A Note on the Inefficiency of Competitive Markets For Quality Goods

John R. Schroeter
Iowa State U, johns@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/econ_las_staffpapers
Part of the Categorical Data Analysis Commons, Econometrics Commons, and the Statistical Models Commons

Recommended Citation
http://lib.dr.iastate.edu/econ_las_staffpapers/246
A Note on the Inefficiency of Competitive Markets For Quality Goods

Abstract
This note describes and investigates an equilibrium model of a service market in which customers search among many firms for ones offering acceptable combinations of money price and expected waiting time. Although neither firms nor customers possess market power, the noncooperative equilibrium of the model is inefficient: Forcing customers to be more selective in their choices of suppliers can produce Pareto improvements in welfare. This result is due to an externality, in queue accession decisions, which others have identified in related contexts.

Disciplines
Categorical Data Analysis | Econometrics | Statistical Models

This report is available at Iowa State University Digital Repository: http://lib.dr.iastate.edu/econ_las_staffpapers/246
A NOTE ON THE INEFFICIENCY OF
COMPETITIVE MARKETS FOR QUALITY GOODS

John R. Schroeter
Iowa State University

May 1993
A Note on the Inefficiency of Competitive Markets for Quality Goods

Abstract

John R. Schroeter

This note describes and investigates an equilibrium model of a service market in which customers search among many firms for ones offering acceptable combinations of money price and expected waiting time. Although neither firms nor customers possess market power, the noncooperative equilibrium of the model is inefficient: Forcing customers to be more selective in their choices of suppliers can produce Pareto improvements in welfare. This result is due to an externality, in queue accession decisions, which others have identified in related contexts.
This note reviews an equilibrium model of a competitive market in which quality plays an important role. The specific quality dimension discussed is the wait required of customers to obtain an otherwise homogeneous service, although the results could be interpreted more generally. The model, itself, is not entirely new: It differs in only a few respects from one presented by De Vany and Saving [1983]. The analysis here investigates the welfare properties of the model, an aspect which has not been examined satisfactorily. The result is that equilibria of the model are inefficient. Due to the presence of an externality, market interventions can be designed to yield Pareto improvements in welfare. Section I states the model, first briefly in intuitive terms, then formally. In the interests of brevity and simplicity, the discussion is limited to those features needed to establish the desired result. In particular, a detailed characterization of firms' decision rules is omitted. Section II explains the nature of the externality and makes use of some simulation results to demonstrate the inefficiency of equilibrium. Section III briefly summarizes the main finding and relates it to similar results in the literature.

I. THE MODEL

A large number of firms are engaged in the provision of a particular service. Since customer arrivals occur randomly in time, service is time consuming, and capacity is finite, each firm typically will face a queue of customers waiting to be served. The only quality dimension that may vary
across firms is the length of time that customers must wait to secure the service. Customers' preferences with respect to this quality dimension are consistent with the goal of securing one unit of the service at the lowest possible full price, the seller's monetary fee plus the anticipated waiting time weighted by a common time value. Customers have an incentive, therefore, to search across firms for ones offering desirable combinations of service fees and queue lengths. Search is costly, however, so customers' optimal search strategies will reflect the trade-off between the costs and expected benefits of continued search. Firms take aggregate demand and customers' search strategies as given in setting a fee and choosing a service capacity to maximize expected profits per unit time. Aggregate demand is a decreasing function of the expected full cost (service fee plus waiting time cost plus search cost) of securing one unit of the service.

The Queueing Theoretic Model of Service

The stochastic operation of a representative firm is described by the M/M/1/N model of elementary queueing theory. Let the firm's customer arrival process be Poisson with mean interarrival time 1/\(\lambda\). Suppose that service times are independently, identically distributed negative exponential random variables with mean 1/\(S\), so that the average service rate, or service capacity, is \(S\). Arriving customers who find \(N-1\) or fewer customers already in the system, that is, in the queue or being served, will join the queue whereas those who find \(N\) or more customers ahead of them will balk. The steady state solution to the model is a set of system size probabilities; \(q_n\) for \(n = 0, 1, 2, \ldots, N\); which can be thought of as the limiting proportions of time spent in each system state on an infinitely long sample path. These probabilities can be written as functions of the
parameters ρ, S, and N where ρ = λ/S. From these, the following results can readily be deduced. 4

\[ q_N = \text{Prob} \{N \text{ customers in the system}\} \]
\[ = \text{Prob} \{\text{An arrival will balk}\} = \frac{(1-\rho)^N}{1-\rho^{N+1}} \]  
(1)

The proportion of arrivals who join the queue is 1-q_N, so the steady state rate of queue accessions, λ', is given by:

\[ \lambda' = (1-q_N) \lambda = \lambda \frac{1-\rho^N}{1-\rho^{N+1}} \]  
(2)

The expected waiting time for queue joiners, W, is given by:

\[ W = \frac{1}{S} \frac{1-(N+1)\rho^N + N\rho^{N+1}}{(1-\rho^N)(1-\rho)} \]  
(3)

The Customers' Problem

Imagine a group of firms, firms of type i, say, each of which charges a price of p_i, employs service capacity S_i, experiences an arrival rate of λ_i, and faces a balking value of N_i. 5 An arrival to a type i firm who becomes the n^th in the system faces a conditional (upon system size) expected full price of P_i = p_i + v*n/S_i where v is the constant time value. P_i, like all conditional expectations, is a random variable. Denote its mean by \( \bar{P} \), its distribution by \( F_i(\cdot) \) and notice that \( F_i(\cdot) \) will depend on p_i, S_i, N_i, and \( \lambda_i \). Suppose that this full price distribution is known to all customers.

If a potential customer were to search among type i firms, making independent, sequential drawings from the distribution \( F_i(\cdot) \), he/she would do so using a search strategy that minimized the expected full cost of acquiring one unit of the service. 6 Let the dollar equivalent cost of sampling a firm be equal to σ. 7 It is well known that optimal search strategies in this setting are of the reservation price form. A standard
search theory exercise produces the following implicit definition of \( R_i \), the optimal reservation price:\(^8\)

\[
\sigma = F_i(R_i)\left[ R_i - \frac{\int_0^{R_i} x dF_i(x)}{F_i(R_i)} \right] \tag{4}
\]

\( R_i \), the reservation full price, is related to \( N_i \), the balking value or reservation queue length, via:

\[
R_i = p_i + \frac{N_i}{S_i} \tag{5}
\]

\( F_i(R_i) \) is the probability that a conditional expected full price drawn from \( F_i(\cdot) \) is acceptable. This is the event of a queue accession, hence

\[
F_i(R_i) = 1 - q_{N_i} \quad . \text{ Finally, } \int_0^{R_i} x dF_i(x)/F_i(R_i) \text{ is the expected full price conditional on } p_i < R_i. \text{ This is simply } p_i + vW_i. \]

With these substitutions, equation (4) can be rewritten as:

\[
\sigma = (1 - q_{N_i}) \left[ v \frac{N_i}{S_i} - vW_i \right]
\]

which, using (1) and (3) becomes:\(^9\)

\[
\frac{\sigma}{v/S_i} = \frac{1 - \rho_i}{N_i + 1} \left[ N_i - \frac{1 - (N_i + 1) \rho_i}{N_i + N_i \rho_i^N_i + 1} \right] \tag{6}
\]

The minimal value of the expected full cost of acquiring one unit of the service will be:

\[
C_i = p_i + \sigma m_i + vW_i
\]

where \( m_i = 1/F_i(R_i) \) is the expected number of searches undertaken.\(^{10}\) This can be written as:
Customers know the distributions of conditional expected full prices for all types of firms, and so can calculate $N_i$ and $C_i$ for each type. Customers will only search within those groups offering the lowest values for $C_i$ available in the market.

The Firms' Problem

The objective function of firms is expected profits per unit time. Expected revenue per unit time is simply $p_i$ times the steady state rate of queue accessions, $\lambda_i$. Expected costs per unit time are assumed to be a function of $\lambda_i$ and service capacity, $S_i$. In view of equation (2), which specifies $\lambda_i$ as a function of $\lambda_i$, $S_i$, and $N_i$, expected profits per unit time can be written as a function of these variables as well as $p_i$. Firms choose $p_i$ and $S_i$ while taking as given the nature of search behavior by customers and the market determined expected full cost of acquiring the service, $\bar{C}$. For any choices of values for $S_i$ and $p_i$, equation (6) and equation (7) with $C_i = \bar{C}$ will determine values for $\lambda_i$ and $N_i$. These constraints on the representative firm's choice setting allow the expected profit rate function to be expressed in terms of $p_i$, $S_i$, and $\bar{C}$ alone. Assuming that optimal values of $p_i$ and $S_i$ exist and are unique, they can be expressed as functions of $\bar{C}$ alone. Denote these as:

$$p_i = p(\bar{C}) \quad (8)$$
$$S_i = S(\bar{C}) \quad (9)$$

Moreover, since all firms have the same cost function and all face the same value of $\bar{C}$, all will choose the same price and capacity. Consequently, the
"i" subscripts will be dropped hereafter.

Aggregate Demand

The model, to this point, consists of equations (6), (7), (8), and (9). To close the model, it remains to introduce the aggregate demand function and specify the relationship between the arrival rate of customers to the market as a whole and to the representative firm. Customer arrivals to the market are assumed to be a Poisson process with expected interarrival time $1/\lambda$, where $\lambda$ is a decreasing function of the market determined expected full cost of acquiring one unit:

$$\lambda = \lambda(c) \quad (10)$$

In steady state equilibrium, the aggregate rate of searching will be $M\lambda$, where $M$ is the fixed number of firms. For this rate of searching to be supported by the aggregate arrival process, it is necessary that $Am = M\lambda$ where $m$, as previously defined, is the expected number of searches per customer. This condition can be written as:

$$\rho = \frac{A}{SM} \frac{1-p^{N+1}}{1-p^N} \quad (11)$$

The model is now complete. It consists of equations (6), (7), (8), (9), (10), and (11) with endogenous variables $\overline{c}$, $p$, $S$, $N$, $\rho$, and $\lambda$.

II. INEFFICIENCY OF EQUILIBRIUM

This section will demonstrate that the model's individually rational search strategy (characterized, for our purposes, by the value for $N$ which solves equation (6)) is socially inefficient. Greater total surplus can be derived by imposing the requirement that all customers search more intensively (adopt lower reservation queue lengths) than is consistent with individual optimization. The reason for this incompatibility of individual
and social objectives is an external diseconomy in customers' queue accession decisions. Customers who are indifferent between joining a queue of a particular length and searching further could balk, leaving themselves no worse off, while improving the distribution of queue lengths for subsequent arrivals.

Suppose that a central authority were to intervene in equilibrium by simply requiring that all customers search with a lower reservation queue length. How things would change as a result is a complicated question and one which we are not equipped to address since we have not explored the details of the firms' pricing and capacity choice rules. To effect the demonstration of the inefficiency of equilibrium more simply, we will consider an even more heavy-handed method of intervention. The central authority will impose a lower balking value and simultaneously suspend firms' decision rules. Firms will be directed to maintain the same service capacity as in the initial equilibrium and to adjust price so that $\bar{C}$, the expected full cost of obtaining the service, remains constant. Since firms will not choose $p$ and $S$ to maximize expected profits per unit time, the specific forms of equations (8) and (9) are not a concern. Since $\bar{C}$ remains constant, $A$ will remain unchanged regardless of the form of the demand function (10). Moreover, since neither the steady state rate of provision of the service nor the expected full cost are affected, customers, as a class, will be indifferent with respect to the intervention. Thus any increase in surplus that might obtain will accrue to firms.

With $\bar{C}$, hence $A$, and $S$ held constant, equation (11) will dictate the response in $p$ to the imposed reduction in $N$. One can easily show that (11) implies:
As the balking value goes down, for fixed M, S, and A, customers search more times on average so \( \lambda \) (hence \( p \)) goes up. To see what change in \( p \) will be required to maintain constant \( \bar{C} \), consider the effect that the change in \( N \), and the attendant response in \( p \), will have on expected waiting and search costs per customer. Use equation (7) to express these in "search equivalents" as:

\[
\tilde{C} = \frac{C}{\bar{C}} = \frac{1 - p^{N+1}}{1 - (N+1)p} + \frac{v/S}{\sigma} w
\]

or, using (11):

\[
\tilde{C} = \frac{1}{\alpha} \rho + \frac{1}{\beta} SW
\]

where \( \alpha = \Lambda/S\cdot M \) and \( \beta = \sigma\cdot S/v \). The total derivative of \( \tilde{C} \) with respect to \( N \) is:

\[
\frac{d\tilde{C}}{dN} = \frac{1}{\alpha} \frac{dp}{dN} + \frac{1}{\beta} \left[ S \frac{\partial W}{\partial N} + S \frac{\partial W}{\partial p} \frac{dp}{dN} \right]
\]

The parameter \( \alpha \) is the overall arrival rate divided by the industry service rate. Clearly this will be positive and must be less than one for a steady state equilibrium to exist. The parameter \( \beta \) is the cost of one search divided by the value of the expected duration of service. \( \beta \) can thus be thought of an an index of search costs relative to waiting costs and can assume any positive value. It can be shown that \( S \frac{\partial W}{\partial N} \) and \( S \frac{\partial W}{\partial p} \) are both positive.\(^{16}\) Recalling that \( dp/dN < 0 \), expression (13) fails to unambiguously sign \( d\tilde{C}/dN \) in general.\(^{17}\) This is not surprising, of course, since, for a given aggregate arrival rate, a decrease in \( N \) will decrease waiting costs and increase search costs. The effect on the sum of waiting and search costs is unclear.
To demonstrate the inefficiency of equilibrium, we need consider only the effects on waiting plus search costs of a change in the equilibrium balking value. That is, our interest is restricted to the signs of \( \frac{d\tilde{C}}{dN} \) evaluated at equilibrium values of \( \rho \) and \( N \). For arbitrary values of \( \alpha \) and \( \beta \), corresponding equilibrium values of \( \rho \) and \( N \) are determined by equations (6) and (11). Since the system (6) and (11) is highly nonlinear, further progress along an analytical course is difficult. One can easily proceed, however, through resort to simulation methods. The table displays, for several pairs of values for \( \alpha \) and \( \beta \), equilibrium values for \( \rho \) and \( N \), determined by (6) and (11), and the associated values of \( \frac{d\tilde{C}}{dN} \). Note that \( \frac{d\tilde{C}}{dN} \) is positive in equilibrium, at least for the wide ranges of values for \( \alpha \) and \( \beta \) examined. Thus a decrease in \( N \) would reduce \( \tilde{C} \) (waiting costs decrease more than search costs increase) and allow \( \rho \) to be increased without changing \( \tilde{C} \). \( S, A, M, \) and hence \( \lambda' \) are unchanged so firms costs are unchanged. As noted above, consumers are as well off as before the intervention since \( \tilde{C} \) and \( A \) have not changed. Yet \( \rho \) has increased so expected profits per unit time are greater for each firm. In equilibrium, the probability of a balk, \( q_N \), varies directly with \( \alpha \) and inversely with \( \beta \). Moreover, the table reveals that the equilibrium values of \( \frac{d\tilde{C}}{dN} \) vary directly with \( q_N \). The gain in efficiency to be derived by reducing the balking value is greater the more frequently the balking value is encountered.

III. SUMMARY

In the model presented here, customers search independently among many firms for ones offering suitably low full prices of service. Firms, too,
act independently in selecting prices and service capacities to maximize their expected profit rates. Moreover, individual agents possess no market power: Customers regard the distribution of full prices as beyond their influence and firms view the market determined expected full cost of service and the aggregate rate of customer arrivals as parameters. Yet the equilibrium of the model is inefficient. This result is due to an externality in customers' queue accession decisions. An individual's criterion for joining a queue does not reflect the costs to be imposed on subsequent arrivals. Search externalities of this general sort, and the consequent non-optimality of equilibria, have been identified in other settings. Examples include papers by Diamond [1981 and 1982]. The contribution of this note has been to demonstrate the presence of the phenomenon in a queueing-theoretic, search-equilibrium model of a competitive market for a quality good.
Simulation Results.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\rho$</th>
<th>$N$</th>
<th>$q_N$</th>
<th>$\frac{d\bar{C}}{dN}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>0.25</td>
<td>0.223</td>
<td>1.363</td>
<td>0.103</td>
<td>0.529</td>
</tr>
<tr>
<td>0.4</td>
<td>0.25</td>
<td>0.512</td>
<td>1.503</td>
<td>0.219</td>
<td>0.806</td>
</tr>
<tr>
<td>0.6</td>
<td>0.25</td>
<td>0.904</td>
<td>1.720</td>
<td>0.336</td>
<td>1.029</td>
</tr>
<tr>
<td>0.8</td>
<td>0.25</td>
<td>1.465</td>
<td>2.143</td>
<td>0.454</td>
<td>1.324</td>
</tr>
<tr>
<td>0.2</td>
<td>0.50</td>
<td>0.212</td>
<td>1.665</td>
<td>0.060</td>
<td>0.153</td>
</tr>
<tr>
<td>0.4</td>
<td>0.50</td>
<td>0.467</td>
<td>1.867</td>
<td>0.145</td>
<td>0.266</td>
</tr>
<tr>
<td>0.6</td>
<td>0.50</td>
<td>0.788</td>
<td>2.168</td>
<td>0.238</td>
<td>0.381</td>
</tr>
<tr>
<td>0.8</td>
<td>0.50</td>
<td>1.217</td>
<td>2.735</td>
<td>0.343</td>
<td>0.558</td>
</tr>
<tr>
<td>0.2</td>
<td>0.75</td>
<td>0.207</td>
<td>1.945</td>
<td>0.037</td>
<td>0.064</td>
</tr>
<tr>
<td>0.4</td>
<td>0.75</td>
<td>0.445</td>
<td>2.189</td>
<td>0.102</td>
<td>0.130</td>
</tr>
<tr>
<td>0.6</td>
<td>0.75</td>
<td>0.732</td>
<td>2.549</td>
<td>0.180</td>
<td>0.209</td>
</tr>
<tr>
<td>0.8</td>
<td>0.75</td>
<td>1.104</td>
<td>3.226</td>
<td>0.275</td>
<td>0.337</td>
</tr>
<tr>
<td>0.2</td>
<td>1.00</td>
<td>0.204</td>
<td>2.213</td>
<td>0.023</td>
<td>0.031</td>
</tr>
<tr>
<td>0.4</td>
<td>1.00</td>
<td>0.432</td>
<td>2.489</td>
<td>0.074</td>
<td>0.076</td>
</tr>
<tr>
<td>0.6</td>
<td>1.00</td>
<td>0.699</td>
<td>2.897</td>
<td>0.141</td>
<td>0.135</td>
</tr>
<tr>
<td>0.8</td>
<td>1.00</td>
<td>1.036</td>
<td>3.668</td>
<td>0.228</td>
<td>0.235</td>
</tr>
</tbody>
</table>

For example, from an equilibrium with $\alpha = 0.8$, $\beta = 0.75$, $\rho = 1.104$, and $N = 3.226$ a 1 unit reduction in the balking value will reduce expected search plus waiting costs per customer by 33.7% of the cost of one search.
Notes

*Associate Professor, Department of Economics, Iowa State University.

My understanding of the issues discussed herein has been enhanced through correspondence with Arthur De Vany.

1 De Vany and Saving [1983] claim to prove that equilibria are efficient, but their efficiency criterion is inappropriate in this setting. (See note 14.)

2 In this application of Kendall's notation, the "M"'s identify the interarrival and service times as i.i.d. negative exponential random variables. The "1" indicates that there is one "server." The "N" signifies that the system size is limited to N customers.

3 It remains to be shown, of course, that a balking rule of this form is, in fact, consistent with individually rational search strategies.

4 Gross and Harris [1974] provide an excellent reference for those results of the M/M/1/N model which are stated here without proof. The discussion in this paper will be limited to the case \( p \neq 1 \), although this is purely a matter of convenience. With \( p = 1 \), most of the expressions to follow would still apply as long as they were interpreted as their limiting values as \( p \) approaches 1.

5 Firms establish a price and a service rate that do not vary as the length of the queue varies. We will ultimately see that, in equilibrium, all firms make identical choices of \( p \) and \( S \) and face the same values for \( \lambda \) and \( N \). For now, we allow that these variables may differ across firms. It turns out to be a bit awkward, however, to think of each firm as having its own unique set of values for \( p, S, \lambda, \) and \( N \) since we will soon need to imagine customers making independent sequential drawings from the distributions of full prices implied by \( p, S, \lambda, \) and \( N \). Thus we model the
population of firms as being comprised of several large "groups." Firms differ across groups but are identical within groups.

6If certain types or groups of agents held a comparative advantage in information acquisition, markets for their search services would arise. It is conceivable that the externality identified here could be internalized through this market mechanism. The present analysis assumes that there is no more efficient alternative to search conducted independently by each customer.

7The full cost of observing the queue at a firm may have a time as well as a money component. As long as the value of time is assumed constant, nothing is lost by expressing the full cost of search in terms of its equivalent dollar cost.

8Equation (4) expresses equality between the expected marginal costs and benefits of search. Since $F_i(*)$ is discontinuous, equation (4) may not have a solution. More formally, $R_i$ is determined as

$$R_i = \sup \{B: h(B) > \sigma\} \text{ where } h(B) = F_i(B)\left[1 - \int_0^B dF_i(x)/F_i(B)\right].$$

9We will proceed as if the value of $N_i$ which, for given $P_i$, $\sigma$, and $S_i$, solves (6) is the reservation queue length even though this value need not be an integer. More formally, $N_i$ would be the greatest integer such that the right hand side of (6) is at least as great as the left hand side.

10The number of independent drawings from $F_i(*)$ before and including the first success (the first drawing of a full price less than or equal to $R_i$) is a random variable with the geometric distribution. Its mean, $m_i$, is equal to one over the probability of success in a single trial.

11Firms in this model are not price takers in the conventional sense, but "expected-full-cost-of-acquiring-one-unit takers." This is a small departure from De Vany and Saving [1983] who apparently have in mind "expected-full-price $(P)$ taking" behavior by firms.
Customers, like firms, are risk neutral since their decision to commence search depends only on the expected full cost of obtaining the service. Since drawings from the full price distribution are independent, the expected incremental cost of obtaining the service remains \( \overline{C} \) after any number of unsuccessful searches. Thus a customer, having once begun search, will not stop until a conditional expected full price below the reservation price is found. De Vany and Saving [1983] take aggregate demand to be a function of the unconditional expected full price, \( \overline{P} \), rather than \( \overline{C} \). Thus the costs of search do not affect aggregate demand in their specification.

\( M \) could be made endogenous by adding a zero expected profit condition. Note that searching customers regard observations of conditional expected full prices to be independent drawings from the steady state distribution of system sizes. If \( M \) is finite, these random variables would, in fact, be correlated. An improbably long queue at one firm makes more likely the observation of an improbably long queue at some other firm since both of these events could be the result of a realized market demand rate above the mean rate of \( A \). Assume that \( M \) is sufficiently large so that customers can ignore the potential for adapting their expectations of queue length through learning based on this correlation of system sizes across firms.

De Vany and Saving claim to establish an efficiency result for their model however their social benefit function takes no explicit account of search costs. Thus, they miss the possibility for the beneficial tradeoffs between search and waiting cost which can be effected through adoption of alternate balking rules. Moreover, it's clear that search costs must be included in the efficiency criterion if meaningful welfare properties are to be established. Without this feature, the "optimum" of the model would
involve all customers searching until an empty queue is found. Agents would thereby economize on waiting costs to the greatest degree possible while engaging in clearly wasteful expenditures of search resources.

The inefficiency of equilibrium result is robust. Our strategy here is to demonstrate it using a simple special case.

The expressions for these derivatives are:

\[
\frac{\partial W}{\partial N} = \frac{-\rho^N \left( (1-\rho^N) + N \ln \rho \right)}{(1-\rho^N)^2} > 0,
\]

and

\[
\frac{\partial W}{\partial \rho} = \frac{(1-\rho^N)^2 - N^2 \rho N (1-\rho)^2}{(1-\rho^N) (1-\rho)^2} > 0 \quad \text{for } \rho \neq 1, N = 1, 2, \ldots
\]

Expected waiting time increases as \( N \) increases and as \( \lambda \) increases relative to \( S \).

The ambiguity does not disappear when use is made of the specific expressions for \( \partial \rho / \partial N, \partial W / \partial N, \) and \( \partial W / \partial \rho \).

These can be rewritten, to involve only \( \alpha, \beta, \rho, \) and \( N \) as:

\[
\beta = \frac{1-\rho^N}{1-\rho} \left[ N - \frac{1-(N+1)\rho^N + N \rho N+1}{(1-\rho^N) (1-\rho)} \right]
\]

(6)

\[
\rho = \frac{\alpha \left( \frac{1-\rho^N}{1-\rho} \right)}{1-\rho^N}
\]

(11)

Since \( \alpha \) and \( \beta \) involve variables that are endogenous to the full model, not all pairs of values for these parameters are necessarily consistent with equilibrium. The desired result actually obtains for all pairs of values within the permissible ranges, and so for all equilibrium points.

Greater relative demand intensity increases the frequency of balking for fixed \( \beta \). Lower costs of search, relative to waiting, increases the frequency of balking for fixed \( \alpha \).

A comment by Diamond to a paper by Mortensen [1982] also demonstrates the presence of a similar externality in an equilibrium model of matching.


