University of Toronto
Department of Economics

Working Paper 490

Customer Relationship and Sales

By Shouyong Shi

June 03, 2013

Customer Relationship and Sales

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

This version: May 2013

Abstract

I analyze a search equilibrium in a large market where customer relationship based on past trade arises endogenously together with service priority and sales. Specifically, there exists a unique equilibrium where it is optimal for a buyer to make repeat purchases from the related seller and optimal for a seller to give service priority to the related buyer. Customer relationship always improves welfare by reducing search frictions, but the equilibrium is socially efficient only when the buyer/seller ratio in the market is below a critical level. When the buyer/seller ratio exceeds this critical level, the equilibrium is inefficient because it fails to induce the coexistence of trading priority for related buyers and partial mixing of buyers for related sellers. Customer relationship induces price variations for individual sellers over time even when market conditions do not change. A seller posts a (high) regular price to sell to the related buyer and, once the seller loses the relationship, the seller posts a (low) sale price to sell to unrelated buyers until he gains a relationship. I also examine how market conditions affect the aggregate stock of relationships, markups, the size and the duration of a sale.

Keywords: Customer relationship; Sales; Directed search.

[*] Address: 150 St. George Street, Toronto, Ontario, Canada, M5S 3G7. A version of this paper has been presented at the Econometric Society Winter meeting (Chicago, 2012), the Tokyo Conference of the OEIO (Tokyo, 2012), Boston University (2012), and the NBER Summer Institute (2012). I thank the participants of those conferences and seminars for comments. I also thank Oleksiy Kryvtsov for insightful conversations on the topic. Financial support from the Canada Research Chair, the Bank of Canada Fellowship, and the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged. The opinion expressed here is my own and does not reflect the view of the Bank of Canada.

1. Introduction

Sellers often give trading priority to repeat buyers. For example, a restaurant may give the reservation priority to patrons, and an airline company may allow frequent flyers to select seats in advance and board earlier. Such customer relationship is often informal, and neither side of the relationship is abided by any agreement to continue the trade in the future. Although common in practice, customer relationship has not been formally modeled as an endogenous outcome of an equilibrium. In this paper, I demonstrate that customer relationship arises in the equilibrium of a large market with search frictions and implies a seller's optimal pricing strategy that involves sales. I analyze whether the equilibrium with relationships is socially efficient and examine how market conditions affect the stock of relationships, markups, the size and the duration of a sale.

A formal model of customer relationship in a large market is useful for the following reasons. First, repeat purchases are prevalent in retail transactions. Although the marketing literature has emphasized customer loyalty, it has taken customer loyalty as a primitive of the model (e.g., Blattberg and Sen, 1974). Endogenizing customer relationship as an equilibrium outcome in a large market helps understanding why it is prevalent. Second, a theory of customer relationship can help explaining some salient features of price dynamics at the micro level of retailers. In particular, prices exhibit large variability over time for the same item at a given store, sales account for a significant part of this variability, and prices excluding sales are sticky.[1] When competition for repeat purchases is important, it is natural that a seller may hold sales from time to time to attract buyers. Once a relationship is formed, a seller may keep price relatively stable while using non-price instruments, such as service priority, to maintain the relationship.

I study a market with a large number of buyers and sellers, where a seller in each period tries to 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, but individuals cannot coordinate

[1] See Klenow and Kryvtsov (2008) and Nakamura and Steinsson (2008).

their decisions. It is possible that a seller may fail to get a visitor and a buyer may fail to be selected by the seller he visits. Hence, there is a trade-off between the price of the good and the probability of trade. An important restriction is that a price posted by a seller must be available to all buyers. That is, a seller is not allowed to offer one price to a subset of buyers and another price to other buyers. This restriction is meant to capture the reality that a sale is commonly understood as a price cut available to all buyers.

Sellers and buyers may want to form relationships in order to increase the probability of trade. If a buyer bought from a seller in the previous period, the buyer is related to the seller, and the seller can choose whether to give trading priority to the related buyer over unrelated buyers. If the seller gives priority to the related buyer, a relationship is said to exist between the two individuals. A relationship breaks if the two fail to trade with each other in a period. Thus, there is nothing hard-wired about customer loyalty in this model; instead, the two individuals in a relationship can trade with anyone they want to.

I characterize a unique 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. The necessary and sufficient condition for this equilibrium with customer relationship to exist is that the number of buyers per seller in the market is below a critical level. Like any relationship, the relationship in this model generates benefits to the partners intertemporally and has implications on prices. A relationship enables a buyer to secure trade with the related seller in the future. Knowing this benefit of a relationship to the buyer, the related seller posts a high price that makes the related buyer just slightly prefer visiting him to visiting other sellers. An unrelated seller posts a low price in order to attract buyers and acquire a relationship. In the equilibrium, related buyers only visit their related sellers and unrelated buyers only visit sellers who have no related buyers.

Individual prices follow stylized dynamics in the equilibrium. A seller posts the high price as long as he has a related buyer and, after he loses the relationship, the seller posts the low price until he gains a relationship. The high price is the regular price and the low price is the sale price, not only because the high price is the one paid most often by buyers, but also because the low price is posted by an unrelated seller in the intention to revert to the high price later.[2] In each period, some unrelated sellers acquire the relationship as they succeed in trade, and some related sellers lose the relationship as they fail to trade due to their related buyers' absence from the market. These flows of individuals between the two types endogenously determine the stock of relationships in the market.

I compare the equilibrium with the social optimum that maximizes social welfare under the same search frictions as in the market. The constrained social optimum always allocates a related buyer to visit only the related seller and gives trading priority to a related buyer. Giving priority to related buyers is socially efficient because it reduces the extent of search frictions. The social optimum may also allocate unrelated buyers to visit both related and unrelated sellers. This partial mixing is socially efficient when the buyer/seller ratio exceeds a critical level, which is lower than the critical level below which the equilibrium exists. Thus, when the buyer/seller ratio is low, the equilibrium is socially efficient. When the buyer/seller ratio is intermediate, the equilibrium is inefficient because it does not feature partial mixing. When the buyer/seller ratio is high, the market outcome is inefficient because the equilibrium with priority ceases to exist.

The model has precise predictions on the frequency, the duration and the price discount of sales. In the steady state, the frequency of sales, the duration of a sale and the fraction of related sellers in the market depend only on the extensive margin of the market, which is represented by the buyer/seller ratio and the probability that a buyer is active in a period. These features of sales do not depend on the intensive margin of the market such as the value and the cost of the good. In contrast, prices, markups and the size of the discount depend on both margins. Moreover, prices and sales have the following features. First, the sale price implies a negative markup, i.e., a markdown when a relationship is sufficiently important. Second, goods with a higher profit margin have higher variability in prices over time even for the same seller, because they have larger markups in regular prices and larger markdowns in sales. Third, the regular price is relatively more stable than the sale

[2] Thus, it should not be surprising if a restaurant or airline company increases price or reduces the quality of service once it has obtained enough regular customers.

price. Fourth, the duration of a sale reflects changes in the buyer/seller ratio better than prices and markups, since the latter two can respond to such changes non-monotonically. Finally, an increase in demand arising from a higher probability that a buyer is active in the market can reduce the average price and markup.

In this model, search is directed in the sense that buyers know the terms of trade before choosing which seller to visit. Directed search is a reasonable assumption because buyers in many markets have information about the product and the price at particular stores and choose where to shop, rather than randomly search among all sellers. A growing literature has used directed search to model ex ante competition in a market with search frictions, dating back at least to Peters (1984, 1991) and Montgomery (1991). The particular approach in Peters (1991) is used here to derive each individual's trading probability endogenously as a function of strategies, which captures explicitly how a seller's decision on priority affects the selling probability.[3] The capacity constraint that underlies the trade-off between price and the trading probability is clearly presented in the two examples in the service industry mentioned earlier. For consumer products, the inventory cost also constrains the number of units available in a store and creates the possibility of stock-out.

This paper contributes to the literature of directed search as follows. First, this paper focuses on customer relationship which the literature has not studied. Second, customer relationship serves as an endogenous state variable in an individual's decision problem, which induces price dynamics and sales that are absent in the literature of directed search. Third, by allowing prices to depend on the endogenous relational type, the model generates a price differential, despite that all sellers are identical and all buyers are identical in exogenous characteristics. Moreover, the equilibrium generates a positive relationship between price and the trading probability across sellers. This may help reconciling the lack of strong evidence for the negative relationship between price and the seller's trading

[3]Other examples of this approach are Julien et al. (2000), Burdett et al. (2001), Albrecht et al. (2006), Galenianos and Kircher (2009), and Kircher (2009). For the alternative approach to directed search that assumes a matching function, see Moen (1997), Acemoglu and Shimer (1999), Shi (2009), Gonzalez and Shi (2010), Guerrieri et al. (2010), and Menzio and Shi (2011). Moreover, some models of directed search in the labor market have incorporated trading priority (e.g., Shi, 2002, 2006, and Shimer, 2005), but they have assumed that prices can be conditional on individuals' type.

probability with the prediction of a canonical model of directed search.

This paper is also related to the literature on sales, which has a few strands. The first strand relies on the result that when each seller faces a discontinuous demand curve, the equilibrium features a mixed strategy in posted prices (Shilony, 1977). Some authors interpret price reductions in the mixed strategy as sales, e.g., Salop and Stiglitz (1982) and Varian (1980).[4] The second strand is based on Sobel (1984). In this strand, a constant flow of buyers enter 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 reaches a critical level, some sellers cut prices to clear this stock. A sale lasts for one period in this model.[5] The third strand 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 the product (e.g., Bagwell, 1987).

I emphasize customer relationship, instead. As a mechanism to generate sales, customer relationship has its intuitive appeal and is complementary to those examined in the literature above. The theory helps explaining the regularity and the duration of sales. Also, it can generate a sale as a markdown, which is consistent with the data (e.g., Dutta et al., 2002) but missing in many of the models above. Finally, the literature on sales typically models a few sellers and/or buyers, and some of the models are difficult to be made dynamic. In contrast, the model here has an infinite horizon and many (in fact, infinitely many) sellers and buyers, which can be useful for analyzing large retail markets. To focus on the link between customer relationship 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.

In an analysis of customer relationship, Gourio and Rudanko (2011) argue that treating

[4]Note that the mixed strategy can be purified when there are a large number of sellers and buyers, as in Burdett and Judd (1983).

[5]Related to but different from Sobel's (1984) model, Lazear (1986) assumes that a seller does not know the value of the good to the buyers. After failing to sell the good, a seller updates his beliefs on buyers' valuation downward and reduces the price to clear the inventory of goods.

customer relationship as a type of intangible capital for a firm is useful for macro studies, but they assume that firms' pricing decisions depend exogenously on the customer base. Albrecht et al. (2011) incorporate directed search into Sobel's (1984) model to show that sales arise in a non-stationary equilibrium. An important ingredient in their model is that buyers are heterogeneous in tastes, as in Sobel's model. In contrast, in my model, all individuals on each side of the market are identical and sales are consistent with the stationary equilibrium. More importantly, I focus on customer relationship as a motivation for sales that is absent in Albrecht et al.

It is important to clarify that this paper is not about price discrimination or pricing strategies that are dependent on buyers' exogenous types. Instead, the emphasis is on how trading strategies can depend on bilateral trading histories and how this dependence generates price variations over time for the same seller even in the absence of the variations in demand and costs.[6] Because of search frictions, it is not feasible for a seller to offer priority to any arbitrary buyer before search to entice the buyer to visit, but the seller can give priority to the particular buyer related to him. This priority treatment arising from past trade is a natural notion of customer relationship. Also, it should be evident that if a seller picks up a buyer randomly to give priority *after* buyers visit, such priority is not able to direct search because a buyer does not know beforehand whether he will be the one to be chosen. To direct search, the priority must be either publicly announced or implicitly understood as the way to play the game. Finally, the mechanism for customer relationship in this paper differs from the locked-in effect arising from a fixed cost of switching sellers. Such a fixed cost does not exist here. In particular, if the buyer/seller is sufficiently high, the priority and the price differential will cease to exist in this model, but a fixed cost of switching will continue to generate the locked-in effect.

[6]If a seller is allowed to price discriminate, it is easy to generate a differential between the prices posted by a seller for related and unrelated buyers. For such models in the labor market with wage differentials, see Shi (2002, 2006) and Shimer (2005). Note that a price differential exists in these models even if the economy lasts for only one period. In contrast, a one-period version of the current model features no price differential because there is no future for a relationship to be paid off.

2. A Model of Directed Search with Customer Relationship

2.1. 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, for simplicity, fix $b \in (0, \infty)$. All individuals discount future at a rate $r > 0$. In each period, a seller can produce one indivisible unit of good at a cost $c \geq 0$. Because goods are perishable, a seller produces a good only after meeting a buyer. In any given period, a buyer has the need to consume a good with probability $\lambda \in (0, 1)$. These taste shocks occur at the beginning of the period, which are *iid* among the buyers and across time.[7] Call a buyer an active buyer if he has the need to consume, and an inactive buyer if he does not have the need. In each period, the number of active buyers is λN. The utility of consumption is U $(> c)$ to an active buyer. I refer to the number of active buyers per seller, $b\lambda$, as the *extensive margin* of the demand, as opposed to the *intensive margin* of the market which is determined by U and c.

Although all sellers are identical in the production capacity and cost, and all active buyers have the same taste, individuals may have different histories of trade. 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. For brevity, I refer to this history of trade as an individual's *type*. As such, an individual's type is endogenous. A seller's history is publicly observed. The focus is on the equilibria in which it is optimal for a related seller to give priority to his related buyer. That is, if a buyer visits the seller related to him, the seller chooses to trade with the buyer instead of an unrelated visitor. A related seller can offer a different price than an unrelated seller does.

In each period, sellers simultaneously post the terms of trade without knowing which buyer is active in the period. After observing the terms of trade, each active buyer chooses which seller to visit without knowing other buyers' choices. Then, a seller chooses one

[7] Alternatively, one can assume that a buyer exits the market with probability $1 - \lambda$ in each period and is replaced by a new buyer. The results are similar to the ones here if each seller must post a price without knowing the identities of the buyers who exited.

visitor to trade with. A seller must sell a good for the same price independently of the number and the type of buyers who will visit the seller. As explained in the introduction, this assumption is intended to capture the fact that a sale is available to all buyers and to make the results robust. However, a seller is allowed to give priority to the related buyer, as described above. If all visitors are unrelated to the seller, the seller randomly selects one to trade with.

A relationship between a buyer and a seller is informal. A buyer is free to shop at any seller, and a seller gives priority to the related buyer only when it is optimal to do so. Moreover, a relationship ends when the buyer fails to visit the seller, either because the buyer is inactive in the period or because the buyer chooses to visit another seller. This assumption keeps the analysis tractable by reducing an individual's history that is relevant for choices from potentially an infinite sequence to $\{0, 1\}$.[8]

The fraction of related sellers, denoted ρ, is an endogenous aggregate state variable. Because each related seller has one and only one related buyer, the number of related buyers in the market is equal to ρM, and the expected number of active related buyers is $\lambda\rho M$. The number of unrelated buyers is $N - \rho M$ and the expected number of active unrelated buyers 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.

2.2. A buyer's decision and payoff

Denote a buyer's (endogenous) type as $i \in \{0, 1\}$, where the buyer has a related seller if $i = 1$ and no related seller if $i = 0$. For a type 0 buyer, the type of the seller whom he can visit is $j \in \{0, 1\}$, where a seller $j = 1$ has a related buyer and a seller of type $j = 0$ does not. For a type 1 buyer, the type of the seller whom he can visit is $j \in \{0, 1, s\}$, where $j = s$ is the particular seller related to the buyer and $j = 1$ is a seller related to someone else.

[8] A generalization is to allow a related seller to keep the relationship with a positive probability if the related buyer is inactive and if no other buyer visits the seller in the period. This extension requires a buyer's taste shock to be public information. If the realization of a buyer's taste shock is private information, it may be rational for a seller to terminate a relationship when the related buyer does not show up. However, explicitly modeling this decision under private information is complicated.

When active, a type i buyer chooses the probability $\theta_{ij}(\rho, p) \in [0, 1]$ with which the buyer visits each seller of type j who posts price p. Let $v_{ij}(\rho, p)$ denote the value to an active type i buyer from such a visit. Let $V_i^a(\rho)$ denote the maximum value that an active type i buyer can obtain in the market, which is referred to as the *market value* for such a buyer. Clearly, $v_{ij}(\rho, p) \leq V_i^a(\rho)$ for all $ij \in \{00, 01, 10, 11, 1s\}$. If $v_{ij}(\rho, p) < V_i^a(\rho)$, it is optimal for a type i buyer not to visit a type j seller; i.e., $\theta_{ij}(\rho, p) = 0$ is optimal. If $v_{ij}(\rho, p) = V_i^a(\rho)$, then a type i buyer is indifferent between visiting a type j seller and obtaining the market value; i.e., the optimal choice is $\theta_{ij}(\rho, p) \in [0, 1]$. The optimal decisions in the two cases can be put in a compact form:

$$\theta_{ij}(\rho, p) \geq 0 \text{ and } v_{ij}(\rho, p) \leq V_i^a(\rho), \tag{2.1}$$

where the two inequalities hold with complementary slackness.

For $\theta_{ij}(\rho, p) > 0$ to be optimal, the combination (j, p) must attain the market value for a type i buyer. To express this requirement formally, let P_0 be the set of prices posted by all unrelated sellers and P_1 by all related sellers. The market value for an active buyer is

$$\begin{aligned} V_0^a(\rho) &= \max_{j \in \{0,1\}} [\max_{p \in P_j} v_{0j}(\rho, p)], \\ V_1^a(\rho) &= \max \left\{ v_{1s}(\rho, p_1),\ \max_{j \in \{0,1\}} [\max_{p \in P_j} v_{1j}(\rho, p)] \right\}, \end{aligned} \tag{2.2}$$

The first equation is for a type 0 buyer. The inner maximization selects among type j sellers the offer that gives the highest value to the buyer, and the outer maximization selects between the two groups of sellers. Similarly, the second equation is for a type 1 buyer. The outermost maximization is the buyer's choice between the particular seller related to him and all other sellers.

Since a large market is the subject of interest here, I will analyze the equilibrium in the limit where $M, N \to \infty$, with a positive and finite ratio $b = N/M$. In this limit, the number of active and related buyers, $\lambda \rho M$, and the number of active and unrelated buyers, $\lambda(N - \rho M)$, both become deterministic and approach infinity. Note that θ is the probability that a buyer visits an individual seller. In the limit, θ_{ij} approaches zero, except for $j = s$, and so it is not a convenient object to use to characterize the limit equilibrium.

For the limit, it is standard in the literature to use the *queue length* to describe the buyers' strategies, which is defined as the expected number of buyers of each type who visit a seller.[9] In particular, for a seller who posts price p in the aggregate state ρ, let $q_{ij}(\rho, p)$ denote the queue length of type i visitors who are unrelated to the seller, where $j \in \{0, 1\}$. If a seller has a related buyer, this particular buyer is counted separately from the queue length. Intuitively, q_{ij} is equal to the total number of active type i buyers multiplied by the probability that each of these buyers visits a type j seller. That is,

$$q_{1j}(\rho, p) = \lambda \rho M \theta_{1j}(\rho, p) \text{ and } q_{0j}(\rho, p) = \lambda(N - \rho M)\theta_{0j}(\rho, p), \quad j \in \{0, 1\}.$$

The queue length remains finite when $M, N \to \infty$, provided that θ does not exceed the order of magnitude $1/M$. A buyer's optimal decision in (2.1) can be expressed as

$$\theta_{1s}(\rho, p) \geq 0 \text{ and } v_{1s}(\rho, p) \leq V_1^a(\rho), \tag{2.3}$$

$$q_{i1}(\rho, p) \geq 0 \text{ and } v_{i1}(\rho, p) \leq V_i^a(\rho), \quad i \in \{0, 1\}, \tag{2.4}$$

$$q_{i0}(\rho, p) \geq 0 \text{ and } v_{i0}(\rho, p) \leq V_i^a(\rho), \quad i \in \{0, 1\}. \tag{2.5}$$

The two inequalities on each line hold with complementary slackness.

A seller treats all visitors who are unrelated to him equally. Following standard derivations, the following lemma describes the probability of trade for an individual and the constraints on the queue length (see Appendix A in Shi, 2011, for a proof):

Lemma 2.1. *A type* 0 *seller succeeds in trade with probability* $1 - e^{-(q_{10}+q_{00})}$ *and a visitor to a type* 0 *seller succeeds in trade with probability* $\frac{1-e^{-(q_{10}+q_{00})}}{q_{10}+q_{00}}$. *Suppose that a type* 1 *seller gives priority to his related buyer. Then, a type* 1 *seller succeeds in trade with his related buyer with probability* $\lambda\theta_{1s}$ *and with an unrelated visitor with probability*

[9] The characterization of a buyer's strategy by the queue length is justified by the focus on symmetric equilibria where all buyers of the same type respond to a seller's posting (including deviations from an equilibrium) in the same way (see Peters, 1991).

$(1-\lambda\theta_{1s})[1-e^{-(q_{11}+q_{01})}]$. *An unrelated visitor to a type* 1 *seller succeeds in trade with probability* $(1-\lambda\theta_{1s})\frac{1-e^{-(q_{11}+q_{01})}}{q_{11}+q_{01}}$. *Moreover, the queue length satisfies*

$$(1-\theta_{1s})\rho\lambda - [\rho q_{11} + (1-\rho)q_{10}] = 0, \tag{2.6}$$

$$(b-\rho)\lambda - [\rho q_{01} + (1-\rho)q_{00}] = 0. \tag{2.7}$$

To explain the trading probabilities, consider a type 0 seller first. The number of type i visitors to a type 0 seller is a random variable generated by the visiting probability, $\theta_{i0}(\rho, p)$, and the expected value of this number is q_{i0}. Since all buyers of the same type i use the same visiting strategy, the distribution of type i visitors over all type 0 sellers is generated by the urn-ball process, where type 0 sellers are the "urns" and type i buyers are the "balls". In the limit $M, N \to \infty$, a type 0 seller receives no visitor of type i with probability $e^{-q_{i0}}$. Thus, a type 0 seller succeeds in trade with probability $1 - e^{-(q_{10}+q_{00})}$. Since the expected number of visitors to a type 0 seller is $(q_{10} + q_{00})$, an arbitrary visitor is chosen by the seller to trade with probability $\frac{1-e^{-(q_{10}+q_{00})}}{q_{10}+q_{00}}$. The explanation for a type 1 seller is similar, except that a type 1 seller gives trading priority to his related buyer who will show up for his related seller with probability $\lambda\theta_{1s}$. To explain the constraints on the queue length, note that a type 1 buyer's visiting probabilities should add up to one across all the sellers, i.e., $\theta_{1s} + (\rho M - 1)\theta_{11} + (1-\rho)M\theta_{10} = 1$. Taking the limit $M, N \to \infty$ on this requirement yields (2.6). Similarly, (2.7) is the limit version of the adding-up constraint on a type 0 buyer's visiting probabilities across the sellers.

Let me calculate buyers' value functions. First, denote $V_i(\rho)$ as a type i buyer's value function at the *end of the previous period*.[10] Then,

$$V_i(\rho) = \frac{1}{1+r}\left[\lambda V_i^a(\rho) + (1-\lambda)V_0(\rho_{+1})\right], \quad i \in \{0, 1\}. \tag{2.8}$$

The subscript "+1" indicates 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

[10]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.

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 and obtains the value at the end of the current period, $V_0(\rho_{+1})$. Discounting the expected value of these two cases yields the value at the end of the previous period.

Second, I calculate v_{1j}, the value for a type 1 buyer when visiting a type j seller. If the buyer visits seller s, the buyer will be chosen by the seller with certainty. If p is the price posted by the related seller, then the buyer's value of visiting the related seller is

$$v_{1s}(\rho, p) = U - p + V_1(\rho_{+1}), \tag{2.9}$$

where $(U-p)$ is net utility from consumption in the period and $V_1(\rho_{+1})$ the continuation value $V_1(\rho_{+1})$. If the buyer visits a seller who is not related to the buyer, then the buyer's value is the same as the value for a buyer who visits the same seller but who is not related to any seller. That is, the following equalities hold for all p:

$$v_{11}(\rho, p) = v_{01}(\rho, p) \quad \text{and} \quad v_{10}(\rho, p) = v_{00}(\rho, p). \tag{2.10}$$

Third, I calculate $v_{0j}(\rho, p)$, the value for a type 0 buyer who visits a type j seller posting price p. For $j = 0$, i.e., a seller without a related buyer, all visitors are chosen by the seller with equal probability. If a buyer is chosen by such a type 0 seller, the trade yields the utility $U - p$ to the buyer in the period. In addition, the buyer becomes related to the seller, and the relationship changes the value for the buyer by $V_1(\rho_{+1}) - V_0(\rho_{+1})$. Denote the buyer's surplus of trade at price p as

$$D_b(\rho_{+1}, p) = U - p + V_1(\rho_{+1}) - V_0(\rho_{+1}).$$

With the trading probabilities in Lemma 2.1, I have

$$v_{00}(\rho, p) = \frac{1 - e^{-(q_{10}+q_{00})}}{q_{10} + q_{00}} D_b(\rho_{+1}, p) + V_0(\rho_{+1}), \tag{2.11}$$

$$v_{01}(\rho, p) = (1 - \lambda\theta_{1s}) \frac{1 - e^{-(q_{11}+q_{01})}}{q_{11} + q_{01}} D_b(\rho_{+1}, p) + V_0(\rho_{+1}), \tag{2.12}$$

where $q_{i0} = q_{i0}(\rho, p)$, $q_{i1} = q_{i1}(\rho, p)$ for $i \in \{0, 1\}$, and $\theta_{1s} = \theta_{1s}(\rho, p)$.

2.3. A seller's decision and payoff

For a seller of type $j \in \{0, 1\}$, let $J_j(\rho)$ be the value function that is measured at the *end of the previous period*. Consider first a type 0 seller who posts a price p_0. Because the seller does not have a related buyer, the seller selects all visitors with equal probability. The probability that the seller succeeds in trade is $1 - e^{-(q_{10}+q_{00})}$, as given in Lemma 2.1. A trade yields profit $p_0 - c$ to the seller in the current period and, relative to no trading, the seller's value at the end of the period changes by $[J_1(\rho_{+1}) - J_0(\rho_{+1})]$. Let $D_s(\rho_{+1}, p)$ denote a seller's surplus of trade at price p:

$$D_s(\rho_{+1}, p) = p - c + J_1(\rho_{+1}) - J_0(\rho_{+1}).$$

A type 0 seller's value function obeys:

$$\begin{aligned} J_0(\rho) \quad &= \tfrac{1}{1+r}\max_{p_0}\left\{\left[1 - e^{-(q_{10}+q_{00})}\right] D_s(\rho_{+1}, p_0) + J_0(\rho_{+1})\right\} \\ \text{s.t.} \quad &(2.5), \qquad \text{where } q_{i0} = q_{i0}(\rho, p_0),\ i \in \{0, 1\}. \end{aligned} \tag{2.13}$$

By incorporating (2.5) as a constraint, the seller explicitly takes into account the effect of the price choice on the queue length of visitors. Similarly, since a type 1 seller succeeds in trade with probability $1 - (1 - \lambda\theta_{1s})e^{-(q_{11}+q_{01})}$ (see Lemma 2.1), then

$$\begin{aligned} J_1(\rho) \quad &= \tfrac{1}{1+r}\max_{p_1}\left\{\left[1 - (1 - \lambda\theta_{1s})e^{-(q_{11}+q_{01})}\right] D_s(\rho_{+1}, p_1) + J_0(\rho_{+1})\right\} \\ \text{s.t.} \quad &(2.3) \text{ and } (2.4), \qquad \text{where } (\theta_{1s}, q_{i1}) = (\theta_{1s}, q_{i1})(\rho, p_1),\ i \in \{0, 1\}. \end{aligned} \tag{2.14}$$

In both (2.13) and (2.14), the seller takes as given the market value and future value functions for a buyer. Also, the seller takes as given his own future value functions, $J_0(\rho_{+1})$ and $J_1(\rho_{+1})$, as is standard in dynamic programming.

A type 1 seller can choose whether to give priority to the related buyer. If the seller does not give such priority, the seller will obtain the value J_0 as an unrelated seller does. It is optimal for a seller to give priority to the related buyer if and only if $J_1 \geq J_0$.

2.4. Equilibrium definition

With many sellers and buyers, 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 deviations) 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. As explained by Burdett et al. (2001), the symmetric equilibrium is a reasonable object to focus on in a large market without coordination. In the symmetric equilibrium, the set of prices posted by type j sellers is $P_j = \{p_j\}$, where $j \in \{0, 1\}$. 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 $i \in \{0, 1\}$.

In the equilibrium, the fraction of related sellers, ρ, is endogenously determined by the flows of sellers between the two types. In a period, a seller becomes unrelated if the seller fails to sell, which occurs with probability $(1-\lambda\theta_{1s})e^{-(q_{11}+q_{01})}$. An unrelated seller becomes related if the seller succeeds in trade, which occurs with probability $1 - e^{-(q_{10}+q_{00})}$. The fraction of related sellers changes between the current and the next period by

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

An *equilibrium with customer relationship* consists of buyers' choices θ_{ij} and 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) and value functions (J_0, J_1), and the fraction of related sellers ρ that satisfy:
(i) 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 (2.3) – (2.7), while (V_i^a, V_i, v_{ij}) satisfy (2.2) and (2.8) – (2.12);
(ii) p_0 solves (2.13), p_1 solves (2.14), and (J_0, J_1) satisfy (2.13), (2.14) and $J_1 \geq J_0$;
(iii) ρ satisfies (2.15).

It is important to repeat that the inequality $J_1 \geq J_0$ in (ii) requires it to be optimal for a seller to give priority to the related buyer. That is, the buyer-seller relationship is endogenous rather than being imposed.

3. Equilibrium Characterization

3.1. Narrowing down the set of equilibria

Because $\rho < 1$, it is easy to see that competition among sellers implies that a buyer enjoys a strictly positive surplus in an equilibrium. That is, $D_b(\rho_{+1}, p_j) > 0$ for $j \in \{0, 1\}$. The following lemma is proven in Appendix A:

Lemma 3.1. *An equilibrium with customer relationship satisfies: (i)* $\theta_{11} = 0$*,* $\theta_{1s} = 1$*, and* $\theta_{10} = 0$*; (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) says that a related buyer does not visit someone else's related seller. The reason is that, in comparison with the buyer's own related seller, someone else's related seller posts the same price and does not give priority to the buyer. With $\theta_{11} = 0$, the result $\theta_{1s} = 1$ implies $\theta_{10} = 0$. The result $\theta_{1s} = 1$ says that a related buyer chooses to visit only his related seller (say, seller A). This result is easy to understand if buyers unrelated to seller A choose not to visit seller A. In this case, whenever $\theta_{1s} < 1$, seller A can induce the related buyer to visit him with probability one by cutting the price slightly. Because this marginal change in the price has no consequence on other buyers' visiting probability to seller A, which remains at zero in this case, the discrete increase in the buyer's visiting probability increases the expected payoff to seller A. The result $\theta_{1s} = 1$ is also intuitive if buyers unrelated to seller A choose to visit seller A with positive probability. For an unrelated buyer to visit seller A, the buyer must obtain at least the same payoff from such a visit as visiting another seller. Since the buyer related to seller A has the priority relative to an unrelated buyer, the related buyer obtains strictly higher payoff from visiting seller A than visiting any other seller. Again, $\theta_{1s} = 1$ in this case. The result $\theta_{1s} = 1$ implies that, *in the equilibrium*, a relationship breaks up only when the related buyer receives a shock that makes him inactive in the period.[11]

The result (ii) in Lemma 3.1 says that an unrelated buyer visits an unrelated seller with positive probability. This result is easy to understand: Since an unrelated seller does not discriminate the visitors, he can always attract unrelated buyers with a sufficiently low price. The result (iii) says that if unrelated buyers visit related sellers with positive probability, then the equilibrium payoff to a related buyer must be strictly higher than

[11] The assumption $\lambda < 1$ is needed to prevent all sellers from being related in the steady state (see (2.15)). Endogenous breakups can be generated by allowing each buyer to contact two or more sellers in each period, in a way similar to labor-search models by Albrecht et al. (2006) and Galenianos and Kircher (2009). This extension is interesting but is left for future research. Also, see footnote 8.

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. The reason is that when a related seller's only potential buyer is the related buyer, a price increase does not crowd out the seller's potential visitors. In this case, it is optimal for the related seller to raise price until the related buyer just slightly prefers visiting him.

With Lemma 3.1, there are only two possibilities of an equilibrium with customer relationship. The two differ in an unrelated buyer's optimal choice:
Partial mixing: An unrelated buyer mixes between the two types of sellers, i.e., $q_{01} > 0$.
Complete separation: An unrelated buyer visits only unrelated sellers, i.e., $q_{01} = 0$.
In both cases, the equilibrium satisfies $q_{00} > 0$, $\theta_{1s} = 1$, $q_{10} = q_{11} = 0$, $V_1^a(\rho) = \max_p v_{1s}(\rho, p)$, 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. The following lemma is proven in Appendix A:

Lemma 3.2. *An equilibrium with partial mixing (and priority) does not exist.*

An equilibrium with partial mixing requires a related seller to attract both the related buyer and unrelated buyers. Because a seller's type (related or unrelated) is public information, an unrelated visitor to a seller who has a related buyer expects to be chosen to trade with a low priority. An unrelated buyer will visit a related seller only if the seller cuts price significantly below the price posted by an unrelated seller. The price cut defeats the seller's purpose of benefiting from the relationship. If forming a relationship is worthwhile for a seller, then setting price high to attract only the related buyer yields higher expected profit to a related seller than cutting price to attract both types of buyers. With the high price, the related seller squeezes all the rent from the relationship. Attracting unrelated buyers is optimal for a related seller only if there are so many buyers in the market that a seller's trading probability with an unrelated buyer is significantly higher than λ. In this case, however, giving priority to the related buyer is no longer optimal for the seller. By treating all buyers equally, instead, the seller can attract unrelated buyers more easily without cutting price below the one posted by unrelated sellers. Therefore, partial mixing and priority do not coexist in an equilibrium.

3.2. The equilibrium

I have shown that the only possible equilibrium with customer relationship is one with complete separation. To characterize the equilibrium, let me reformulate sellers' maximization problems. In a type 0 seller's problem, (2.13), the constraint (2.5) in the case $q_0 > 0$ requires the seller to give a type 0 visitor the same expected surplus as in the market. Expressing this constraint explicitly transforms a type 0 seller's problem as follows:

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

where D_s is the seller's surplus and D_b the buyer's surplus of trade defined earlier. Also as explained before, the seller takes as given the market value, $V_0^a(\rho)$, and future value functions, $(V_0, V_1, V_0^a, J_0, J_1)(\rho_{+1})$. 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 $\Delta = D_s + D_b$ is the joint surplus of the match given as

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

The first equation in (3.2) states that a type 0 buyer's gain from participating in the market, relative to staying inactive for a period, is equal to a fraction e^{-q_0} of the joint surplus of a match. The second equation in (3.2) states that a type 0 buyer's surplus from a trade, D_b, is equal to this gain from participating in the market multiplied by a factor $\frac{q_0}{1-e^{-q_0}}$. Together, the two conditions imply that the share of a type 0 buyer's surplus from a trade is equal to $\frac{q_0}{e^{q_0}-1}$. Equation (3.3) states that a type 0 seller's expected surplus from a trade is a share $[1-(1+q_0)e^{-q_0}]$ of the joint surplus of a match.

Note that the trading probability and the surplus share for an individual are both endogenous and only functions of the queue length. A buyer's trading probability when visiting an unrelated seller, $\frac{1-e^{-q_0}}{q_0}$, and the buyer's share of the match surplus, $\frac{q_0}{e^{q_0}-1}$, are decreasing functions of q_0. Intuitively, if buyers visit unrelated sellers with higher probability, there is higher congestion of buyers for such a seller. Each unrelated visitor succeeds in trade less likely in this case and the buyer who succeeds in trade obtains a smaller fraction of the match surplus. Moreover, because the matching probability for an unrelated visitor and the buyer's share of the match surplus are only functions of q_0, the two have a one-to-one positive relation with each other. Similarly, an unrelated seller's trading probability and surplus share are increasing functions of q_0.

With complete separation, a type 1 seller attracts only the related buyer (since $q_1 = 0$). Denote $\bar{p}$ as the price that makes the related buyer indifferent between visiting the related seller and visiting other sellers; i.e., $\bar{p}$ satisfies $v_{1s}(\rho, \bar{p}) = V_0^a(\rho)$. As explained after Lemma 3.2, the optimal price for attracting only the related buyer is $\bar{p} - \varepsilon$, where $\varepsilon > 0$ is arbitrarily small. For convenience, I will simply refer to $\bar{p}$ as a type 1 seller's optimal choice of the price. Solving the related buyer's indifference condition, I get:

$$\bar{p} = U + V_1(\rho_{+1}) - V_0^a(\rho). \tag{3.5}$$

Because $\bar{p}$ does not attract any type 0 buyer, a type 1 seller's value function satisfies:

$$(1+r)J_1(\rho) = J_0(\rho_{+1}) + \lambda \left[\bar{p} - c + J_1(\rho_{+1}) - J_0(\rho_{+1})\right]. \tag{3.6}$$

For complete separation to be an equilibrium, a type 1 seller should not gain from deviating to a price that attracts both the related buyer and unrelated buyers.[12] Such a deviation does not affect a related buyer's visiting strategy, who still gets the selection priority from the seller. Let $\tilde{p}_1$ be the price in the best deviation of this kind, $\tilde{q}_1$ the queue length of unrelated buyers attracted by the deviation, and $\tilde{J}_1(\rho)$ be the deviating seller's value function. By modifying (2.14), I can express the best choices $(\tilde{p}_1, \tilde{q}_1)$ as the solution

[12] Lemma 3.2 states that there cannot be an equilibrium in which all type 1 sellers attract both types of buyers. This does not automatically imply that a single type 1 seller's deviation to attract both types of buyers is not profitable.

to the following problem:

$$(1+r)\tilde{J}_1(\rho) = J_0(\rho_{+1}) + \max_{(p_1,q_1)} \left[1-(1-\lambda)e^{-q_1}\right] D_s(\rho_{+1}, p_1) \tag{3.7}$$

subject to

$$(1-\lambda)\frac{1-e^{-q_1}}{q_1} D_b(\rho_{+1}, p_1) = V_0^a(\rho) - V_0(\rho_{+1}). \tag{3.8}$$

As a form of (2.4) with $q_1 > 0$, (3.8) requires that if p_1 attracts unrelated buyers, then the expected surplus to such a buyer should be equal to the buyer's expected surplus in the market.[13] The deviation is not profitable if and only if $\tilde{J}_1(\rho) \leq J_1(\rho)$. Equivalently, this requires that the expected surplus generated by the deviation $(\tilde{p}_1, \tilde{q}_1)$ should not exceed that by $\bar{p}$, as expressed below:

$$\left[1-(1-\lambda)e^{-\tilde{q}_1}\right] D_s(\rho_{+1}, \tilde{p}_1) \leq \lambda D_s(\rho_{+1}, \bar{p}). \tag{3.9}$$

Moreover, an equilibrium with complete separation must satisfy $J_1(\rho) \geq J_0(\rho)$ in order for a type 1 seller to give priority to the related buyer.

In addition to the above conditions on prices and value functions, an equilibrium requires q_0 and ρ to be determined and satisfy $q_0 > 0$ and $\rho \in (0, \min\{1, b\})$. With complete separation, the adding-up constraint (2.7) yields:

$$q_0 = H(\rho) \equiv \lambda(b-\rho)/(1-\rho). \tag{3.10}$$

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

$$\rho_{+1} = G(\rho) \equiv \lambda\rho + (1-\rho)\left[1-e^{-H(\rho)}\right]. \tag{3.11}$$

Note that (3.10) and (3.11) are independent of variables other 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.

[13] The deviation is a one-period deviation. As such, it does not change the deviator's future value functions that appear in the suprlus $D_s(\rho_{+1}, p)$, which are still $J_0(\rho_{+1})$ and $J_1(\rho_{+1})$ rather than the ones with tilde. As an implication of dynamic programming, a deviation is not profitable if and only if no one-period deviation is profitable.

To sum up, an equilibrium with customer relationship consists of prices $(p_0, \bar{p})$, queue length q_0, value functions $(J_0(\rho), J_1(\rho))$ and the fraction of related sellers ρ that satisfy (3.2), (3.3), (3.5), (3.6), (3.9), $J_1(\rho) \geq J_0(\rho)$, (3.10) and (3.11). In this equilibrium, all active buyers have the same market value, $V_1^a(\rho) = V_0^a(\rho)$. Adding the superscript $*$ to indicate the steady state, I prove the following proposition in Appendix B:

Proposition 3.3. *(i) There exists $B(\lambda)$, defined by (B.10) in Appendix B, such that the steady state exists if and only if $b \leq B(\lambda)$. Moreover, $B(\lambda) > \lambda - \frac{1-\lambda}{\lambda}\ln(1-\lambda) > 1$ for all $\lambda \in (0,1)$, and $B(1) = 1$.*
(ii) The steady state is unique and locally stable. Equilibrium ρ and q_0 depend only on (b, λ). Moreover, $0 < \rho^* < \min\{1, b\lambda\}$, $q_0^* > 0$, $\frac{d\rho^*}{db} > 0$, $\frac{dq_0^*}{db} > 0$, and $\frac{d\rho^*}{d\lambda} > 0$.
(iii) For ρ close to ρ^*, the equilibrium satisfies $\Delta(\rho) > 0$ and

$$p_0 = U - \frac{q_0 e^{-q_0}}{1 - e^{-q_0}}\Delta(\rho_{+1}), \tag{3.12}$$

$$p_1 = \bar{p} = U - e^{-q_0}\Delta(\rho_{+1}) > \max\{p_0, c\}, \tag{3.13}$$

$$V_1(\rho) = V_0(\rho) = \frac{1}{1+r}\left[V_0(\rho_{+1}) + \lambda e^{-q_0}\Delta(\rho_{+1})\right], \tag{3.14}$$

$$\Delta(\rho) = U - c + \frac{1}{1+r}\left[\lambda(1 - e^{-q_0}) - 1 + (1 + q_0)e^{-q_0}\right]\Delta(\rho_{+1}), \tag{3.15}$$

where $q_0 = H(\rho)$ and $\rho_{+1} = G(\rho)$ are given by (3.10) and (3.11).

The steady state with customer relationship exists if and only if the buyer/seller ratio b does not exceed the critical level $B(\lambda)$. This existence condition comes from the requirement that it should be optimal for a type 1 seller to give priority to the related buyer, i.e., $J_1 \geq J_0$. The condition is intuitive. To a type 1 seller, the benefits of giving priority to his related buyer are that the seller can guarantee a trade whenever the related buyer is active and that the seller can charge a relatively high price. The opportunity cost is that the priority for the related buyer drives away unrelated buyers. When the number of buyers per seller is small, unrelated buyers are difficult to attract regardless of whether a seller gives priority to the related buyer. In this case, the benefits of giving priority to the related buyer outweigh the opportunity cost for a type 1 seller. On the other hand, if the number of buyers per seller is high, giving priority to the related buyer can reduce a type 1 seller's trading probability significantly as it eliminates a large number of unrelated buyers' visit to the seller. In this case, it is optimal for a type 1 seller to give no priority to his related buyer and treat all buyers equally.

Specifically, when a type 1 seller chooses whether to give priority to the related buyer, the seller compares the expected share of the match surplus from the two options. If the seller gives priority, the seller's trading probability is λ. In such a trade, the seller sets price as such that the related buyer's surplus from the trade, $U - \bar{p}$, is the same as the expected surplus that the buyer can get from visiting another (unrelated) seller, which is $e^{-q_0}\Delta$. Thus, by giving priority, the seller's share of the match surplus is $(1 - e^{-q_0})$ and the seller's expected share is $\lambda(1 - e^{-q_0})$. If the seller does not give priority, the seller's trading probability is $(1 - e^{-q_0})$, the share of the match surplus is $(1 - \frac{q_0}{e^{q_0}-1})$, and the expected share is $(1 - e^{-q_0})(1 - \frac{q_0}{e^{q_0}-1})$. By giving priority to the related buyer, the seller gets a higher expected share of the surplus if and only if $\lambda > 1 - \frac{q_0}{e^{q_0}-1}$. This condition puts an upper bound on the queue length of buyers for an unrelated seller and, hence, on the number of buyers per seller.

The condition $b \leq B(\lambda)$ is also sufficient for (3.9) to be met. Namely, if it is optimal for type 1 sellers to give priority to related buyers, then it is not profitable for a type 1 seller to deviate to attract both types of buyers. To explain why this is the case, note that when the number of buyers per seller is low, the price posted by a type 0 seller is likely to be low. If a type 1 seller deviates to attract unrelated buyers, competing against type 0 sellers for unrelated buyers entails a large price cut from $\bar{p}$. This loss in profit outweighs the benefit of the small increase in the trading probability induced by the deviation.

Because $B(\lambda) \geq 1$ for all λ, a sufficient condition for the equilibrium with customer relationship to exist is that the number of buyers does not exceed the number of sellers. Moreover, since $B(\lambda)$ depends only on λ, whether the equilibrium exists is independent of the intensive margin of the market (c, U) and the discount rate r. This independence arises from the fact that (c, U, r) affect the size of the match surplus but not an individual's trading probability or the share of the match surplus.

Similarly, the fraction of related sellers, ρ, and the queue length, q_0, depend only on the extensive margin of the demand, (b, λ), and not on other parameters such as (c, U, r). Given 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 q_0, each seller's trading probability is uniquely determined, which determines the flows of sellers into and out of relationships and, hence, the dynamics of ρ. These two relations are given above as $q_0 = H(\rho)$ and $\rho_{+1} = G(\rho)$, which uniquely determine ρ and q_0. Because the only parameters in these relations are (b, λ), the solutions for (ρ, q_0) to these relations depend only on (b, λ) and not on other parameters such as (c, U, r).

The fraction of related sellers in the steady state increases in the extensive margin of demand. A higher b or λ leads to more trades in each period, which turn more unrelated sellers into related ones and so increase the fraction of related sellers 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^*. However, it is ambiguous whether a higher λ increases q_0^*. On the one hand, a higher λ increases the general demand by increasing the number of active buyers. This effect increases q_0^*. On the other hand, a higher λ tilts demand toward related sellers by making a larger fraction of buyers related to some sellers. This effect reduces the queue length of buyers for an unrelated seller. The overall effect of λ on q_0^* depends on the values of λ and b.

The unique steady state is locally stable; namely, there is a unique saddle path along which the fraction of related sellers converges to the steady state level ρ^*. If $\rho > \rho^*$, the flow of related sellers who lose relationships in a period is greater than the flow of unrelated sellers who acquire relationships through trade. In this case, there is a positive net flow of sellers out of relationships, and so 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 than the flow of unrelated sellers who acquire relationships through trade. In this case, the fraction of related sellers increases toward the steady state.

4. Efficiency Properties of the Equilibrium

Because the equilibrium features priority for related buyers and complete separation between related and unrelated individuals, it is important to examine whether the equilibrium is socially efficient. For this purpose, I characterize the allocation chosen by a social planner to maximize social welfare under the same matching frictions as in the market.[14] The planner knows each individual's trading history in the previous period that determines whether the individual is related or unrelated. For a related buyer, the planner chooses the probability for the buyer to visit the related seller, θ_{1s}, and queue lengths (q_{11}, q_{10}). For an unrelated buyer, the planner chooses queue lengths (q_{01}, q_{00}). Because of coordination frictions, the planner is constrained to treat all individuals of the same type equally, as in the market. Thus, when the planner allocates a buyer to visit a seller who is unrelated to the buyer, the allocation cannot depend on the identity of the buyer or the seller. This constraint is implicit in the notation of queue lengths that can depend only on the buyer's and seller's type. Similarly, the probability that a visitor to a seller is chosen to trade can depend only on whether or not the visitor is related to the seller, but not on the visitor's identity. For the moment, I assume that the planner gives priority to the related buyer when the buyer visits his related seller. Later I will verify that such priority is indeed socially efficient. The matching probabilities are given in Lemma 2.1.

Because all individuals are risk-neutral, social welfare can be measured by the discounted sum of surpluses in all matches divided by the number of sellers. Let $W(\rho)$ denote this social welfare function, measured at the end of the previous period. To calculate W, recall that a seller fails to trade in a period with probability $(1-\lambda\theta_{1s})e^{-(q_{11}+q_{01})}$ if the seller is related and with probability $e^{-(q_{10}+q_{00})}$ if the seller is unrelated. The average probability

[14] All proofs for this section are relegated to Appendix C.

of trade per seller in a period is:

$$\psi = 1 - \rho(1 - \lambda\theta_{1s})e^{-(q_{11}+q_{01})} - (1-\rho)e^{-(q_{10}+q_{00})}.$$

The joint surplus of a match in a period is $U - c$. Thus, the social welfare function and the planner's optimal choices solve the following problem:

$$(1+r)W(\rho) = \max_{(\theta_{1s},q_{11},q_{10},q_{01},q_{00})} \left[\psi(U-c) + W(\rho_{+1})\right], \tag{4.1}$$

subject to (2.6), (2.7), (2.15), $\theta_{1s} \in [0,1]$, and $q_{ij} \geq 0$ for $ij \in \{11, 10, 01, 00\}$. These constraints have the same meanings as in the equilibrium.

Consider the following candidate for the social optimum:

$$\left.\begin{array}{l} \theta_{1s} = 1, \quad q_{11} = q_{10} = 0; \\ q_{00} > 0; \\ q_{01} \geq 0 \text{ and } q_{01} \geq q_{00} + \ln(1-\lambda), \text{ with complementary slackness.} \end{array}\right\} \tag{4.2}$$

This candidate allocates a related buyer to visit only the buyer's related seller and an unrelated buyer to visit unrelated sellers with positive probability, as in the equilibrium. In contrast to the equilibrium, the candidate may or may not choose to partially mix related and unrelated individuals. Precisely, the candidate allocates an unrelated buyer to visit related sellers with positive probability if and only if $q_{00} > -\ln(1-\lambda)$. Let me delay the explanation why this condition is needed for efficiency and, instead, turn to the comparison of efficiency between the allocations with and without priority.

To simplify the comparison, I focus on the steady state. The results also hold in a neighborhood of the steady state. Comparing steady-state welfare amounts to comparing the average trading probability per seller, ψ. In any allocation without priority, all sellers are the same, all buyers are the same, the queue length of buyers per seller is $b\lambda$, and the average trading probability per seller is $1 - e^{-b\lambda}$. The following lemma shows that giving priority to the related buyer is socially efficient, in addition to other results:

Lemma 4.1. *Focus on the steady state. The average trading probability per seller is* $\psi = \rho$. *The allocation in (4.2) with trading priority for related buyers yields higher welfare* *than the allocation without priority. Moreover, the allocation in (4.2) induces* $q_{01} > 0$ *if and only* $b > B_e(\lambda)$ *where*

$$B_e(\lambda) \equiv \lambda - \frac{1-\lambda}{\lambda}\ln(1-\lambda). \tag{4.3}$$

$B_e(\lambda)$ *satisfies* $1 < B_e(\lambda) < -\frac{1}{\lambda}\ln(1-\lambda)$ *for all* $\lambda \in (0,1)$*, and* $B_e(0) = B_e(1) = 1$.

The result $\psi = \rho$ in the steady state is straightforward to explain. In any period, the fraction of sellers who succeed in trade is ψ. Because related sellers who fail to trade lose their relationship, the fraction of related sellers in the next period will be $\rho_{+1} = \psi$. In the steady state, the fraction of related sellers is constant over time, and so $\rho = \psi$.

To explain why giving priority to related buyers is socially efficient, let me first explain the condition for partial mixing in (4.2), i.e., $q_{01} = q_{00} + \ln(1-\lambda)$ if $1 - \lambda > e^{-q_{00}}$, and $q_{01} = 0$ if $1 - \lambda \leq e^{-q_{00}}$. This efficiency condition arises from maximizing the marginal contribution of an unrelated buyer to the number of trades and, hence, to social welfare. To see why, note that allocating a buyer to match with a seller increases the number of trades if and only if the seller is not already matched. Thus, the marginal contribution of an unrelated buyer to the number of trades is $(1-\lambda)e^{-q_{01}}$ when the buyer is assigned to match with a related seller and $e^{-q_{00}}$ when the buyer is assigned to match with an unrelated seller. If $1 - \lambda > e^{-q_{00}}$, social welfare is maximized by equalizing the marginal contribution of an unrelated buyer across sellers, i.e., by setting $q_{01} = q_{00} + \ln(1-\lambda)$. If $1 - \lambda < e^{-q_{00}}$, such equalization is not possible, because the marginal contribution of an unrelated buyer is higher when the buyer is assigned to match with unrelated sellers than with related sellers, even if only one unrelated buyer is assigned to match with related sellers. In this case, all unrelated buyers should be assigned to match with only unrelated sellers.

It is socially efficient to give priority to related buyers because doing so reduces the extent of matching frictions. The number of trades would be maximized if the planner could allocate all individuals by their identities and assign each individual buyer to match with a particular seller. This would ensure the shorter side of the market to be always matched. However, this allocation is not feasible because matching frictions prevent the planner from allocating unrelated individuals by their identities. Nevertheless, it is feasible to assign a related buyer to the particular seller related to him. For this reason, giving priority to the related buyer in the match with the related seller reduces the extent of matching frictions and increases the number of trades. This efficiency gain from priority is evident when $\lambda \geq 1 - e^{-q_{00}}$. In this case, because a related buyer is active with high probability, the priority gives a related seller a higher trading probability than an unrelated seller. At the same time, related sellers do not create congestion for unrelated sellers by competing for unrelated buyers. When $\lambda < 1 - e^{-q_{00}}$, partially mixing can increase efficiency, but the probability that an unrelated buyer is assigned to a related seller is still lower than that to an unrelated seller. Thus, related sellers do not crowd out unrelated sellers by as much as in the allocation without priority. Again, the number of trades and, hence, welfare is higher with priority than without.

It is intuitive that partial mixing is efficient if and only if the number of buyers per seller is sufficiently high, i.e., if $b > B_e(\lambda)$. When buyers are abundant, there is high congestion of buyers for each unrelated seller. Since a related seller has positive probability of failing to trade with the related buyer, as $\lambda < 1$, allocating an unrelated buyer to match with related sellers with positive probability increases the total number of trades in this case. Since $B_e(\lambda) > 1$ for all $0 < \lambda < 1$, a necessary condition for partial mixing to be socially efficient is that the number of buyers exceeds the number of sellers.

The following proposition confirms the allocation in (4.2) as the social optimum and compares the equilibrium with the social optimum:

Proposition 4.2. *In a neighborhood of the steady state, the social optimum is the one with customer relationship as described by (4.2). Moreover, $B(\lambda) > B_e(\lambda)$. Thus, the equilibrium is socially efficient when $b \leq B_e(\lambda)$ and socially inefficient when $b > B_e(\lambda)$. The inefficiency of the equilibrium arises from the absence of mixing when $B_e(\lambda) < b \leq B(\lambda)$, and from the absence of priority when $b > B(\lambda)$.*

It is not surprising that the allocation in (4.2) is socially efficient, given the above explanation on how the allocation minimizes the difference in an unrelated buyer's social marginal contribution across sellers. It is also not surprising that the equilibrium is socially efficient when $b \leq B_e(\lambda)$: In this case, both the equilibrium and the social optimum call for complete separation between related and unrelated individuals.

When $b > B_e(\lambda)$, the equilibrium is inefficient because it fails to induce the coexistence of priority and partial mixing. I have explained this failure for Lemma 3.2. In the case $B_e(\lambda) < b \leq B(\lambda)$, the equilibrium has priority but no mixing, because a related seller would rather keep price at the high level $\bar{p}$ to attract only the related buyer. In the case $b > B(\lambda)$, the equilibrium with customer relationship ceases to exist. Because there are a large number of buyers per seller in this case, a seller's trading probability with an unrelated buyer can be higher than with the related buyer. By eliminating priority, a related seller can attract a buyer more easily without a large price cut.

The source of the inefficiency is the restriction that a seller must offer one price to all buyers. With this restriction, a related seller ends up choosing either to give priority to attract only the related buyer or to attract all buyers without priority. If a related seller can post prices conditional on the buyer's type, then the seller will be able to use a relatively low price to attract unrelated buyers and offer a high price and priority to the related buyer. It can be shown that such conditional pricing can restore efficiency. This efficiency role of conditional pricing resembles the one established by Shi (2002, 2006) and Shimer (2005) in related models of the labor market with directed search.

5. Relationship, Prices and Sales

I now analyze prices and sales in the equilibrium and conduct comparative statics.

5.1. Dynamic patterns of pricing and matching

In each period, a related seller posts the high price $p_1 = \bar{p}$ and an unrelated seller posts the low price $p_0 < \bar{p}$. The high price p_1 is paid by buyers not only in most trades in any given period but also most often over time to any given seller. Thus, I call p_1 the *regular price* and p_0 the *sale price*. The price posted by a seller varies over time between the two prices even though there is no change in market conditions. A seller posts the regular price as long as he has a related buyer and, once the seller loses the relationship, he holds a sale at price p_0 until he gains a relationship. The fact that an unrelated seller cuts the price to p_0 with an intention to revert to a higher price p_1 in the future accords with the common sense that p_0 is a sale price. Outside the steady state, the two prices are not constant; instead, they are functions of the fraction of related sellers in the market. As this aggregate state approaches the steady state, so do the two prices.

The consideration for holding a sale is intertemporal. By holding a sale, a seller intends to attract buyers and form a relationship, but the relationship will pay off only in the future through a higher (regular) price. Moreover, it is possible that the sale price is lower than the marginal cost of the good (see Corollary 5.1 below). Specifically, when the rent to the seller from a relationship, measured by $J_1 - J_0$, is large enough, an unrelated seller may find it optimal to incur the temporary loss of a markdown in order to form the relationship. In contrast, the regular price p_1 is always strictly higher than the marginal cost.

The intertemproal consideration also implies that price and the selling probability are positively correlated across sellers when λ is high. In particular, when λ is close to one, a seller who posts the high price p_1 in the equilibrium has a selling probability close to one while a seller who posts the low price p_0 has a non-negligible probability of failing to sell. At first glance, this positive relationship between price and the selling probability seems to contradict the fundamental trade-off between the two variables in a directed-search framework. But such a contradiction does exist in the model. In fact, if a type 0 seller contemplates posting a price higher than p_0, the queue length of visitors q_0 will indeed be lower, and so will the seller's probability to sell. The positive relationship between price and the selling probability in the equilibrium is fully consistent with the trade-off in directed search, because the relationship arises across different types of sellers. A related seller had a related buyer but an unrelated seller does not. By offering priority to explore the relationship, a related seller is able to both charge a higher price and sell the good more quickly. Thus, controling only for exogenous characteristics of sellers is not enough for testing whether sellers make the trade-off between price and the selling probability. One must also control for endogenous heterogeneity among sellers, such as the seller's trading history that determines the seller's type in this model.[15]

5.2. Frequency, duration and size of sales

The frequency of sales is equal to the fraction of trades that occur at the sale price. Because the number of trades at the sale price in each period is $M(1-\rho)(1-e^{-q_0})$ and the total number of trades in a period is $M\rho$ (see Lemma 4.1), the *frequency of sales* is $1-\lambda$. Most of the trades occur at the regular price if λ is close to one. The duration of a sale is equal to the expected length of time that an unrelated seller takes to succeed in trade. Because an unrelated seller's trading probability is $1-e^{-q_0}$, the *duration of a sale* is $1/(1-e^{-q_0})$. The markup implied by each price is

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

Weighted by the frequency of trades, the average markup is $(1-\lambda) \times markup0 + \lambda \times markup1$. The *size of the price discount* of a sale is $1 - \frac{p_0}{p_1}$.

The frequency of sales in the steady state depends only on how frequently the seller loses a relationship, which is $1-\lambda$. The duration of a sale depends only on (b, λ) and not on (c, U, r), because the queue length q_0 has this property (see Proposition 3.3). In contrast, markups and the size of the price discount of a sale depend on both the extensive and the intensive margin of the market. The following corollary describes additional features of comparative statics (see Appendix D for a proof):[16]

Corollary 5.1. *Let $r \to 0$. The steady state has the following features:*
(i) $markup1 > 0$. Also, $markup1$ increases in U/c and b.
(ii) If $\lambda \leq 1/2$, then $markup0 > 0$ and it increases in U/c and b. If $\lambda > 1/2$, then (a) there exists $b_2 < B(\lambda)$ such that $markup0 < 0$ if and only if $b \in (0, b_2)$; (b) $markup0$ decreases in U/c; and (c) there exists $b_1 < b_2$ such that $markup0$ decreases in b if and only if $b \in (0, b_1)$.
(iii) The size of the discount of a sale increases in U/c. If $\lambda > 1/2$, the size of the discount

[15] A buyer's trading probability is positively related to the price, which is standard in directed search.
[16] A change in a parameter may also induce dynamics in the stock of relationships, ρ. Such dynamics are examined in Shi (2011) but omitted here to economize on space.

increases in b when b is small and decreases in b when b is close to $B(\lambda)$.
(iv) The duration of a sale decreases in b. If $b < 1$, then there exists $\lambda_1 > 0$ such that the duration of a sale decreases in λ if and only if $\lambda \in (0, \lambda_1)$.

The behavior of the regular price is as expected. The positive markup implied by the regular price, $markup1$, reflects the rent that a related seller gets from a relationship. Setting the regular price below the marginal cost would defeat the purpose of forming a relationship. When the value of the good increases relative to the cost, $markup1$ increases because the profit margin increases. Also, as the number of buyers per seller increases, $markup1$ increases to capture the benefit to a seller of the higher demand.

The sale price behaves similarly to the regular price only when $\lambda \leq 1/2$. When $\lambda > 1/2$, however, the sale price behaves quite differently. First, when the number of buyers per seller is low, the sale price implies a negative markup, that is, a markdown on the marginal cost. Second, when the value of the good increases relative to the cost, $markup0$ increases if and only if it is positive. Putting the result differently, if the sale price implies a markdown, then the markdown is larger when the value of the good relative to the cost is larger. Third, $markup0$ exhibits a U-shaped response to the buyer/seller ratio. When the buyer/seller ratio is sufficiently small, the markdown implied by the sale price increases when the buyer/seller ratio increases. After the buyer/seller ratio increases above the critical level b_1, the markdown starts to decrease; i.e., the markup starts to increase but is still negative. When the buyer/seller ratio increases further to pass the critical level b_2, the markup implied by the sale price becomes positive and increases with b.

To explain why the sale price behaves so differently, note first that when $\lambda > 1/2$, a relationship generates a trade for the related seller with high probability. This motivates an unrelated seller to cut price deeply in order to form a relationship. If the number of buyers per seller is also low, the fierce competition among sellers for buyers induces an unrelated seller to cut price below the marginal cost. As the value of the good relative to the cost increases, the surplus of trade increases, and so the expected return to a relationship increases. This large expected gain in the future motivates an unrelated seller to offer a larger markdown to attract buyers to form a relationship. Similarly, when the number of buyers per seller increases from sufficiently low levels, an unrelated seller increases the markdown to attract the increasing number of buyers. When the number of buyers per sellers is sufficiently large, however, buyers are readily available and there is no need for an unrelated seller to offer a markdown. In this case, the markup implied by the sale price is positive and increases with the buyer/seller ratio.

The size of the price discount of a sale reflects these differences between the sale price and the regular price. When the value of the good relative to the cost increases, the size of the discount increases because an unrelated seller reduces the sale price to attract buyers and a related seller increases the regular price to explore the larger profit margin. When the number of buyers per seller is sufficiently small, an increase in this number increases both the markdown and the markup, thus widening the gap between the regular and the sale price. When the number of buyers per seller is sufficiently high, the incentive to offer a price cut to form a relationship is weak, in which case an increase in the number of buyers per seller reduces the price discount in a sale.

In contrast to the non-monotonic responses to b by the markdown and the size of the discount, the duration of the discount monotonically decreases in b. This is because an increase in the buyer/seller ratio increases the queue length of buyers for each unrelated seller. Although an increase in demand through b may induce unrelated sellers to offer deeper discounts, a seller finds a trade more quickly and so a sale ends more quickly.

The above responses of the equilibrium to market conditions have several implications. First, if customer relationship is important in the retail market, then markups change with the marginal cost, in contrast to the constant markup in macro models based on monopolistic competition. Second, goods with a higher profit margin have a higher variability in prices over time even for the same seller, because the gap between the regular price and the sale price is larger for such goods. Third, the regular price is relatively more stable than the sale price. Because a related seller's trading probability is higher than an unrelated seller's, 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 increased cost, only a smaller change in the regular price is needed than in the sale price. Fourth, the duration of a sale is a more reliable indicator of a change in the buyer/seller ratio than prices and markups, since it responds to such a change monotonically in the intuitive direction.

I now discuss the response of the equilibrium to changes in λ. As explained for (ii) of Proposition 3.3, an increase in λ both increases the extensive margin of demand for goods and tilts demand toward related sellers. As a result, a higher λ does not necessarily increase the queue length of buyers for an unrelated seller, in contrast to the effect of a higher b. Specifically, if $b < 1$ and if λ is sufficiently small, an increase in λ has a stronger effect of tilting demand than increasing the general demand. In this case, a higher λ reduces the queue length of buyers for an unrelated seller and increases the duration of a sale, as stated in (iv) of the above corollary.

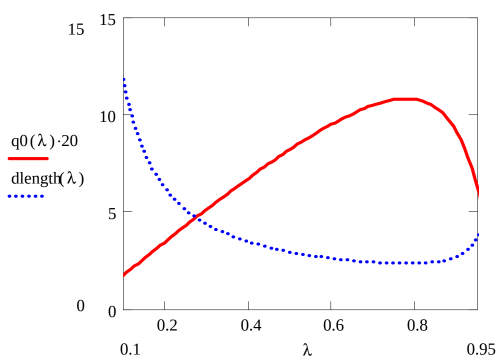

Figure 1.1
Queue length and duration of a sale

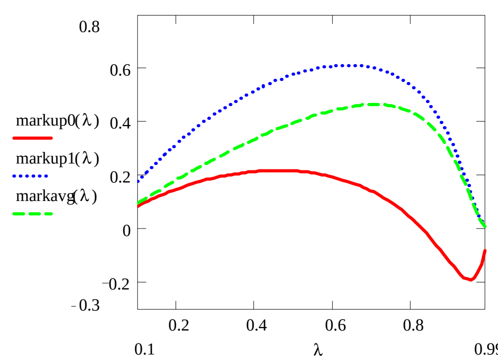

Figure 1.2
Markups

A related difference between λ and b is the effect on the average price and the average markup. Although an increase in b can induce non-monotonic responses in the sale price, the regular price always increases in b. Since most trades occur at the regular price, the average price and the average markup increase in b. In contrast, an increase in λ can lead to non-monotonic responses in the regular price, the average price and the average markup. As an illustration, consider the following parameter values: $r = 7.545 \times 10^{-4}$, $U = 1$, $c = 0.311$, $\lambda = 0.956$, $b = 0.895$.[17] Figures 1.1 and 1.2 illustrate the response

[17]See Shi (2011) for the calibration that led to these parameter values.

of the steady state to changes in λ from 0.1 to 0.95, while other parameters are fixed at the values above. The queue length of buyers for an unrelated seller has a hump-shaped dependence on λ, which results in a U-shaped dependence on λ of the duration of a sale depicted in Figure 1.1 as *dlength* (in weeks). Figure 1.2 shows that the two markups depend on λ non-monotonically. In addition, the average markup (depicted as *markavg*) is hump-shaped.

The hump-shaped dependence of the average markup on λ reflects the fact that the return of a relationship to a seller is larger at intermediate values of λ than at both high and low values of λ. This return is low at low values of λ because the related buyer is unlikely to be active in the future. This return is also low at high values of λ because, when almost all sellers have a related buyer, competition among them drives down the regular price. In contrast, at intermediate values of λ, there is a sizable benefit of giving priority to the related buyer and charging a high regular price. It is remarkable that prices and markups can fall when the demand increases through λ. This result offers an explanation for the finding by Chevalier et al. (2003) that prices of particular items in a supermarket chain typically fall when the items experience peak seasonal demand. Specifically, the particular items are likely demanded by a store's regular customers. An increase in the demand for such items increases competition for the regular customers and reduces prices.

6. Conclusion

Modeling customer relationship in a model of directed search, I prove that there exists a unique equilibrium in which a seller gives priority to the related buyer and a buyer makes repeat purchases from the related seller. Customer relationship always improves welfare, but the equilibrium is socially efficient only when the buyer/seller ratio in the market is below a critical level. When the buyer/seller ratio exceeds this critical level, the equilibrium is inefficient because it fails to induce the coexistence of trading priority for related buyers and partial mixing of buyers for related sellers. Customer relationship induces price variations for individual sellers over time even when market conditions do not change. A seller posts a (high) regular price to sell to the related buyer and, once the seller loses the relationship, the seller posts a (low) sale price to sell to unrelated buyers until he gains a relationship. I also examine how market conditions affect the aggregate stock of relationships, markups, the size and the duration of a sale.

Endogenous entry of sellers and dynamics are abstracted from this paper, but they can be incorporated (see Shi, 2011). With endogenous entry of sellers, in particular, changes in the intensive margin of the market can affect the stock of relationships and the duration of a sale by affecting the buyer/seller ratio. Another feasible extension of the model is to allow each buyer to contact two or more sellers in each period, which will lead to endogenous breakups of relationships in the equilibrium.

In addition to explaining customer relationship, this model has several features that can be useful for future work to build models where aggregate fluctuations are consistent with price dynamics at the micro level. The model has many buyers and sellers, it is dynamic, and it endogenizes customer relationship. Consistent with the microdata, the model shows that prices exhibit large variability over time for the same item at a given store, sales account for a significant part of this variability, and the regular price is less responsive to market conditions than the sale price. Moreover, an increase in the particular extensive margin of demand, λ, can reduce the average markup. This result can be useful for generating counter-cyclical markups that are central to business cycle models which focus on demand shocks. Finally, the model reveals that the duration of a sale responds to market conditions more accurately than prices do. A macro model with sales should explicitly incorporate the duration of a sale as part of price adjustments.

Appendix

A. Proofs of Lemmas 3.1 and 3.2

Let me prove Lemma 3.1 first. Because $D_b(\rho_{+1}, p_1) > 0$, comparing (2.9) and (2.12) yields $v_{1s} > v_{01}$ $(= v_{11})$. Then, a type 1 buyer's optimal decision 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 $\theta_{1s} = 1$, 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 $\theta_{1s} = 1$. 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 (2.14), I compute the payoff to a seller who has a related buyer in this case as

$$J_1(\rho) = \frac{1}{1+r} \max_{p_1} \left\{ \lambda \theta_{1s} D_s(\rho_{+1}, p_1) + J_0(\rho_{+1}) \right\}, \text{ s.t. } (2.3).$$

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 other buyers' choices, because the seller only attracts his related buyer in this case. But this price reduction will increase the seller's value, which implies that $\theta_{1s} < 1$ cannot be an equilibrium outcome.

To prove (ii) in 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 (2.11) 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) in 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) in 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 $\theta_{01} = 0$. This completes the proof of Lemma 3.1.

Now turn to Lemma 3.2. Suppose, contrary to the lemma, that an equilibrium with partial mixing (and priority) exists. Let a type 1 seller's optimal choice be p_1 and the induced queue length of unrelated buyers be q_1. The pair (p_1, q_1) solves a problem similar to (3.7), except that $\tilde{J}_1$ is replaced with J_1 and that all value functions are the ones generated under the supposition that type 1 sellers attract both types of buyers. The additional constraint on a type 1 seller is that the choice of attracting both types of buyers should weakly dominate the choice of attracting only the related buyer with price $\bar{p}$:

$$\left[1-(1-\lambda)e^{-q_1}\right] D_s(\rho_{+1}, p_1) \geq \lambda D_s(\rho_{+1}, \bar{p}). \tag{A.1}$$

If (A.1) is binding, then the type 1 seller can gain by deviating to $\bar{p}$ that attracts only the related buyer. Thus, for an equilibrium with partial mixing to exist, (A.1) cannot be binding. In this case, the first-order condition of a type 1 seller's problem yields:

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

where Δ is defined in (3.4). Moreover, with $J_1(\rho)$ replacing $\tilde{J}_1(\rho)$ in (3.7), the seller's Bellman equation yields:

$$(1+r)J_1(\rho) - J_0(\rho_{+1}) = \frac{\left[1-(1+q_1)e^{-q_1}\right]\left[1-(1-\lambda)e^{-q_1}\right]^2 \Delta(\rho_{+1})}{(1-\lambda)\left[1-e^{-q_1}\right]^2 + \lambda\left[1-(1+q_1)e^{-q_1}\right]}. \tag{A.3}$$

A type 0 seller's optimization problem is (3.1), which yields the optimality conditions (3.2) and (3.3). By equating the two expressions for $[V_0^a(\rho) - V_0(\rho_{+1})]$ in (3.2) and (A.2), I get the following relation between q_0 and q_1:

$$q_0 = h(q_1) \equiv q_1 + \ln\left\{\frac{1}{1-\lambda}\left[1 + \frac{\lambda}{1-\lambda}\frac{\left[1-(1+q_1)e^{-q_1}\right]}{(1-e^{-q_1})^2}\right]\right\}. \tag{A.4}$$

It can be verified that $h'(q) > 1$ for all $q > 0$.

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. To start, I substitute $(1+r)J_0(\rho)$ from (3.3) and $(1+r)J_1(\rho)$ from (A.3) to rewrite the relation $J_0(\rho) > J_1(\rho)$ as

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

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 (A.4), I rewrite the above condition further as $f(a, q_1) < 0$, where a temporarily denotes $a = \frac{1}{1-\lambda}$ and f temporarily denotes

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

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

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

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:

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

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

B. Proof of Proposition 3.3

Let me first prove part (ii) of Proposition 3.3. The dynamics of ρ are determined by (3.11), given the initial value of ρ. For each value of ρ on the dynamic path, q_0 is given by (3.10). Because (3.10) and (3.11) depend only on (b, λ), the steady state and the dynamics of (ρ, q_0) depend only on (b, λ). Let me solve the steady state of (ρ, q_0), denoted (ρ^*, q_0^*). In the steady state, $\rho_{+1} = \rho = \rho^*$, and so (3.11) and $H(\rho) = q_0$ yield:

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

Rewriting (3.10) 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:

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

The steady-state values, (ρ^*, q_0^*), solve $\rho^* = \rho 1(q_0^*) = \rho_2(q_0^*)$. It is easy to verify that $\rho 1'(q) > 0$, $\rho 2'(q) < 0$, $\rho 1(0) = 0 < b\lambda = \rho 2(0)$, and $\rho 1(\infty) = \frac{1}{2-\lambda} > 0 = \rho 2(\infty)$. Thus, there exists a unique $q_0^* \in (0, \infty)$ that solves $\rho 1(q_0^*) = \rho 2(q_0^*)$. This result and part (i) proven below establish that the steady state is unique. 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 is stated in part (ii) of the proposition. Also, it is straightforward to use (B.1) and (B.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 (ii), I show that the steady state of ρ is locally stable. Because ρ is the only aggregate state variable, this result proves that the steady state is locally stable. The steady state of ρ is locally stable if and only if $|G'(\rho^*)| < 1$, where G is defined in (3.11). Using (3.10) to compute $H'(\rho) = (q_0 - \lambda)/(1-\rho)$ first and then computing 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 $q_0 = H(\rho) > 0$. In particular, $|G'(\rho^*)| < 1$.

For part (i) of Proposition 3.3, let me first presume $\Delta(\rho_{+1}) > 0$ (which will be verified in part (iii)) and find the conditions for $J_1(\rho) \geq J_0(\rho)$ and (3.9). Consider the condition $J_1(\rho) \geq J_0(\rho)$. Using (3.5) and V_0^a in (3.2), I compute:

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

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

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

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

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

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

Note that $q_a(\lambda)$ is an increasing function. Moreover, because $-\ln(1-\lambda) > \lambda$ for all $\lambda \in (0,1)$, it can be verified that $\frac{q}{e^q-1} > 1-\lambda$ at $q = -\ln(1-\lambda)$, which implies $q_a(\lambda) > -\ln(1-\lambda) > \lambda$ for all $\lambda \in (0,1)$. When $\lambda \to 0$, $q_a(\lambda) \to 0$ and $\frac{q_a(\lambda)}{\lambda} \to 2$. When $\lambda \to 1$, $q_a(\lambda) \to \infty$ and $(1-\lambda)q_a(\lambda) \to 0$.

Now consider (3.9), the condition that requires it to be not optimal for a type 1 seller to attract both types of buyers. Still maintain $\Delta(\rho_{+1}) > 0$. Because the queue length of unrelated buyers attracted by the best deviation of this kind, $\tilde{q}_1$, satisfies (A.2), consistency between (A.2) and (3.2) requires $q_0 = h(\tilde{q}_1)$, where h is defined in (A.4). The deviating price $\tilde{p}_1$ can be retrieved from (3.8). With $V_0^a - V_0 = e^{-q_0}\Delta(\rho_{+1})$, the best deviation is

$$\tilde{q}_1 = h^{-1}(q_0), \quad \tilde{p}_1 = U + V_1(\rho_{+1}) - V_0(\rho_{+1}) - \frac{\tilde{q}_1 e^{-q_0}\Delta(\rho_{+1})}{(1-e^{-\tilde{q}_1})(1-\lambda)}. \tag{B.7}$$

Substituting $\tilde{p}_1$ from (B.7), $\bar{p}$ from (B.3), and $q_0 = h(\tilde{q}_1)$, I can write (3.9) equivalently as $f(\tilde{q}_1) \leq 0$, where f denotes:

$$f(q) = e^{2q} - \left[2 - 2\lambda + \lambda^2 + (1+\lambda)q\right] e^q + (1-\lambda)(1-\lambda+q). \tag{B.8}$$

Compute:

$$\begin{aligned} f'(q) &= 2e^{2q} - \left[3 - \lambda + \lambda^2 + (1+\lambda)q\right] e^q + 1 - \lambda \\ f''(q) &= e^q \left[4e^q - 4 - \lambda^2\right]. \end{aligned}$$

Thus, $f''(q) > 0$ if and only if $q > \ln(1+\frac{\lambda^2}{4})$. Because $f'(0) = -\lambda^2 < 0$ and $f'(\infty) = \infty$, the property of f'' implies that there exists $q_4 > \ln(1+\frac{\lambda^2}{4})$ such that $f'(q) > 0$ if and only if $q > q_4$. In turn, because $f(0) = 0$ and $f(\infty) = \infty$, the property of f' 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 $\tilde{q}_1 \leq q_5(\lambda)$. Because $h(q)$ defined in (A.4) is an increasing function and $q_0 = h(\tilde{q}_1)$, then $\tilde{q}_1 \leq q_5(\lambda)$ if and only if

$$q_0 \leq q_b(\lambda) \equiv h(q_5(\lambda)) \in (0, \infty). \tag{B.9}$$

Lemma B.1 below proves that if the deviation indeed attracts type 0 buyers (i.e., if $\tilde{q}_1 > 0$) and if it is optimal for a type 1 seller to give priority (i.e., if $q_0 \leq q_a(\lambda)$), then $q_b(\lambda) \geq q_a(\lambda)$ for all $\lambda \in [0,1]$. If $q_0 \leq q_a(\lambda)$, then $q_0 \leq q_b(\lambda)$, and so it is not profitable for a type 1 seller to deviate to attract both types of buyers.

Therefore, the requirement $J_1(\rho) \geq J_0(\rho)$ and the requirement (3.9) are both satisfied if and only if $q_0 \leq q_a(\lambda)$. To express this condition in terms of parameters, recall that the steady-state value q_0^* satisfies $\rho 2(q_0^*) = \rho 1(q_0^*)$. Because $\rho 2(q) < \rho 1(q)$ if and only if $q > q_0^*$, the steady state satisfies $q_0^* \leq q_a(\lambda)$ if and only if $\rho 2(q_a(\lambda)) \leq \rho 1(q_a(\lambda))$, which can be rewritten as $b \leq B(\lambda)$ where

$$B(\lambda) \equiv \frac{\lambda[1 - e^{-q_a(\lambda)}] + (1-\lambda)q_a(\lambda)}{\lambda[2 - \lambda - e^{-q_a(\lambda)}]}. \tag{B.10}$$

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 (B.6) that $q_a(\lambda)$ depends only on λ and not on other parameters such as (c, U, r). Thus, $B(\lambda)$ depends only on λ. Because $q_a(1) \to \infty$ and $\lim_{\lambda \to 1}(1-\lambda)q_a(\lambda) = 0$, then $B(1) = 1$. Moreover, if q_a is treated as a separate argument from λ in (B.10), then B can be written as $\tilde{B}(\lambda, q_a(\lambda))$. For any given $\lambda \in (0,1)$, the derivative of $\tilde{B}(\lambda, q)$ with respect to q has the same sign as $(2-\lambda)e^q - 1 - q + \lambda$, which is strictly positive because $e^q > 1 + q$. For all $\lambda < 1$, because $q_a > -\ln(1-\lambda)$ and the value of B at $q_a = -\ln(1-\lambda)$ is equal to $\lambda - \frac{1-\lambda}{\lambda}\ln(1-\lambda)$, then $B(\lambda) > \lambda - \frac{1-\lambda}{\lambda}\ln(1-\lambda) \geq 1$, where the second inequality follows from $-\ln(1-\lambda) \geq \lambda$.

For part (iii) of Proposition 3.3, recall that $\bar{p}$ in (3.5) is derived from the relation $v_{1s}(\rho, \bar{p}) = V_0^a(\rho)$. Hence, (2.8) implies $V_1(\rho) = V_0(\rho)$. Substituting V_0^a from the first condition in (3.2) into (2.8) for $i = 0$ yields the expression for V in (3.14). Combining the two conditions in (3.2) to eliminate V_0^a yields the expression for p_0 in (3.12), and substituting the result $V_1 = V_0$ into (B.3) yields the expression for $\bar{p}$ in (3.13). 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 (3.15) and verify $\Delta(\rho) > 0$ for ρ near the steady-state value ρ^*, I subtract (3.3) from (B.5) to get:

$$J_1(\rho) - J_0(\rho) = \frac{1}{1+r}\left[\lambda(1 - e^{-q_0}) - 1 + (1+q_0)e^{-q_0}\right]\Delta(\rho_{+1}). \tag{B.11}$$

The definition of $\Delta(\rho)$ in (3.4) and the result $V_1 = V_0$ then imply the dynamic equation for Δ in (3.15). Setting $\Delta(\rho) = \Delta(\rho_{+1}) = \Delta^*$ and $q_0 = q_0^*$ in (3.15), I obtain:

$$\Delta^* = \frac{(1+r)(U-c)}{1+r-\lambda(1-e^{-q_0^*})+1-(1+q_0^*)e^{-q_0^*}}. \tag{B.12}$$

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, $\Delta(\rho) > 0$.

With (B.12) and (3.13), 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 positive, indeed. Thus, if ρ is close to ρ^*, then $\bar{p} > c$. This completes the proof of Proposition 3.3. **QED**

Lemma B.1. *If $q_0 \le q_a(\lambda)$ and $\tilde{q}_1 > 0$, then $q_b(\lambda) \ge q_a(\lambda)$.*

Proof. Assume $q_0 \le q_a(\lambda)$ and $\tilde{q}_1 > 0$. Define q_{a1} by $h(q_{a1}) = q_a$, where $h(q)$ is defined by (A.4). Because $h(q_5) = q_b$ and $h(q)$ is strictly increasing, $q_b \ge q_a$ if and only $q_5 \ge q_{a1}$. By the proof above, $f(q)$ defined by (B.8) satisfies $f(q) < 0$ if and only if $q \in (0, q_5)$. Thus, $q_5 \ge q_{a1}$ if and only if $f(q_{a1}) \le 0$. I prove $f(q_{a1}) \le 0$ in the following steps.

First, I obtain a restriction on λ that is needed for the maintained assumptions $q_0 \le q_a(\lambda)$ and $\tilde{q}_1 > 0$ to hold. Note that because $h(q_{a1}) = q_a \ge q_0 = h(\tilde{q}_1)$ and $h(q)$ is strictly increasing, then $q_{a1} \ge \tilde{q}_1 > 0$. Also note that the function $\frac{\left[1-(1+q)e^{-q}\right]}{(1-e^{-q})^2}$ is strictly increasing and its value at $q = 0$ is $1/2$. Because $q_{a1} > 0$, (A.4) implies

$$q_a = h(q_{a1}) > q_{a1} + \ln\frac{2-\lambda}{2(1-\lambda)^2}.$$

Equivalently, $q_{a1} < q_a - \ln\frac{2-\lambda}{2(1-\lambda)^2}$. Because $q_{a1} > 0$, this condition requires $q_a > \ln\frac{2-\lambda}{2(1-\lambda)^2}$. With the definition of q_a in (B.6), I express this requirement on q_a as

$$1-\lambda < \left.\frac{q}{e^q-1}\right|_{q=\frac{2-\lambda}{2(1-\lambda)^2}} = \frac{2(1-\lambda)^2}{\lambda(3-2\lambda)}\ln\frac{2-\lambda}{2(1-\lambda)^2}.$$

Write this condition further as

$$\lambda(\frac{3}{2}-\lambda) - (1-\lambda)\ln\frac{2-\lambda}{2(1-\lambda)^2} < 0.$$

It can be verified that this is equivalent to $\lambda \in (0, \lambda_A)$, where $\lambda_A \approx 0.7735$. Because $q_a(\lambda)$ is an increasing function, then $q_a > \ln \frac{2-\lambda}{2(1-\lambda)^2}$ if and only if $q_a(\lambda) \in (0, q_A)$, where $q_A = q_a(\lambda_A) \approx 2.4811$.

Second, I find a sufficient condition for the desired result $f(q_{a1}) \leq 0$. The function $f(q)$ has the property that if $f(Q) \leq 0$ for some $Q > 0$, then $f(q) < 0$ for all $q \in [0, Q]$. Because $q_{a1} < q_a - \ln \frac{2-\lambda}{2(1-\lambda)^2}$, then $q_{a1} < q_a + \ln(1-\lambda)$. Thus, a sufficient condition for $f(q_{a1}) \leq 0$ is $f(q_a + \ln(1-\lambda)) \leq 0$.

Third, I verify $f(q_a(\lambda) + \ln(1-\lambda)) \leq 0$ for all $\lambda \in (0, \lambda_A)$. Note that $e^{q_a + \ln(1-\lambda)} = (1-\lambda)e^{q_a} = 1 - \lambda + q_a$, where the second equality follows from using the definition of q_a to substitute e^{q_a}. Then,

$$\begin{aligned}
&f(q_a(\lambda) + \ln(1-\lambda)) \\
&= [\lambda(1-\lambda) + (1+\lambda)q_a(\lambda)] \left\{ -\ln(1-\lambda) - \tfrac{\lambda[\lambda + q_a(\lambda)][1-\lambda+q_a(\lambda)]}{\lambda(1-\lambda)+(1+\lambda)q_a(\lambda)} \right\}.
\end{aligned}$$

Clearly, $f(q_a(\lambda) + \ln(1-\lambda)) \leq 0$ if and only if the expression in $\{.\}$ is negative. Using (B.6) to substitute $\lambda = 1 - \frac{q_a}{e^{q_a}-1}$, I can write the expression in $\{.\}$ above as $L(q_a)$, where

$$L(q) = \ln\left(\frac{e^q - 1}{q}\right) - \frac{[1-(1+q)e^{-q}][1+q-(1+2q)e^{-q}]}{(1-e^{-q})[2-(3+q)e^{-q}+e^{-2q}]}.$$

This function has no parameter. It can be verified that $L(0) = 0$ and $L(q) < 0$ for all $q \in (0, q_A)$. Thus, $f(q_a(\lambda) + \ln(1-\lambda))$ is negative for all $\lambda \in (0, \lambda_A)$. That is, $f(q_{a1}) \leq 0$ whenever $q_0 \leq q_a(\lambda)$ and $\tilde{q}_1 > 0$. **QED**

C. Proofs for Section 4

C.1. Proof of Lemma 4.1

In the steady state, (2.15) yields $\rho(1 - \lambda\theta_{1s})e^{-(q_{11}+q_{01})} = (1-\rho)\left[1 - e^{-(q_{10}+q_{00})}\right]$. Substituting into the expression for ψ yields $\psi = \rho$.

Consider the allocation described by (4.2), with trading priority for related buyers. I prove that $q_{01} > 0$ if and only if $b > B_e(\lambda)$, where $B_e(\lambda)$ is defined by (4.3) in the lemma. Suppose $q_{01} > 0$ first. Then $q_{01} = q_{00} + \ln(1-\lambda)$, which indeed satisfies $q_{01} > 0$ iff $q_{00} > -\ln(1-\lambda)$. In the case $q_{01} > 0$, substituting $q_{01} = q_{00} + \ln(1-\lambda)$ into the steady-state version of (2.15) yields $\rho = 1 - e^{-q_{00}}$, which is denoted as $\rho 3(q_{00})$. The constraint

(2.7) yields: $(\frac{b}{\rho}-1)\lambda = \frac{1}{\rho}q_{00} + \ln(1-\lambda)$. Substituting $\rho = 1-e^{-q_{00}}$ into the right-hand side of this condition yields

$$\rho = b\left[1+\frac{1}{\lambda}\left(\ln(1-\lambda)+\frac{q_{00}}{1-e^{-q_{00}}}\right)\right]^{-1} \equiv \rho 4(q_{00}).$$

The level of q_{00} in the steady state solves $\rho 3(q) = \rho 4(q)$. It is easy to verify that $\rho 3'(q) > 0$, $\rho 4'(q) < 0$ and $\rho 3(\infty) = 1 > 0 = \rho 4(\infty)$. Thus, there is at most one solution to $\rho 3(q_{00}) = \rho 4(q_{00})$. This solution exists and satisfies $q_{00} > -\ln(1-\lambda)$ if and only if $\rho 3(q) < \rho 4(q)$ at $q = -\ln(1-\lambda)$. At $q = -\ln(1-\lambda)$, $\rho 3(q) = \lambda$ and $\rho 4(q) = b\left[1-\frac{1-\lambda}{\lambda^2}\ln(1-\lambda)\right]^{-1}$. Thus, the solution for q_{00} exists and satisfies the requirement $q_{00} > -\ln(1-\lambda)$ in the current case if and only if $b > B_e(\lambda)$, where $B_e(\lambda)$ is defined by (4.3). Suppose $q_{01} = 0$ next. Then, $q_{00} \leq -\ln(1-\lambda)$. In this case, a derivation similar to that in Appendix B shows that q_{00} solves $\rho 1(q) = \rho 2(q)$, where $\rho 1(q)$ is defined by (B.1) and $\rho 2(q)$ by (B.2). Moreover, this solution for q_{00} satisfies the requirement $q_{00} \leq -\ln(1-\lambda)$ if and only if $\rho 1(q) \geq \rho 2(q)$ at $q = -\ln(1-\lambda)$. This requirement is equivalent to $b \leq B_e(\lambda)$.

Because $\lim_{\lambda\to 0}\frac{1}{\lambda}\ln(1-\lambda) = -1$ and $\lim_{\lambda\to 1}(1-\lambda)\ln(1-\lambda) = 0$, then $B_e(0) = B_e(1) = 1$. For all $\lambda \in (0,1)$, the fact $-\ln(1-\lambda) > \lambda$ implies $B_e(\lambda) > 1$. This fact also implies

$$B_e(\lambda) < -\ln(1-\lambda) - \frac{1-\lambda}{\lambda}\ln(1-\lambda) = \frac{-1}{\lambda}\ln(1-\lambda), \text{ all } \lambda \in (0,1).$$

I now prove that the allocation in (4.2) together with priority for related buyers yields higher welfare than the allocation without priority. As explained immediately before Lemma 4.1, this amounts to proving $\psi > 1-e^{-b\lambda}$. Because $\psi = \rho$, it suffices to prove $\rho > 1-e^{-b\lambda}$. Consider first the case where $b > B_e(\lambda)$. In this case, $q_{01} > 0$, and q_{00} solves $\rho 3(q) = \rho 4(q)$, where $\rho 3$ and $\rho 4$ are defined above. Denote q_c temporarily as the solution to $\rho 3(q_c) = 1-e^{-b\lambda}$. Then, $\rho > 1-e^{-b\lambda}$ if and only if $\rho 4(q_c) > 1-e^{-b\lambda}$. Solving $q_c = b\lambda$, I rewrite $\rho 4(q_c) > 1-e^{-b\lambda}$ as $-\ln(1-\lambda) > \lambda$, which is satisfied for all $\lambda \in (0,1]$. Consider next the case where $b \leq B_e(\lambda)$. In this case $q_{01} = 0$, $q_{00} \leq -\ln(1-\lambda)$, and q_{00} solves $\rho 1(q) = \rho 2(q)$, where $\rho 1$ and $\rho 2$ are defined by (B.1) and (B.2). Denote q_c as the solution to $\rho 1(q_c) = 1-e^{-b\lambda}$. Then, $\rho > 1-e^{-b\lambda}$ if and only if $\rho 2(q_c) > 1-e^{-b\lambda}$. Solving q_c as $q_c = -\ln[1-(1-\lambda)(e^{b\lambda}-1)]$, I express $\rho 2(q_c) > 1-e^{-b\lambda}$ as $g(b) > 0$, where

$$g(b) = b\lambda e^{b\lambda} - \lambda(e^{b\lambda}-1) + \ln[1-(1-\lambda)(e^{b\lambda}-1)].$$

It is clear that $g(0)=0$. Also, $g'(b)>0$ if and only if $\frac{b\lambda}{e^{b\lambda}-1}-(b\lambda+1-\lambda)(1-\lambda)>0$. The expression on the left-hand side of this inequality is strictly decreasing in b. This expression is equal to $1-(1-\lambda)^2>0$ at $b=0$ and to $-\infty$ at $b=\infty$. Thus, there exists $b_c \in (0,\infty)$ such that the expression is strictly positive if and only if $b<b_c$. That is, $g(b)$ is increasing for $b<b_c$ and decreasing for $b>b_c$. Since $g(0)=0$, then $g(b)>0$ for all $b \leq B_e(\lambda)$ if and only if $g(B_e(\lambda))>0$. To prove $g(B_e(\lambda))>0$, note that since $B_e(\lambda)<\frac{-1}{\lambda}\ln(1-\lambda)$, then $\ln[1-(1-\lambda)(e^{B_e\lambda}-1)]>-B_e\lambda$, and so

$$g(B_e)>B_e\lambda e^{B_e\lambda}-\lambda(e^{B_e\lambda}-1)-B_e\lambda=(B_e-1)\lambda e^{B_e\lambda}>0,$$

where the last inequality follows from $B_e>1$. **QED**

C.2. Optimality conditions of the social planner's problem

In the social planner's problem, (4.1), let β_1 be the Lagrangian multiplier of (2.6) and β_0 the Lagrangian multiplier of (2.7). Then, the socially efficient choice of θ_{1s} satisfies:

$$\begin{aligned}&[U-c+W'(\rho_{+1})]\,e^{-(q_{11}+q_{01})}-\beta_1\\&=0 \text{ if } \theta_{1s}\in(0,1);\ \leq 0 \text{ if } \theta_{1s}=0;\ \geq 0 \text{ if } \theta_{1s}=1.\end{aligned} \tag{C.1}$$

Similarly, the efficient choices of $(q_{11},q_{10},q_{01},q_{00})$ satisfy:

$$(1-\lambda\theta_{1s})\,[U-c+W'(\rho_{+1})]\,e^{-(q_{11}+q_{01})}\leq\beta_1 \quad \text{and} \quad q_{11}\geq 0 \tag{C.2}$$

$$[U-c+W'(\rho_{+1})]\,e^{-(q_{10}+q_{00})}\leq\beta_1 \quad \text{and} \quad q_{10}\geq 0 \tag{C.3}$$

$$(1-\lambda\theta_{1s})\,[U-c+W'(\rho_{+1})]\,e^{-(q_{11}+q_{01})}\leq\beta_0 \quad \text{and} \quad q_{01}\geq 0 \tag{C.4}$$

$$[U-c+W'(\rho_{+1})]\,e^{-(q_{10}+q_{00})}\leq\beta_0 \quad \text{and} \quad q_{00}\geq 0, \tag{C.5}$$

where the two inequalities on each line hold with complementary slackness. The envelope condition is:

$$\begin{aligned}(1+r)W'(\rho)=\ &[U-c+W'(\rho_{+1})]\left[e^{-(q_{10}+q_{00})}-(1-\lambda\theta_{1s})e^{-(q_{11}+q_{01})}\right]\\&+\beta_1\,[\lambda(1-\theta_{1s})-q_{11}+q_{10}]+\beta_0\,(q_{00}-q_{01}-\lambda)\,.\end{aligned} \tag{C.6}$$

C.3. Proof of Proposition 4.2

I prove that Proposition 4.2 holds in the steady state. For the moment, assume $W'(\rho) > 0$ in the social optimum so that $U - c + W'(\rho_{+1}) > 0$ (which will be verified later). I establish (i) - (iv) below that together prove that the social optimum is characterized by (4.2):

(i) $\beta_1 > 0$ and $\beta_0 > 0$, where β_1 and β_0 are the Lagrangian multipliers of (2.6) and (2.7): Constraints (2.6) and (2.7) imply that all queue lengths are finite. Since $U - c + W'(\rho_{+1}) > 0$, then (C.3) implies $\beta_1 > 0$ and (C.5) implies $\beta_0 > 0$.

(ii) At least one of q_{00} and q_{10} is strictly positive near the steady state: If $q_{00} = 0$ and $q_{10} = 0$, then (2.15) implies $\rho = 0$ in the steady state. In this case, (2.7) is violated because it becomes $b\lambda = 0$.

(iii) $\theta_{1s} = 1$ and $q_{11} = q_{10} = 0$: If $\theta_{1s} = 1$, then $q_{11} = q_{10} = 0$ by (2.6). So, it suffices to prove $\theta_{1s} = 1$. Suppose $\theta_{1s} < 1$, to the contrary. Because a related buyer visits unrelated sellers with probability $1 - \theta_{1s} > 0$, at least one of q_{11} and q_{10} is strictly positive. Also, with $\theta_{1s} < 1$, (C.1) implies $\beta_1 \geq [U - c + W'(\rho_{+1})]\, e^{-(q_{11}+q_{01})}$ and $\theta_{1s} \geq 0$, with complementary slackness. Consider first the case where one of the two inequalities is strict, i.e., either $\beta_1 > [U - c + W'(\rho_{+1})]\, e^{-(q_{11}+q_{01})}$ or $\theta_{1s} > 0$. Then, $q_{11} = 0$ by (C.2). Since at least one of q_{11} and q_{10} is strictly positive when $\theta_{1s} < 1$, then $q_{10} > 0$, and so (C.3) holds with equality. This equality and the inequality, $\beta_1 \geq [U - c + W'(\rho_{+1})]\, e^{-q_{01}}$, together yield $q_{01} \geq q_{10} + q_{00} > 0$. Thus, (C.4) holds with equality. With this equality, the inequality $(1 - \lambda\theta_{1s})e^{-q_{01}} < e^{-(q_{10}+q_{00})}$, and (C.5), I have the following contradiction:

$$\beta_0 = (1 - \lambda\theta_{1s})\, [U - c + W'(\rho_{+1})]\, e^{-q_{01}} < [U - c + W'(\rho_{+1})]\, e^{-(q_{10}+q_{00})} \leq \beta_0.$$

Thus, $\theta_{1s} < 1$ is not socially efficient if either $\beta_1 > [U - c + W'(\rho_{+1})]\, e^{-(q_{11}+q_{01})}$ or $\theta_{1s} > 0$.

Now consider the remaining case where $\beta_1 = [U - c + W'(\rho_{+1})]\, e^{-(q_{11}+q_{01})}$ and $\theta_{1s} = 0$. With this expression for β_1, (C.4) implies $\beta_1 \leq \beta_0$, where the inequality becomes equality if $q_{01} > 0$. If $q_{01} > 0$, then $\beta_1 = \beta_0$. If $q_{01} = 0$, then $q_{00} > 0$ and so (C.5) holds with equality. This equality and (C.3) together imply $\beta_1 \geq \beta_0$, which is consistent with $\beta_1 \leq \beta_0$ only if $\beta_1 = \beta_0$. Thus, the current case has $\beta_1 = \beta_0$. Because at least one of q_{10} and q_{00} is strictly positive, as shown in (ii) above, (C.3) and (C.5) have at least one equality.

This equality, together with $\beta_0 = \beta_1$, implies $[U - c + W'(\rho_{+1})]e^{-(q_{10}+q_{00})} = \beta_1$. Since $\beta_1 = [U - c + W'(\rho_{+1})]\, e^{-(q_{11}+q_{01})}$ in the current case, then $q_{11} + q_{01} = q_{10} + q_{00}$. With this result, adding up (2.6) and (2.7) yields $q_{10} + q_{00} = b\lambda$. Substituting this result into (2.15) solves $\rho = 1 - e^{-b\lambda}$ in the steady state. Then, $\psi = 1 - e^{-b\lambda}$. Thus, the allocation in this case yields the same welfare level as the allocation without priority. By Lemma 4.1, this allocation is not socially efficient. Hence, $\theta_{1s} < 1$ is not socially efficient.

(iv) $q_{00} > 0$, $\beta_1 \geq \beta_0$, $q_{01} \geq 0$ and $q_{01} \geq q_{00} + \ln(1-\lambda)$, where the last two inequalities hold with complementary slackness: Because $q_{10} = 0$ by (iii), then $q_{00} > 0$ by (ii), and (C.3) yields $\beta_1 \geq \beta_0$. The complementary slackness on q_{01} follows from (C.4), (C.5) and $q_{00} > 0$, as explained prior to Lemma 4.1.

Now I prove that $W'(\rho) > 0$ in the steady state, which has been assumed above. Substituting (4.2) into (C.6), I get:

$$\begin{aligned}(1+r)W'(\rho) &= [U - c + W'(\rho_{+1})]\,[e^{-q_{00}} - (1-\lambda)e^{-q_{01}}] + \beta_0\,(q_{00} - q_{01} - \lambda)\\ &= [U - c + W'(\rho_{+1})]\,[(q_{00} - q_{01} + 1 - \lambda)\,e^{-q_{00}} - (1-\lambda)e^{-q_{01}}]\,.\end{aligned}$$

To obtain the second equality, I have substituted β_0 from (C.5) with the knowledge that $q_{00} > 0$ and $q_{01} \geq 0$. If $q_{01} > 0$, then $q_{01} = q_{00} + \ln(1-\lambda)$, in which case the above equation in the steady state yields

$$W'(\rho) = \frac{(U-c)[-\ln(1-\lambda) - \lambda]}{r + 1 - [-\ln(1-\lambda) - \lambda]e^{-q_{00}}}.$$

Since $q_{00} > -\ln(1-\lambda)$ in this case (with $q_{01} > 0$) and $-\ln(1-\lambda) > \lambda$ for all $\lambda \in (0, 1]$, then

$$1 - [-\ln(1-\lambda) - \lambda]e^{-q_{00}} > (1-\lambda)\left[\frac{1}{1-\lambda} + \lambda + \ln(1-\lambda)\right].$$

The expression in [.] on the right-hand side is strictly increasing in λ, and its value is equal to 1 at $\lambda = 0$. Thus, the expression is positive for all λ. This shows that if $q_{01} > 0$ then $W'(\rho) > 0$ for all $\lambda \in (0, 1]$. If $q_{01} = 0$, then $W'(\rho)$ in the steady state is given as

$$W'(\rho) = \frac{(U-c)[(q_{00} + 1 - \lambda)e^{-q_{00}} - 1 + \lambda]}{r + 2 - \lambda - (q_{00} + 1 - \lambda)e^{-q_{00}}}.$$

The derivative of the expression $[(q + 1 - \lambda)e^{-q} - 1 + \lambda]$ with respect to q is positive if and only if $q < \lambda$. Thus, the expression is maximized at $q = \lambda$. Also, the expression at $q = 0$ is equal to 0. Because $q_{00} < -\ln(1-\lambda)$ in the current case (with $q_{01} = 0$), then a sufficient condition for $(q_{00} + 1 - \lambda)e^{-q_{00}} - 1 + \lambda > 0$ is $[(q + 1 - \lambda)e^{-q} - 1 + \lambda]_{q=-\ln(1-\lambda)} > 0$. It is easy to verify that the latter condition is satisfied, noting that $-\ln(1-\lambda) > \lambda$. Moreover, because $(q + 1 - \lambda)e^{-q}$ is maximized at $q = \lambda$, then

$$2 - \lambda - (q_{00} + 1 - \lambda)e^{-q_{00}} > 2 - \lambda - (\lambda + 1 - \lambda)e^{-\lambda} = 2 - \lambda - e^{-\lambda} > 0.$$

Therefore, if $q_{01} = 0$, then $W'(\rho) > 0$ also holds.

Finally, for all $\lambda \in (0, 1)$, the inequality $B(\lambda) > B_e(\lambda)$ follows from part (i) of Proposition 3.3. The comparison between the equilibrium and the social optimum stated in Proposition 4.2 is evident. **QED**

D. Proof of Corollary 5.1

Let $r \to 0$. In the steady state, (3.12) - (3.15) yield:

$$\begin{aligned} &markup0 = (\tfrac{U}{c} - 1)[1 - \tfrac{1}{g(q_0,\lambda)}], \\ &markup1 = (\tfrac{U}{c} - 1)[1 - \tfrac{1}{(2-\lambda)e^{q_0} - (1+q_0-\lambda)}], \\ &1 - \tfrac{p_0}{p_1} = (1 - \tfrac{c}{U})[\tfrac{q_0}{1-e^{-q_0}} - 1] \Big/ \Big[\tfrac{c}{U} + (2-\lambda)e^{q_0} - (2 - \lambda + q_0)\Big], \end{aligned} \tag{D.1}$$

where the function g temporarily denotes

$$g(q, \lambda) = \frac{e^q - 1}{q}\Big[2 - \lambda - (1 + q - \lambda)e^{-q}\Big].$$

It is easy to verify the properties of $markup1$ stated in (i) of the corollary. To prove the properties of $markup0$ stated in (ii) of the corollary, let me first examine $g(q, \lambda)$. Note that $g(0, \lambda) = 1$. Substituting $\lambda = 1 - \frac{q_a}{e^{q_a}-1}$ from (B.6) yields $g(q_a, \lambda) = \frac{e^{q_a}-1}{q_a} > 1$. Compute the derivatives of $g(q, \lambda)$ with respect to q:

$$\begin{aligned} &q^2 g_1(q, \lambda) = [(q - 1)e^q + 1](2 - \lambda - e^{-q}) + (q - \lambda)[1 - (1 + q)e^{-q}], \\ &\tfrac{\partial}{\partial q}[q^2 g_1(q, \lambda)] = qe^{-q}\left[(2 - \lambda)e^{2q} + q - \lambda - 1\right]. \end{aligned}$$

Note that $\lim_{q\to 0}[q^2 g_1(q, \lambda)] = 0$. The expression $[(2 - \lambda)e^{2q} + q - \lambda - 1]$ in the derivative $\frac{\partial}{\partial q}[q^2 g_1(q, \lambda)]$ is strictly increasing in q, its value at $q = 0$ is $1 - 2\lambda$, and it approaches ∞ as $q \to \infty$. If $\lambda \le 1/2$, then $\frac{\partial}{\partial q}[q^2 g_1(q, \lambda)] > 0$ for all $q > 0$, and so $q^2 g_1(q, \lambda) > \lim_{q\to 0}[q^2 g_1(q, \lambda)] = 0$. This further implies that $g(q, \lambda) > g(0, \lambda) = 1$ for all $q > 0$. Since

q_0 is an increasing function of b (see Proposition 3.3), then $g(q_0, \lambda) > 1$ for all $b > 0$ when $\lambda \le 1/2$. If $\lambda > 1/2$, on the other hand, $\frac{\partial}{\partial q}[q^2 g_1(q, \lambda)]$ is negative if and only if q is sufficiently small. Since $\lim_{q \to 0}[q^2 g_1(q, \lambda)] = 0$, there exists $q_A > 0$ such that $g_1(q, \lambda) < 0$ if and only if $q \in (0, q_A)$. In turn, there exists $q_B > q_A$ such that $g(q, \lambda) < g(0, \lambda) = 1$ if and only if $q \in (0, q_B)$. Moreover, $q_B < q_a$ because $g(q_a, \lambda) > 1$. Since q_0 is an increasing function of b and $\lim_{b \to B(\lambda)} q_0(b) = q_a$ (see the proof of Proposition 3.3), there exist b_1 and b_2, with $0 < b_1 < b_2 < B(\lambda)$, such that $q_0(b_1) = q_A$ and $q_0(b_2) = q_B$. Moreover, $g_1(q_0, \lambda) < 0$ if and only if $b \in (0, b_1)$, and $g(q_0, \lambda) < 1$ if and only if $b \in (0, b_2)$.

From (D.1), it is evident that $markup0 < 0$ if and only $g(q_0, \lambda) < 1$. Thus, if $\lambda \le 1/2$, then $markup0 > 0$; if $\lambda > 1/2$, then $markup0 < 0$ if and only if $b \in (0, b_2)$. This proves (a) in (ii) of Corollary 5.1. The property (b) in (ii) follows from that fact that $markup0$ increases in U/c if and only if $markup0 > 0$. For (c) in (ii) of the corollary, note that markups depend on b only through q_0. If $\lambda \le 1/2$, then $markup0$ increases in q_0 and, hence, in b. If $\lambda > 1/2$, $markup0$ decreases in q_0 and, hence, in b if and only if $b \in (0, b_1)$, where $b_1 < b_2$ is defined above.

It is clear from the expression for $(1 - \frac{p_0}{p_1})$ in (D.1) that the size of the price discount of a sale increases in U/c. The dependence of the size of the discount on b, stated in (iii) of the corollary, follows from the dependence of the two markups on b.

For (iv) of Corollary 5.1, it is evident that the duration of a sale decreases in λ if and only if q_0 increases in λ. Recall that q_0 solves the equation $\rho 1(q) = \rho 2(q)$, where $\rho 1$ is defined by (B.1) and $\rho 2$ by (B.2). From this equation it can be verified that q_0 increases in λ if and only if $q_0 > b\lambda^2$, which is equivalent to $\rho 1(b\lambda^2) < \rho 2(b\lambda^2)$ and, hence, to

$$0 < b(2 - \lambda - e^{-b\lambda^2}) - \left[1 + b\lambda(1 - \lambda) - e^{-b\lambda^2}\right].$$

At $\lambda = 0$, the expression on the right-hand side is equal to $b > 0$; at $\lambda = 1$, the expression is equal to $(b - 1)(1 - e^{-b})$. Moreover, the derivative of the expression with respect to λ is $-2b\left[1 - \lambda + \lambda(1 - b)e^{-b\lambda^2}\right]$. If $b < 1$, this derivative is negative and the expression at $\lambda = 1$ is negative. In this case, there is a unique $\lambda_1 \in (0, 1)$ such that the expression is positive if and only if $\lambda \in (0, \lambda_1)$. **QED**

References

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

[2] Albrecht, J., Gautier, P.A., and S. Vroman, 2006, "Equilibrium Directed Search with Multiple Applications," Review of Economic Studies 73, 869-891.

[3] Albrecht, J., Postel-Vinay, F. and S. Vroman, 2011, "An Equilibrium Search Model of Sales," manuscript, Georgetown University.

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

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

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

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

[8] 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.

[9] 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.

[10] Galenianos, M. and P. Kircher, 2009, "Directed Search with Multiple Job Applications," Journal of Economic Theory 144, 445-471.

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

[12] Gourio, F. and L. Rudanko, 2011, "Customer Capital," manuscript, Boston University.

[13] Guerrieri, V., Shimer, R. and R. Wright, 2010, "Adverse Selection in Competitive Search Equilibrium," Econometrica 78, 1823-1862.

[14] Julien, B., Kennes, J. and I. King, 2000, “Bidding for Labor,” Review of Economic Dynamics 3, 619–649.

[15] Kircher, P., 2009, “Efficiency of Simultaneous Search,” Journal of Political Economy 117, 861-913.

[16] 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.

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

[18] Menzio, G. and S. Shi, 2011, “Efficient Search on the Job and the Business Cycle,” Journal of Political Economy 119, 468-510.

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

[20] Moen, E.R., 1997, “Competitive Search Equilibrium,” Journal of Political Economy 105, 385-411.

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

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

[23] Peters, M., 1984, “Bertrand equilibrium with Capacity Constraints and Restricted Mobility,” Econometrica 52, 1117-1127.

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

[25] 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.

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

[27] Shi, S., 2006, “Wage Differentials, Discrimination and Efficiency,” European Economic Review 50, 849-875.

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

[29] Shi, S., 2011, "Customer Relationship and Sales," Working paper 429, University of Toronto.

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

[31] Shimer, R., 2005, "The Assignment of Workers to Jobs in an Economy with Coordination Frictions," Journal of Political Economy 113, 996-1025.

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

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