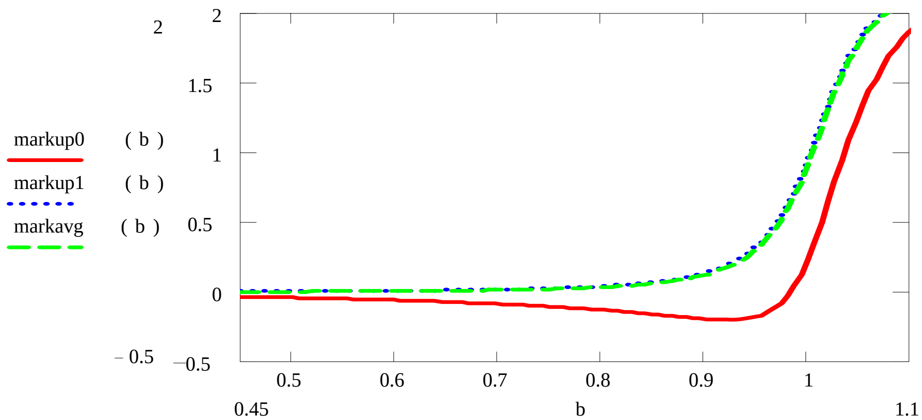

Figure 3.1. Markups as functions of b

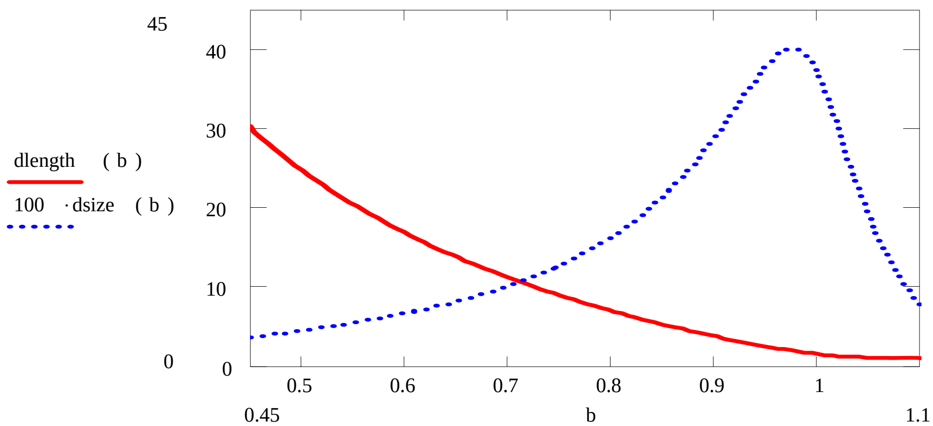

Figure 3.2. Sale duration and discount as functions of b

There are a number of notable features. First, the markdown feature in Figure 1 extends from the baseline value of b to all values of $b \leq 0.984$. The absolute value of the markdown reaches the maximum at $b = 0.92$. When b continues to increase above 0.92, a seller's incentive to use markdowns to attract buyers becomes increasingly weak, and so the markup implied by the sale price eventually becomes positive. Second, the markup implied by the regular price is positive and increases in b for all values of b. Together with the behavior of the markup implied by the sale price, this implies that the percentage price discount of a sale, $dsize$, has a hump-shaped dependence on b, as depicted in Figure 3.2. This price discount increases when b increases from low values, reaches the maximum

[11] All these values of b satisfy the condition $b \leq B(\lambda)$ required for the steady state with complete separation to exist.

around $b = 0.98$ and then starts to fall as b continues to increase. Third, in contrast to the non-monotonic dependence of the markdown and the price discount of a sale on b, the duration of the discount, *dlength*, monotonically decreases in b, as depicted in Figure 3.2. This is because the duration of the discount is a decreasing function of the queue length of buyers for each unrelated seller and, as shown in Proposition 3.3, this queue length increases monotonically with the buyer/seller ratio.

Note that reductions in b generate large, non-linear increases in the duration of a sale. For example, when b decreases from the baseline value 0.895 to 0.7 and then to 0.5, the duration of a sale increases from 4 weeks to 11.3 weeks and then to 24.8 weeks. In contrast to the non-monotonic responses of prices and the price discount, this monotonic increase in the duration of a sale with the reduction of the buyer/seller ratio is a more reliable indicator of the change in the market condition.

4.2.3. Steady-state responses to changes in λ

The third experiment is to change the probability that a buyer is active in a period, λ. Although λ affects the extensive margin of the demand as does b, it is different from b. While b affects the demand for related and unrelated sellers evenly, λ affects the demand unevenly. For λ ranging from 0.1 to 0.95, Figure 4.1 depicts the queue length of buyers per unrelated seller, q_0, and the duration of a sale in the steady state, while Figure 4.2 depicts the markups in the steady state. The queue length q_0 depends on λ non-monotonically: it increases in λ when λ increases from low values, reaches the maximum around $\lambda = 0.78$ and then decreases as λ increases further. This means that at low values of λ, the general effect of a higher λ in increasing the demand dominates but at high values of λ, the effect of a higher λ in shifting the demand to related sellers dominates. Because the duration of a sale is a decreasing function of q_0, the hump-shaped response of q_0 to λ implies that the duration of a sale has a U-shaped dependence on λ. That is, the duration of a sale is higher at both high and low values of λ than at intermediate values of λ.

The markups also depend on λ non-monotonically. At low values of λ, the markups implied by the regular price and the sale price are both positive, and both increase in λ. The markup implied by the sale price reaches the maximum around $\lambda = 0.45$, after which it falls and eventually becomes negative. The markup implied by the regular price continues to increase until $\lambda = 0.66$, after which it also starts to fall but stays positive.

As a result, the price discount of a sale (not depicted) has hump-shaped dependence on λ. These responses of the two markups to λ reflect the fact that the gain to a seller from building a relationship with a buyer is larger at intermediate values of λ than at both high and low values of λ. At low values of λ, the return to a seller from a relationship is low because the related buyer is unlikely to make a purchase in the future. At high values of λ, there is not much need to build a relationship because a buyer is readily available. At intermediate values of λ, there is a sizable benefit of giving priority to the related buyer and charging a high regular price later. Reflecting this non-monotonic dependence of the benefit of a relationship, a seller's value function (not depicted here) has hump-shaped dependence on λ and reaches the maximum around $\lambda = 0.8$, although a buyer's value monotonically increases in λ.

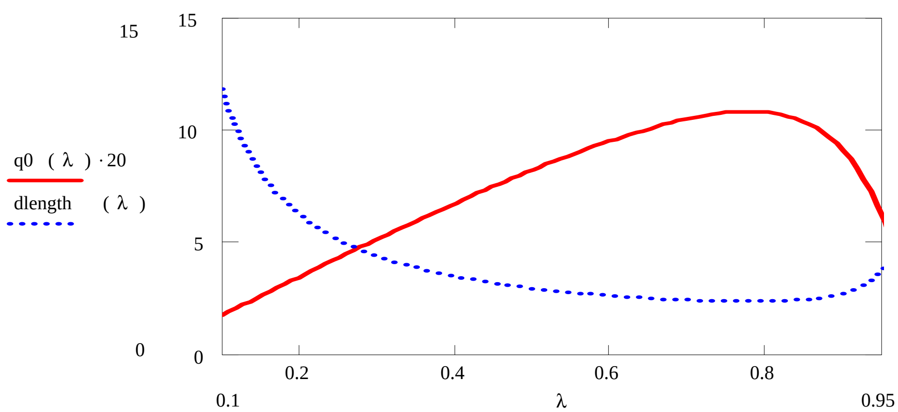

Figure 4.1. Queue length of buyers for an unrelated seller and the duration of a sale as functions of λ

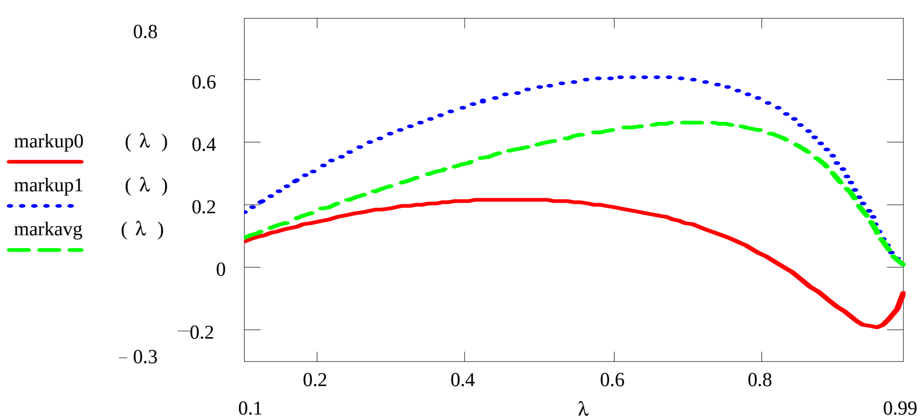

Figure 4.2. Markups as functions of λ

It is remarkable that the average markup can fall with λ when λ is large enough. Because λ is a dimension of the extensive margin of the demand, this result indicates that it is possible for prices to fall when the demand increases. This result is not as perverse as it seems. For example, Chevalier et al. (2003) find that prices of particular items in a supermarket chain typically fall when the items experience peak seasonal demand. The current model offers customer relationship as an alternative explanation for this phenomenon. It suggests that the phenomenon is more likely to arise when the peak demand comes from a store's regular customers than from general customers.

4.3. Dynamic responses of the equilibrium

Suppose that the economy is in the steady state with the baseline parameter values calibrated above. Then there is an unanticipated permanent change in one of the parameters. I compute the dynamics of the equilibrium after this change. The main purpose of this exercise is to check how the dynamics of the stock of customer relationships, ρ, affect short-run responses of the equilibrium and, in particular, whether short-run responses are significantly different from the long-run responses documented in subsection 4.2. It is straightforward to compute the dynamics of the equilibrium after a permanent change in a parameter (see the description at the end of Appendix C).

The equilibrium responds to a change in the cost in the same way in the short run as in the long run, provided that the economy is at the steady state before the cost changes. The reason is that the stock of customer relationships, ρ, and the queue length of buyers for an unrelated seller, q_0, are independent of the cost. The dynamics of these two variables are completely determined by (3.15) and (3.16), where the cost c does not appear. If the economy is in the steady state before the cost changes, then ρ and q_0 will remain in the steady state. So will the frequency and the duration of a sale. As a result, the total surplus of a match, whose dynamics obey (3.20), will have a one-time jump from the old to the new steady state. Prices, markups, and price discounts will also have a one-time jump, and so the dynamics of the equilibrium are completed in one period.

To economize on space, let me analyze the dynamics after a shock to b and omit the analysis on a shock to λ. Consider a permanent increase of b from the baseline value 0.895 to 0.95. For convenience of graphing the dynamics, let the increase in b occur at the beginning of period 2 instead of period 1. Figure 5.1 depicts the dynamics of the queue length q_0 and the fraction of related sellers ρ, while Figure 5.2 depicts the dynamics of the markups. As established in Proposition 3.3, the increase in b increases steady-state values of q_0 and b. Notice that ρ increases monotonically from the initial steady state to the new steady state, but the queue length q_0 overshoots the new steady state immediately after the shock (in period 2) and then decreases toward the new steady state. Overshooting in q_0 occurs because ρ is fixed in period 2 at the old steady-state level, which means that each seller gets a higher expected number of buyers after the increase in b than in the new steady state. Precisely, with $b < 1$, the function $q_0 = H(\rho)$ given by (3.15) is a decreasing function of ρ, and so $\rho < \rho^*$ implies $q_0 > q_0^*$. Overshooting in q_0 implies that the duration of a sale falls by more in period 2 than in the new steady state.

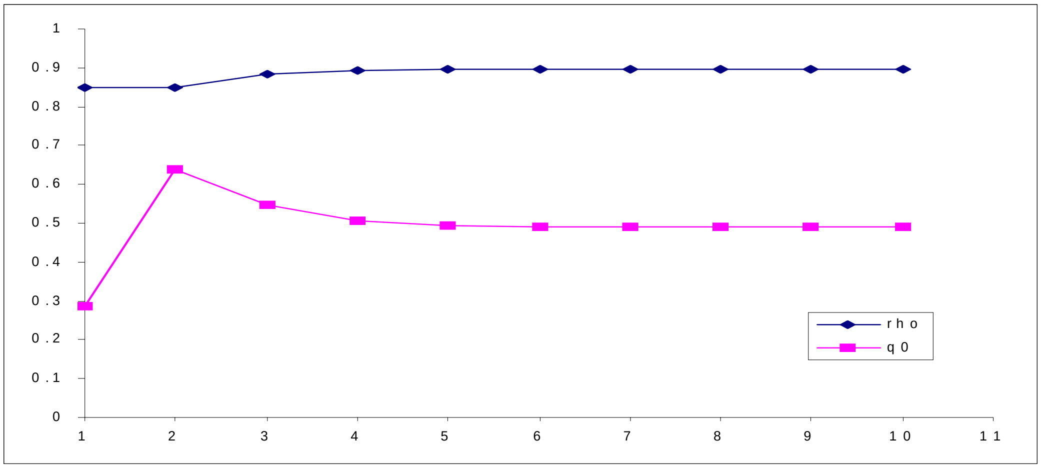

Figure 5.1. Dynamic responses of q_0 and ρ to b

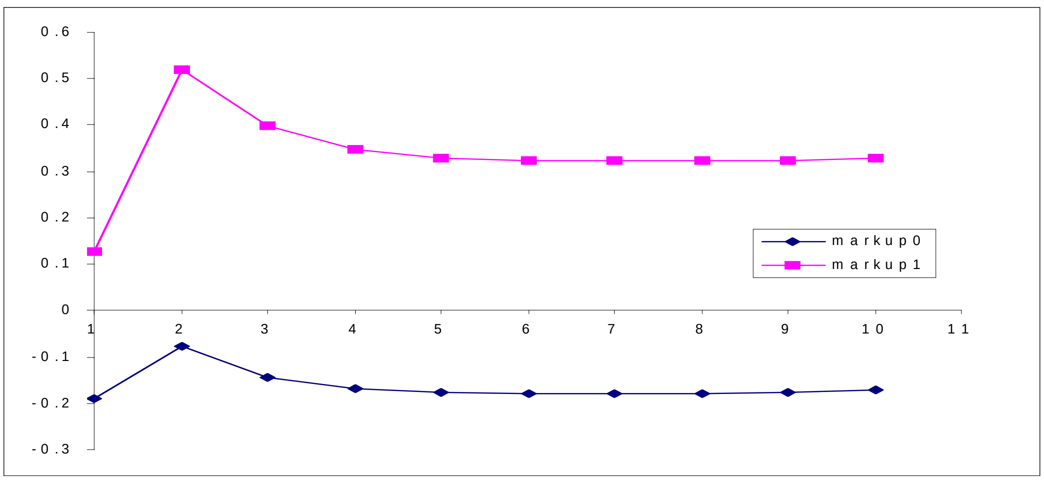

Figure 5.2. Dynamic responses of markups to b

Similarly, because ρ is fixed in period 2 at the old steady state, the increase in the buyer/seller ratio increases a seller's expected profit by more in period 2 than in the new steady state. Sellers are able to charge a higher price in period 2 than in the new steady state. Thus, the markup implied by each price overshoots the new steady state, as depicted in Figure 5.2. Notice that the transition is relatively short: the economy is very close to the new steady state four weeks after the shock. Thus, the shock does not produce persistent differences between short-run and long-run responses.

5. Free Entry of Sellers

In previous sections I have assumed that the buyer/seller ratio b is fixed. In this section, I endogenize b by allowing free entry of sellers into the market. Let k be the cost of entry and, to simplify various expressions, let it be measured at the end of the previous period. If a seller pays the cost k to enter the market, the seller is unrelated to any buyer, and so the seller's value function is $J_0(\rho)$, also measured at the end of the previous period. Free entry implies that the net value of an entry seller is zero, i.e., $J_0(\rho) = k$. Let me focus on the effect of entry on the steady state. With $J_0(\rho)$ in (3.3) and Δ^* in (C.11), the free-entry condition implies the following equation in the steady state:

$$\frac{1 + r - \lambda(1 - e^{-q_0^*})}{1 - (1 + q_0^*)e^{-q_0^*}} + 1 - (1 + r)\frac{U - c}{rk} = 0. \tag{5.1}$$

Once q_0^* is solved from this equation, the steady-state version of (3.16) determines ρ^* with $H(\rho^*) = q_0^*$, and (3.15) solves for the steady-state value of the buyer/seller ratio, b^*. Other steady-state variables can be recovered accordingly.

The following proposition states the condition for the existence of a steady state with endogenous b and the properties of this steady state (see Appendix D for a proof):

Proposition 5.1. *(i) There exists $K > 0$ such that a steady state with endogenous b and complete separation exists if and only if $k \leq K$, in which case the steady state is unique.*
(ii) In the steady state, $dq_0^*/dc > 0$, $d\rho^*/dc > 0$, $db^*/dc > 0$.
(iii) $dp_0^*/dc > 0$, $dp_1^*/dc > 0$, $d(p_1^* - p_0^*)/dc < 0$, and $d(p_1^* - c)/dc < 0$.

The existence of the steady state requires the entry cost not to be very high. If the entry cost is very high, then the expected value for an unrelated seller must increase accordingly in order to cover the entry cost. This requires the matching rate for an unrelated seller to increase and, hence, the queue length of buyers per unrelated seller to increase. Since it is relatively easy to have a trade as an unrelated seller in this case, a related seller finds it optimal to attract unrelated buyers or to give no priority to the related buyer.

An increase in the marginal cost of the good increases the queue length per unrelated seller, q_0^*, the fraction of related sellers in the market, ρ^*, and the buyer/seller ratio, b^*. These effects are easy to explain. An increase in the marginal cost reduces ex post profit. Since the entry cost has not changed, the trading probability must increase for an unrelated seller in order keep the seller's expected profit equal to the entry cost. This requires the queue length for an unrelated seller to increase. As each unrelated seller succeeds in trading more likely than before, the flow of sellers from unrelated ones into related ones increases, and so the fraction of related sellers in the steady state, ρ^*, increases. Also, for the queue length of buyers per unrelated seller to increase, there must be fewer sellers than before; that is, the buyer/seller ratio, b^*, increases. Note that because the increase in the cost increases the queue length, it reduces the duration of a sale, in contrast to the constant queue length when b is fixed.

Part (iii) of Proposition 5.1 reveals that the increase in the cost increases the two prices unevenly and it increases the prices not by the same amount as the increase in the cost itself. First, the sale price increases with the cost by more than the regular price does. This difference between the two prices' responses to the cost is similar to that in subsection 4.2.1, and it is reinforced here by the increase in the buyer/seller ratio induced by the increase in the cost. As a result of this difference, the increase in the cost reduces the percentage of the price discount. Second, the regular price increases by less than the increase in the cost, again reflecting the feature that the regular price is relatively more stable with respect to the change in the cost. This result implies that the markup implied by the regular price falls as the cost increases. The response of the markup implied by the sale price to the cost is ambiguous analytically.

On the quantitative side, I can calibrate the entry cost k so that the steady-state buyer/seller ratio at the baseline value of c is equal to the one identified in Table 1. With this value of k, the markup implied by the sale price is negative and increases with the cost. More importantly, the responses of the two markups to the cost (not graphed here) are very close to the ones depicted in Figure 2. In particular, it is difficult to discern the difference between the responses of the markdown by the sale price in the case with an endogenous b and the case with a fixed b. The response of the markup by the regular price is slightly flatter in the case with an endogenous b than in the case with a fixed b. Thus, endogenizing b does not change the responses of prices to the cost by much. Instead, the main change in the effect is that an increase in the cost reduces the duration of a sale sizably when b is endogenous, in contrast to the constant duration when b is fixed. This contrast between the two cases provides another illustration for the earlier statement that the duration of a sale reflects the market conditions more accurately than prices or markups do.

The effects of an increase in the entry cost are similar to an increase in the cost of a good, and so they are omitted here. The effects of an increase in the probability of a buyer being active, λ, are quite different in the case with an endogenous b from those in the case of a fixed b. It is easy to verify from (5.1) that when b is endogenous, an increase in λ leads to a reduction in q_0^* rather than the non-monotonic response depicted in Figure 4.1. This is because for most values of λ except high λ, an increase in λ increases the demand and induces more sellers to enter the market, which drives down the queue length of buyers per seller. As a result, the duration of a sale increases with λ. Moreover, when b is endogenous, both the magnitude and the response of a markup to λ are different from those depicted in Figure 4.2 for the case with a fixed b. With an endogenous b, the two markups are positive when λ is low and they decrease with λ for all values of λ. The markup implied by the sale price eventually becomes negative as λ becomes sufficiently large. Thus, with free entry of sellers, an increase in λ reduces prices for all values of λ, not just for high values of λ. The increasing part of the response of markups to λ in Figure 4.2 is reversed here by the effect of a falling buyer/seller ratio.

6. Conclusion

I construct a search model to formalize the intuitive idea that sellers hold sales to attract buyers and build customer relationships. The market consists of a large number of buyers and sellers. All sellers sell a homogeneous good and all buyers have the same publicly known valuation of the good. Buyers know the terms of trade offered by sellers before choosing which seller to visit. A buyer is related to a seller if the buyer just bought a good from the seller and the relationship is broken if the buyer fails to continue to buy from the seller. Sellers are restricted to offer the same price to all buyers, but they are allowed to give priority to their related buyers. I prove that there is an equilibrium in which a seller gives priority to the related buyer and a buyer makes repeat purchases from the related seller. In the equilibrium, a seller who does not have a related buyer posts a low (sale) price to attract the buyers who are unrelated to any seller and, once the seller is related to a buyer after a trade, the seller will post a high (regular) price to sell only to the related buyer. The fraction of related sellers is endogenous in the equilibrium. I calibrate the steady state of the model to the data and find that the sale price represents a sizable markdown, and the regular price a sizable markup, on the marginal cost. With the calibrated model, I examine comparative statics and dynamics of the equilibrium with respect to changes in the cost of and the demand for the good.

Aside from formalizing a theory on customer relationships and sales, this model is intended to be a step toward building a macro model in which sales are an important part of price adjustments in aggregate fluctuations. There is still some distance to go from the current model to such a macro model. The model needs to incorporate money and add fixed costs of changing prices in order to make nominal prices sticky. However, the current model holds some promises. In contrast to other models of sales, this model has many buyers and sellers, it is dynamic and it endogenizes customer relationships. The calibration shows that customer relationship can be an important driving force of price fluctuations. It can generate markdowns as well as markups and can explain some puzzling price behavior in the data. Moreover, the regular price is less responsive to shocks than the sale price, which is consistent with the microdata (Klenow and Kryvtsov, 2008, and Nakamura and Steinsson, 2008). Finally, the model shows that the duration of a sale responds to shocks more accurately than prices do. A macro model with sales should explicitly incorporate the duration of a sale as part of price adjustments.[12]

[12]Recently, Guimaraes and Sheedy (2009) and Kehoe and Midrigan (2010) incorporate sales into a macro model with sticky prices. Guimaraes and Sheedy (2009) assume that households are loyal to certain brands exogenously. Kehoe and Midrigan (2010) assume that sale prices and regular prices are governed by different processes and different menu costs.

Appendix

A. Derivation of Trading Probabilities

I calculate the probability that a type-0 buyer B who visits a particular seller A of type-0 is chosen by seller A. Seller A posts price p. Each active type-i buyer visits seller A with probability $\theta_{i0}(\rho, p)$, where $i \in \{0, 1\}$. The queue length of type-1 buyers visiting seller A is $q_{10}(\rho, p) = \lambda\rho M\theta_{10}(\rho, p)$, and the queue length of type-0 buyers visiting seller A is $q_{00}(\rho, p) = \lambda(N - \rho M)\theta_{00}(\rho, p)$. To simplify the notation below, I suppress the dependence of θ_{10}, θ_{00}, q_{10} and q_{00} on (ρ, q).

Each buyer who has a related seller visits seller A with probability $\lambda\theta_{10}$. Because there are ρM such buyers, seller A will be visited by n_1 type-1 buyers with probability $C_{\rho M}^{n_1} (\lambda\theta_{10})^{n_1} (1 - \lambda\theta_{10})^{\rho M - n_1}$, where $C_{\rho M}^{n_1} = (\rho M)!/ [(n_1)! (\rho M - n_1)!]$. In addition, each type-0 buyer visits seller A with probability $\lambda\theta_{00}$. Because there are $N - \rho M - 1$ such buyers other than buyer B, seller A will be visited by n_0 other type-0 buyers with probability $C_{N-\rho M-1}^{n_0} (\lambda\theta_{00})^{n_0} (1 - \lambda\theta_{00})^{N-\rho M-1-n_0}$. If seller A is visited by n_1 type-1 buyers and n_0 type-0 buyers in addition to buyer B, then buyer B will be chosen by seller A with probability $1/(n_1 + n_0 + 1)$. Considering all realizations of n_1 and n_0, I conclude that buyer B will be chosen by seller A with probability $g(1)$, where $g(x)$ is defined for $x \in [0, 1]$ as

$$g(x) = \sum_{n_1=0}^{\rho M} \sum_{n_0=0}^{N-\rho M-1} \frac{1}{n_1 + n_0 + 1} \left[\begin{array}{l} C_{\rho M}^{n_1} (x\lambda\theta_{10})^{n_1} (1 - \lambda\theta_{10})^{\rho M - n_1} \\ \times C_{N-\rho M-1}^{n_0} (x\lambda\theta_{00})^{n_0} (1 - \lambda\theta_{00})^{N-\rho M-1-n_0} \end{array} \right].$$

To compute $g(x)$, let me first compute:

$$\begin{aligned} \tfrac{d}{dx}[xg(x)] &= \sum_{n_1=0}^{\rho M} \left[C_{\rho M}^{n_1} (x\lambda\theta_{10})^{n_1} (1 - \lambda\theta_{10})^{\rho M - n_1}\right] \\ &\quad \times \sum_{n_0=0}^{N-\rho M-1} \left[C_{N-\rho M-1}^{n_0} (x\lambda\theta_{00})^{n_0} (1 - \lambda\theta_{00})^{N-\rho M-1-n_0}\right] \\ &= [x\lambda\theta_{10} + 1 - \lambda\theta_{10}]^{\rho M} [x\lambda\theta_{00} + 1 - \lambda\theta_{00}]^{N-\rho M-1}. \end{aligned}$$

Note that $g(x)$ and $xg(x)$ are bounded for all $x \in [0, 1]$, and so they are integrable. Integrating the above result from $x = 0$ to $x = 1$ yields:

$$xg(x) = \int_0^x [1 - (1 - y)\lambda\theta_{10}]^{\rho M} [1 - (1 - y)\lambda\theta_{00}]^{N-\rho M-1} dy.$$

Taking the limit M, $N \to \infty$ (with $N/M = b$ being fixed) and using the definition of queue lengths, I get: $[1 - (1 - y)\lambda\theta_{10}]^{\rho M} \to e^{-q_{10}(1-y)}$ and $[1 - (1 - y)\lambda\theta_{00}]^{N-\rho M-1} \to e^{-q_{00}(1-y)}$.

Thus, when visiting a seller who does not have a related buyer, a buyer is chosen to trade with by the seller with probability:

$$xg(x) \to \int_0^x e^{-(q_{10}+q_{00})(1-y)}dy = \frac{e^{-(q_{10}+q_{00})(1-x)} - e^{-(q_{10}+q_{00})}}{q_{10}+q_{00}}.$$

Setting $x = 1$ yields $g(1)$ that is the probability used in the main text.

B. Proofs of Lemmas 3.1 and 3.2

Let me prove Lemma 3.1 first. Because $p_1 < U + V_1(\rho_{+1}) - V_0(\rho_{+1})$, comparing (2.3) and (2.6) yields $v_{1s} > v_{01}$ $(= v_{11})$. Then (2.2) implies $\theta_{11} = 0$, as stated in (i) of the lemma. Now, the results $\theta_{10} = 0$ and $\theta_{1s} = 1$ are equivalent to each other. To prove $\theta_{10} = 0$, there are two cases to consider. The first case has $v_{00} \leq v_{01}$. In this case, $v_{10} = v_{00} \leq v_{01} < v_{1s}$. The strict inequality $v_{10} < v_{1s}$ implies $\theta_{10} = 0$. The second case has $v_{00} > v_{01}$. In this case, $\theta_{00} = 1$, and so $q_{01} = 0$. Recall that $q_{11} = 0$ (since $\theta_{11} = 0$). From (2.12), I can compute the payoff to a seller who has a related buyer in this case as

$$J_1(\rho) = \frac{1}{1+r} \max_{p_1} \left\{ \lambda\theta_{1s}[p_1 - c + J_1(\rho_{+1}) - J_0(\rho_{+1})] + J_0(\rho_{+1}) \right\}, \text{ s.t. } (2.7).$$

If $\theta_{1s} < 1$, the seller can increase the payoff by reducing p_1 slightly to induce $\theta_{1s} = 1$. This price reduction does not change any other buyer's choice, because the seller only attracts his related buyer in this case. Thus, the reduction will increase the seller's value, which implies that $\theta_{1s} < 1$ cannot be equilibrium outcome. Now that $\theta_{1s} = 1$, I have $\theta_{10} = 0$.

To prove (ii) of Lemma 3.1, suppose $\theta_{00} = 0$ (i.e., $\theta_{01} = 1$), to the contrary. Since $q_{00} = 0$ in this case, and since $q_{10} = 0$ (as a result of $\theta_{10} = 0$), then (2.5) yields $v_{00} = U - p_0 + V_1(\rho_{+1})$. A seller without a related buyer can set $p_0 = p_1 - \varepsilon$, where $\varepsilon > 0$ is sufficiently small. Doing so will yield $v_{00} > v_{1s} > v_{01}$ and, hence, $\theta_{00} = 1$ that contradicts the supposition $\theta_{01} = 1$.

To prove (iii) of Lemma 3.1, note that $\theta_{01} > 0$ implies $v_{00} \leq v_{01}$. Since $v_{1s} > v_{01}$, then $v_{1s} > v_{00} = v_{10}$ in this case.

For (iv) of Lemma 3.1, note that $v_{1s} \geq v_{10} = v_{00}$, where the inequality follows from $\theta_{1s} = 1$. If $\theta_{01} = 0$, the formula of J_1 in the above proof is valid. If $v_{1s} > v_{00}$ in this case, a seller with a related buyer can raise p_1 slightly without disturbing the outcome $\theta_{1s} = 1$, thus increasing his payoff. Therefore, $v_{1s} = v_{00}$ must hold if $\theta_{01} = 0$.

Now turn to Lemma 3.2. As discussed in the main text, if (3.7) is binding for a type-1 seller, then such a seller can gain by deviating to $\bar{p}$ that attracts only the related buyer, and an equilibrium with partial mixing does not exist. So, suppose that (3.7) is not binding. Then, a type-1 seller's optimal choice q_1 satisfies (3.9), the value function J_1 satisfies (3.10), and the relation $q_0 = h(q_1)$ holds, where h is defined in (3.11). I prove that these conditions lead to the result $J_0(\rho) > J_1(\rho)$, which violates the equilibrium requirement that it be optimal for a type-1 seller to give priority to the related buyer.

Because $(1+r)J_0(\rho)$ is given by (3.3) and $(1+r)J_1(\rho)$ by (3.10), the relation $J_0(\rho) > J_1(\rho)$ is equivalent to:

$$1-(1+q_0)e^{-q_0} > \frac{\left[1-(1+q_1)e^{-q_1}\right]\left[1-(1-\lambda)e^{-q_1}\right]^2}{(1-\lambda)\left[1-e^{-q_1}\right]^2+\lambda\left[1-(1+q_1)e^{-q_1}\right]}.$$

Here I have used the fact that $\Delta(\rho_{+1}) > 0$, where Δ is defined in (3.4). Substituting $q_0 = h(q_1)$ from (3.11), I rewrite the above condition as $f(a, q_1) < 0$, where a temporarily denotes $a = \frac{1}{1-\lambda}$ and f temporarily denotes

$$f(a,q) = \ln\left\{a\left[1+(a-1)\frac{1-(1+q)e^{-q}}{(1-e^{-q})^2}\right]\right\} - (a-1)\frac{\left[q_1+1-(2q_1+1)e^{-q_1}\right]}{(1-e^{-q_1})^2}.$$

Note that $a > 1$. Also, $f(1, q) = 0$ for all $q > 0$. Compute:

$$\frac{\partial f(a,q)}{\partial a} = \frac{1}{a} + \left[a-1+\frac{(1-e^{-q})^2}{1-(1+q)e^{-q}}\right]^{-1} - \frac{\left[q_1+1-(2q_1+1)e^{-q_1}\right]}{(1-e^{-q_1})^2}.$$

From this expression it is easy to verify that

$$\frac{\partial f(1,q)}{\partial a} = 1 - \frac{q}{1-e^{-q}} < 0 \text{ for all } q > 0.$$

It is also easy to verify that $\frac{\partial f(a,q)}{\partial a}$ is decreasing in a for all $q > 0$ and all $a > 1$. Thus, for all $a > 1$ and $q > 0$, the following results hold:

$$\frac{\partial f(a,q)}{\partial a} < \frac{\partial f(1,q)}{\partial a} < 0, \qquad f(a,q) < f(1,q) = 0.$$

This establishes the result $J_0(\rho) > J_1(\rho)$ under partial mixing and, hence, proves that an equilibrium with partial mixing does not exist. **QED**

C. Proof of Proposition 3.3 and the Computation of Dynamics

For part (i) of Proposition 3.3, it is clear that the steady state and the dynamics of ρ are determined by (3.16), given the initial value of ρ. For each value of ρ on the dynamic path, q_0 is given by (3.15). Because (3.15) and (3.16) depend only on (λ, b), the steady state and the dynamics of (ρ, q_0) depend on (λ, b) but not on other parameters such as (c, U). Let me solve the steady state of (ρ, q_0), denoted (ρ^*, q_0^*). In the steady state, $\rho_{+1} = \rho = \rho^*$, and so (3.16) and $H(\rho) = q_0$ yield:

$$\rho^* = \rho 1(q_0^*) \equiv \left[1 + \frac{1-\lambda}{1-e^{-q_0^*}}\right]^{-1}. \tag{C.1}$$

Rewriting (3.15) as $\rho = b\left[1 + \frac{1}{\lambda}(\frac{1}{\rho} - 1)q_0\right]^{-1}$ and substituting $\rho^* = \rho 1(q_0^*)$, I get:

$$\rho^* = \rho 2(q_0^*) \equiv b\left[1 + \frac{1-\lambda}{\lambda}\frac{q_0^*}{1-e^{-q_0^*}}\right]^{-1}. \tag{C.2}$$

The steady-state values, (ρ^*, q_0^*), solve $\rho^* = \rho 1(q_0^*) = \rho_2(q_0^*)$. It is easy to verify that $\rho 1'(q) > 0$, $\rho 1(0) = 0$, $\rho 1(\infty) = \frac{1}{2-\lambda} < 1$, $\rho 2'(q) < 0$, $\rho 2(0) = b\lambda > 0$, and $\rho 2(\infty) = 0$. Thus, there exists a unique $q_0^* \in (0, \infty)$ that solves $\rho 1(q_0^*) = \rho 2(q_0^*)$. The implied solution for ρ^* satisfies $\rho^* \in (0, 1)$ because $\rho 1(q_0^*) \in (0, 1)$. Also, $\rho^* < b\lambda$ because $\rho 2(q_0^*) < b\lambda$. Thus, $0 < \rho^* < \min\{1, b\lambda\}$, as it is stated in part (i) of the proposition. Also, it is straightforward to use (C.1) and (C.2) to verify that $\frac{d\rho^*}{db} > 0$, $\frac{dq_0^*}{db} > 0$, and $\frac{d\rho^*}{d\lambda} > 0$.

Continuing the proof of part (i), I show that the steady state ρ^* is locally stable. This amounts to proving $|G'(\rho^*)| < 1$, where G is defined in (3.16). Using (3.15) to compute $H'(\rho) = (q_0 - \lambda)/(1 - \rho)$ first and then G', I have:

$$G'(\rho) = (1 - \lambda + q_0)e^{-q_0} - (1 - \lambda), \text{ where } q_0 = H(\rho).$$

Clearly, $G'(\rho) > -(1-\lambda) > -1$. Also, $G'(\rho) < q_0 e^{-q_0} < 1$. Thus, $|G'(\rho)| < 1$ for all ρ such that $q_0 = H(\rho) > 0$ and, clearly, for $\rho = \rho^*$.

For part (ii), let me first presume $\Delta(\rho_{+1}) > 0$, which will be verified in part (iii), and find the conditions under which the requirements $J_1(\rho) \geq J_0(\rho)$ and (3.14) are satisfied. Consider the condition $J_1(\rho) \geq J_0(\rho)$. Using (3.8) and V_0^a in (3.2), I can compute:

$$\bar{p} = U + V_1(\rho_{+1}) - V_0(\rho_{+1}) - e^{-q_0}\Delta(\rho_{+1}). \tag{C.3}$$

Then, a type-1 seller's expected surplus and the value function are

$$\lambda\left[\bar{p} - c + J_1(\rho_{+1}) - J_0(\rho_{+1})\right] = \lambda(1 - e^{-q_0})\Delta(\rho_{+1}), \tag{C.4}$$

$$(1 + r)J_1(\rho) = J_0(\rho_{+1}) + \lambda(1 - e^{-q_0})\Delta(\rho_{+1}). \tag{C.5}$$

Comparing (C.5) with (3.3), I express the condition $J_1(\rho) \geq J_0(\rho)$ equivalently as $\frac{q_0}{e^{q_0}-1} \geq 1-\lambda$. Because $\frac{q}{e^q-1}$ is a decreasing function of q for all $q > 0$ and its value lies in $(0,1)$, this condition is equivalent to $q_0 \leq q_a(\lambda)$ where q_a is defined by the following equation:

$$\frac{q_a}{e^{q_a}-1} = 1-\lambda. \tag{C.6}$$

Note that $q_a(\lambda)$ is an increasing function of λ. Moreover, because $e^\lambda < \frac{1}{1-\lambda}$ for all $\lambda \in (0,1)$, it can be verified that $\frac{\lambda}{e^\lambda-1} > 1-\lambda$, which implies $q_a(\lambda) > \lambda$.

Now consider the condition (3.14), still under the presumption that $\Delta(\rho_{+1}) > 0$. Substituting $\tilde{p}_1$ from (3.13), $\bar{p}$ from (C.3), and $q_0 = h(\tilde{q}_1)$, where h is defined in (3.11), I can write (3.14) equivalently as $f(\tilde{q}_1) \leq 0$, where f temporarily denotes:

$$f(q) = e^{2q} - \left[2 - 2\lambda + \lambda^2 + (1+\lambda)q\right] e^q + (1-\lambda)(1-\lambda+q).$$

Note that $f(0) = 0$, $f(\infty) = \infty$ and compute:

$$f'(q) = 2e^{2q} - \left[3 - \lambda + \lambda^2 + (1+\lambda)q\right] e^q + 1 - \lambda.$$

Note that $f'(0) = -\lambda^2 < 0$, $f'(\infty) = \infty$ and compute:

$$f''(q) = e^q \left[4e^q - 4 - \lambda^2\right].$$

Thus, $f''(q) > 0$ if and only if $q > \ln(1 + \frac{\lambda^2}{4})$. With the properties of $f'(0)$ and $f'(\infty)$, this result implies that there exists $q_4 \in (0, \ln(1+\frac{\lambda^2}{4}))$ such that $f'(q) > 0$ if and only if $q > q_4$. In turn, with the properties of $f(0)$ and $f(\infty)$, this result implies that there exists $q_5(\lambda) \in (q_4, \infty)$ such that $f(q) > 0$ if and only if $q > q_5(\lambda)$. That is, $f(\tilde{q}_1) \leq 0$ if and only if $\tilde{q}_1 \leq q_5(\lambda)$. It can be verified that $h(q)$ defined in (3.11) is an increasing function. Because $q_0 = h(\tilde{q}_1)$, then $\tilde{q}_1 \leq q_5(\lambda)$ if and only if

$$q_0 \leq q_b(\lambda) \equiv h(q_5(\lambda)) \in (0, \infty). \tag{C.7}$$

Therefore, the requirement $J_1(\rho) \geq J_0(\rho)$ and the requirement (3.14) are both satisfied if and only if $q_0 \leq Q(\lambda)$, where $Q(\lambda)$ is defined as

$$Q(\lambda) = \min\{q_a(\lambda), q_b(\lambda)\} \in (0, \infty). \tag{C.8}$$

Because $\rho 2(q) < \rho 1(q)$ if and only if $q > q_0^*$, the steady state of the equilibrium satisfies $q_0^* \leq Q(\lambda)$ if and only if $\rho 2(Q(\lambda)) \leq \rho 1(Q(\lambda))$, which can be rewritten as $b \leq B(\lambda)$ where

$$B(\lambda) \equiv \frac{\lambda[1 - e^{-Q(\lambda)}] + (1-\lambda)Q(\lambda)}{\lambda\,[2 - \lambda - e^{-Q(\lambda)}]}. \tag{C.9}$$

Together with part (i), this establishes that a steady state with complete separation exists if and only if $b \leq B(\lambda)$. Under this condition, the steady state is unique. It is clear from the definitions of $q_a(\lambda)$ and $q_b(\lambda)$ that they depend only on λ and not on other parameters such as c and U. Thus, $Q(\lambda)$ and $B(\lambda)$ depend only on λ and not on (c, U).

For part (iii), note that the construction of the price $\bar{p}$ in (3.8) implies $v_{1s}(\rho, \bar{p}) = V_0^a(\rho)$. Hence, (2.1) implies $V_1(\rho) = V_0(\rho)$. Substituting V_0^a from the first condition in (3.2) into (2.1) for $i = 0$ yields the expression for V in (3.19). Combining the two conditions in (3.2) to eliminate V_0^a yields the expression for p_0 in (3.17) and substituting the result $V_1 = V_0$ into (C.3) yields the expression for $\bar{p}$ in (3.18). Because $\frac{q}{1-e^{-q}} > 1$ for all $q > 0$ and $\Delta > 0$, it is clear that $\bar{p} > p_0$. Let me delay the proof of $\bar{p} > c$.

To derive the dynamic equation for Δ in (3.20) and verify $\Delta(\rho) > 0$ for ρ near the steady-state value ρ^*, I subtract (3.3) from (C.5) to get:

$$J_1(\rho) - J_0(\rho) = \frac{1}{1+r}\left[\lambda(1 - e^{-q_0}) - 1 + (1 + q_0)e^{-q_0}\right]\Delta(\rho_{+1}). \tag{C.10}$$

The definition of $\Delta(\rho)$ in (3.4) and the result $V_1 = V_0$ then imply the dynamic equation for Δ in (3.20). Setting $\Delta(\rho) = \Delta(\rho_{+1}) = \Delta^*$ and $q_0 = q_0^*$ in (3.20), I obtain:

$$\Delta^* = \frac{(1+r)(U-c)}{1 + r - \lambda(1 - e^{-q_0^*}) + 1 - (1 + q_0^*)e^{-q_0^*}}. \tag{C.11}$$

Because $r > 0$, $1 > \lambda(1 - e^{-q_0^*})$, $1 > (1 + q_0^*)e^{-q_0^*}$ and $U > c$, then $\Delta^* > 0$. If ρ is close to ρ^*, then $q_0 = H(\rho)$ is close to q_0^*, in which case $\Delta(\rho)$ is close to Δ^* and, hence, positive.

With (C.11) and (3.18), I can deduce that $\bar{p}^* > c$ if and only if

$$(1 + r - \lambda)(1 - e^{-q_0^*}) + 1 - (1 + q_0^*)e^{-q_0^*} > 0.$$

Note that the left-hand is an increasing function of q_0^* and its value is 0 if $q_0^* = 0$. Because $q_0^* > 0$, then the left-hand side is strictly greater positive, indeed. This implies that if ρ is close to ρ^*, then $\bar{p} > c$.

Finally, I describe the procedure of computing the dynamics of the equilibrium. Starting with any initial value of ρ, the dynamic path of ρ can be solved by repeatedly using (3.16). For each ρ on this dynamic path, the value of q_0 is given by $q_0 = H(\rho)$, where H is defined in (3.15). Let $\{\rho_{+t}\}_{t=0}^{\infty}$ be the path of ρ and $\{q_{0,+t}\}_{t=0}^{\infty}$ the path of q_0. To solve the dynamics of other variables, let me start with Δ, the total surplus of a match.

From (3.20), it is clear that Δ obeys a linear difference equation whose coefficients are time-varying due to the dynamics of q_0. Taking T to be a sufficiently large number and choosing $\Delta(\rho_{+T})$ to be sufficiently close to the new steady state level of Δ, I can iterate on (3.20) backward to obtain the sequence $\{\Delta(\rho_{+t})\}_{t=0}^{T}$. Similarly, I can iterate on (3.19) backward to obtain $\{V_i(\rho_{+t})\}_{t=0}^{T}$ and on (3.3) to obtain $\{J_0(\rho_{+t})\}_{t=0}^{T}$. Then, (C.10) gives $\{J_1(\rho_{+t})\}_{t=0}^{T}$. Substituting the dynamic paths of q_0 and Δ into (3.17) and (3.18) yields the dynamic paths of p_0 and p_1. Similarly, substituting the dynamic paths of (ρ, q_0, p_0, p_1) into (3.21) - (3.25) yields the dynamic paths of the probability, length, and the price discount of a sale as well as the markups. **QED**

D. Proof of Proposition 5.1

For (i), note that a steady state with complete separation and endogenous b exists if and only if the steady-state value b^* satisfies $b^* \leq B(\lambda)$, where B is defined in (C.9). This requirement requires equivalently that (5.1) should have a solution for q_0^* which satisfies $q_0^* \leq Q(\lambda)$, where Q is defined in (C.8). Temporarily denote the left-hand side of (5.1) as $LHS(q_0^*)$. Because $(1-e^{-q})$ and $[1-(1+q)e^{-q}]$ are increasing functions of q, then $LHS(q)$ is a decreasing function. Because $LHS(0) = \infty > 0$, (5.1) has a solution for q_0^* that satisfies $q_0^* \leq Q(\lambda)$ if and only if $LHS(Q(\lambda)) \leq 0$. I express this condition as

$$k \leq K \equiv (U-c)(\frac{1}{r}+1) \Big/ \left[\frac{1+r-\lambda(1-e^{-Q(\lambda)})}{1-[1+Q(\lambda)]e^{-Q(\lambda)}}+1\right]. \tag{D.1}$$

Clearly, $K > 0$. Also, under (D.1), the solution for q_0^* to (5.1) is unique, and so the solution for b^* to (3.15) is unique.

For (ii), it is easy to verify from (5.1) that $dq_0^*/dc > 0$. Because $\rho^* = \rho 1(q_0^*)$, where $\rho 1$ is an increasing function given in (C.1), it is clear that $d\rho^*/dc > 0$. From (C.2) I get:

$$b^* = \rho^* \left[1 + \frac{1-\lambda}{\lambda}\frac{q_0^*}{1-e^{-q_0^*}}\right].$$

Because $q/(1-e^{-q})$ is an increasing function of q, the results $dq_0^*/dc > 0$ and $d\rho^*/dc > 0$ imply that $db^*/dc > 0$.

To establish (iii), note that $rJ_0^* = [1-(1+q_0^*)e^{-q_0^*}]\Delta^*$. The free-entry condition $J_0^* = k$ implies: $\Delta^* = rk/[1-(1+q_0^*)e^{-q_0^*}]$. Substituting Δ^* into (3.17) and (3.18) yields:

$$p_0^* = U - \frac{rk\ q_0^*}{(e^{q_0^*}-1)[1-(1+q_0^*)e^{-q_0^*}]}, \qquad p_1^* = U - \frac{rk\ e^{-q_0^*}}{1-(1+q_0^*)e^{-q_0^*}}.$$

Because both expressions are increasing functions of q_0^*, the result $dq_0^*/dc > 0$ implies $dp_0^*/dc > 0$ and $dp_1^*/dc > 0$. The difference between the two prices is:

$$p_1^* - p_0^* = \frac{rk(q_0^* - 1 + e^{-q_0^*})}{(1 - e^{-q_0^*})(e^{q_0^*} - 1 - q_0^*)}.$$

The derivative of this difference with respect to q_0^* has the same sign as that of

$$(1 - e^{-q_0^*})[(2 - q_0^*)e^{q_0^*} - 2 - q_0^*] - (q_0^* - 1 + e^{-q_0^*})[1 - (1 + q_0^*)e^{-q_0^*}].$$

Note that $q - 1 + e^{-q} > 0$ and $1 - (1 + q)e^{-q} > 0$ for all $q > 0$. Also, the function $[(2 - q)e^q - 2 - q]$ is equal to 0 when $q = 0$, and its derivative with respect to q is equal to $-[(q - 1)e^q + 1] < 0$. Thus, the function is negative for all $q > 0$. These results imply that the above expression for $(p_1^* - p_0^*)$ decreases in q_0^*. Because $dq_0^*/dc > 0$, then $d(p_1^* - p_0^*)/dc < 0$. Finally, using the above expression for p_1^* to compute $(p_1^* - c)$ and using (5.1) to substitute $(U - c)$, I obtain:

$$p_1^* - c = rk\left[1 + \frac{(1 + r - \lambda)e^{q_0^*} + \lambda - 1}{e^{q_0^*} - 1 - q_0^*}\right].$$

The derivative of this expression with respect to q_0^* has the same sign as that of

$$(1 - \lambda)(1 - e^{-q_0^*}) - (1 + r - \lambda)q_0^*.$$

Because $r > 0$, this expression is less than $(1 - \lambda)[1 - e^{-q_0^*} - q_0^*] < 0$. Thus, $(p_1^* - c)$ decreases in q_0^* and, hence, decreases in c. **QED**

References

[1] Acemoglu, D. and R. Shimer, 1999, "Efficient Unemployment Insurance," Journal of Political Economy 107, 893-928.

[2] Aguirregabiria, V., 1999, "The Dynamics of Markups and Inventories in Retailing Firms," Review of Economic Studies 66, 275-308.

[3] Bagwell, K., 1987, "Introductory Price as a Signal of Cost in a Model of Repeat Business," Review of Economic Studies 54, 365-384.

[4] Blattberg, R.C. and S.K. Sen, 1974, "Market Segmentation Using Models of Multidimensional Purchasing Behavior," Journal of Marketing 38, 17-28.

[5] Burdett, K., and K. Judd, 1983, "Equilibrium Price Dispersion," Econometrica 51, 955-970.

[6] Burdett, K., Shi, S. and R. Wright, 2001, "Pricing and Matching with Frictions," Journal of Political Economy 109, 1060-1085.

[7] Chevalier, J.A., Kashyap, A.K. and P.E. Rossi, 2003, "Why Don't Prices Rise during Periods of Peak Demand? Evidence from Scanner Data," American Economic Review 93, 15-37.

[8] Dutta, S. Bergen, M., and D. Levy, 2002, "Price Flexibility in Channels of Distribution: Evidence from Scanner Data," Journal of Economic Dynamics and Control 26, 1845-1900.

[9] Gonzalez, F.M. and S. Shi, 2010, "An Equilibrium Theory of Learning, Search and Wages," Econometrica 78, 509-537.

[10] Guimaraes, B. and K.D. Sheedy, 2009, "Sales and Monetary Policy," Working paper, London School of Economics.

[11] Kehoe, P.J. and V. Midrigan, 2010, "Prices Are Sticky After All," Working paper, Federal Reserve Bank of Minneapolis.

[12] Klenow, P. and O. Kryvtsov, 2008, "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent US Inflation?" Quarterly Journal of Economics 123, 863-904.

[13] Lazear, E.P., 1986, "Retail Pricing and Clearance Sales," American Economic Review 76, 14-32.

[14] Milgrom, P. and J. Roberts, 1986, "Price and Advertising Signals of Product Quality," Journal of Political Economy 94, 796-821.

[15] Moen, E.R., 1997, "Competitive Search Equilibrium," Journal of Political Economy 105, 385-411.

[16] Montgomery, J.D., 1991, "Equilibrium Wage Dispersion and Interindustry Wage Differentials," Quarterly Journal of Economics 106, 163-179.

[17] Nakamura, E. and J. Steinsson, 2008, "Five Facts about Prices: A Reevaluation of Menu Cost Models," Quarterly Journal of Economics 123, 1415-1464.

[18] Pesendorfer, M., 2002, "Retail Sales: A Study of Pricing Behavior in Supermarkets," Journal of Business 75, 33-66.

[19] Peters, M., 1991, "Ex Ante Price Offers in Matching Games: Non-Steady State," Econometrica 59, 1425-1454.

[20] Salop, S. and J. Stiglitz, 1982, "The Theory of Sales: A Simple Model of Equilibrium Price Dispersion with Identical Agents," American Economic Review 72, 1121-1130.

[21] Shi, S., 2002, "A Directed Search Model of Inequality with Heterogeneous Skills and Skill-Biased Technology," Review of Economic Studies 69, 467-491.

[22] Shi, S., 2006, "Wage Differentials, Discrimination and Efficiency," European Economic Review 50, 849-875.

[23] Shi, S., 2009, "Directed Search for Equilibrium Wage-Tenure Contracts," Econometrica 77, 561-584.

[24] Shilony, Y., 1977, "Mixed Pricing in Oligopoly," Journal of Economic Theory 14, 373-388.

[25] Slade, M., 1998, "Optimal Pricing with Costly Adjustment: Evidence from Retail-Grocery Prices," Review of Economic Studies 65, 87-107.

[26] Sobel, J., 1984, "The Timing of Sales," Review of Economic Studies 51, 353-368.

[27] Varian, H., 1980, "A Model of Sales," American Economic Review 70, 651-659.