Dynamic Price Competition in Auto-Insurance Brokerage*

Bruno Aurichio† Luis H.B. Braido‡

November 25, 2009

Abstract

Brazilian data on auto-insurance present an intriguing fact: the coexistence of policies being sold with zero and positive brokerage fees. We extend the Bertrand model of price competition to a dynamic environment in which agents face a fixed cost to switch to a new (unmatched) broker. This incumbent's advantage drives unmatched brokers to set fees aggressively. We explicitly derive a symmetric recursive Nash equilibrium in which zero and positive brokerage fees are played with positive probability. We then use the mixed-strategy equilibrium distribution to perform a maximum likelihood estimation of the model parameters. Using data on brokerage fees alone, this structural econometric procedure allows us to identify: (i) the number of brokers effectively bidding to a given consumer; (ii) the insuree's cost of switching brokers; and (iii) the expected life-time discounted profits of matched and unmatched brokers.

Keywords: Brokerage, fees, price competition, structural estimation, auto-insurance.

JEL Classification: D4; L1.

*We are thankful for comments from Marcelo Medeiros, Humberto Moreira, Heleno Pioner, Marcelo Sant'Anna, and seminar participants at IPEA, PUC-Rio, and USP.

†FUCAPE Business School and Getulio Vargas Foundation (FGV-EPGE); aurichio@fucape.br.

‡Getulio Vargas Foundation, Graduate School of Economics (FGV-EPGE); Praia de Botafogo 190, s.1100, Rio de Janeiro, RJ 22253-900, Brazil; lbraido@fgv.br.

1 Introduction

Brokers intermediate sales in many different markets such as real state and insurance. In the automobile insurance market, for instance, they act as retailers for the insurance companies and help consumers to fill out reimbursement papers in the case of a claim. This sector is very competitive in Brazil. Thousands of independent brokers offer policies from different insurance companies and add a brokerage fee to the insurance wholesale price.

In this paper, we analyze Brazilian data on auto-insurance contracts and document a interesting fact: the coexistence of policies being sold with zero and positive brokerage fees. This evidence is not consistent with static models of price competition since Nash equilibrium requires that strategies played with positive mass display the same expected payoff. Therefore, recent versions of Bertrand and searching models—such as Varian (1980), Stahl (1989), Sharkey and Sibley (1993), and Braido (2009)—cannot be used to directly account for this feature of the data.

We introduce dynamics into a Bertrand model of price competition in which insurees face a switching cost when signing a contract with a new unmatched broker. This incumbent's advantage drives unmatched brokers to set fees aggressively and generates a mass of policies with zero brokerage fees. We explicitly derive a symmetric recursive Nash equilibrium for this model and use the mixed-strategy equilibrium distribution to perform a maximum likelihood estimation of the model parameters. Using data on brokerage fees alone, this structural econometric procedure allows us to identify: (i) the number of insurance quotes effectively accessed by insurees before purchasing a policy; (ii) the insuree's cost of switching; and (iii) the expected life-time discounted profits of matched and unmatched brokers. These are important concepts in industrial organization which are difficult to obtain from data on prices.

Our work starts in Section 2 by analyzing a thorough data set recorded by the Brazilian Superintendence of Private Insurance (SUSEP) which conveys population information on auto-insurance contracts. We note that about 20 percent of the analyzed contracts have no brokerage fee. Moreover, the evidence indicates a strong dispersion in fees. Our attempt to understand this evidence relies on a game theory model in which brokers use mixed strategies to set different offers to identical insurees.

Section 3 formally describes our dynamic game of price competition and derives the distribution of price offers that is consistent with a symmetric recursive Nash equilibrium. In Section 4, we stress that the database displays solely the contracted fees, not all bids accessed by the insuree. We then use the mixed-strategy equilibrium distribution and the individual decision rules to derive the data-generating distribution. This distribution is used in Section 5 to identify the parameters of the model through a maximum likelihood estimation. Our results identify that: (i) insurees typically access two offers before purchasing an insurance policy; (ii) they face a switching cost equivalent to 142 BRL (approximately US$ 71 in purchasing power parity); and (iii) the broker's expected lifetime discounted profit is about 269 BRL when trading to a matched consumer, and about 127 BRL when trading to an unmatched consumer. These values are economically meaningful.

2 Data

We use data recorded by the Brazilian Superintendence of Private Insurance (SUSEP) which contain information on brokerage fees and insuree characteristics for every auto-insurance contract signed in Brazil during the first semester of 2003. Table 1 defines all variables used in our analysis. As is stressed there, we restrict attention to comprehensive policies contracted by individuals.[1]

[Table 1]

A key aspect of the data is that policies are sold with different brokerage fees. This feature is very robust and appears in many different subsamples of the data. Figure 1 presents histograms for nominal brokerage fees expressed in BRL, and Figure 2 displays the broker's percentage fees as a fraction of the policy total premium.

[Figures 1-2]

Six different subsamples are presented. Subsample A shows the fees paid to brokers across all insurance contracts in the country. Subsample B focuses on contracts in the metropolitan area of Sao Paulo (the largest in the country). Subsample C presents fees on policies covering nine comparable compact cars with identical power trains in the metropolitan area of Sao Paulo.[2] Next, Subsample D refines Subsample C to include only contracts with positive bonus discount. Insurees with positive bonus discount are particularly interesting for our purposes because they must have signed with a broker in the period before. Subsamples E and F further refine Subsample D to include only

[1] We do not consider: (i) policy endorsements, (ii) noncomprehensive contracts, (iii) policies contracted by firms, and (iv) collective policies covering multiple vehicles.

[2] The vehicle models are: Peugeot 206 1.0L, Clio 1.0L, Gol 1.0L, Fiesta 1.0L, Ka 1.0L, Uno 1.0L, Palio 1.0L, Corsa 1.0L, and Celta 1.0L.

new compact cars and new Volkswagen Gol 1.0L, respectively. Note that the data refer to 2003 year models and that the Volkswagen Gol 1.0L accounts for 30 percent of its class within our data set.

On average, about 20 percent of the insurance policies are sold with zero fees. (This is omitted in Figure 1 for scale reasons.) Furthermore, there is large dispersion in both nominal and percentage brokerage fees. The shapes of the histograms are surprisingly similar across the different subsamples. This suggests that there might be some relevant feature behind the dispersion in brokerage fees paid by consumers.

We will use subsample E in our quantitative analyses throughout the paper and subsample F to check the robustness of our results in Section 5.2. The choice of analyzing more homogeneous subsamples is justified by the fact that we attempt to understand the economics behind the dispersion in brokerage fees which are unrelated to consumer heterogeneity.

In line with this argument, Table 2 presents reduced-form OLS regressions for three different variables. First, we analyze how observable characteristics are correlated to the choice of zero brokerage fees. The following two regressions consider the subsample of contracts with positive fees to analyze the correlation between characteristics and fees in both nominal and percentage values. In all three regressions, we use an enormous amount of dummy variables to capture different characteristics of contracts and insurees. This large number of control variables account for 10 to 23 percent of the dispersion of fees, as follows from the R^2 values. This suggests that a significant part of the dispersion in fees is not related to observed consumer heterogeneity.

[Table 2]

3 A Dynamic Model of Price Competition

Consider a dynamic recursive economy populated by a continuum of insurance brokers and by a finite number of homogeneous insurees. These two sets—brokers and insurees—are fixed over an infinite time horizon. Insurees demand one insurance policy per period, and they must buy it through a broker. They minimize total costs in each period. Since we will focus on recursive Nash equilibria, this is equivalent to minimizing life-time discounted costs.

All policies in the market last for one period. Brokers access the same pool of auto-insurance companies and face the same wholesale prices. They maximize expected life-time discounted profits, and their intertemporal discount factor is $\beta \in (0, 1)$.

In each period, brokers set independent brokerage fees $q \in Q \subset \mathbb{R}$ to be charged to each potential customer. In principle, brokers are not allowed to charge negative prices. In practice, however, they could do so by offering nonpecuniary gifts to consumers. We then allow for negative fees and assume that the data set is censored at $q = 0$. That is, the values of such gifts are not observed by the econometrician.

In the initial node, each insuree is randomly matched to a broker. In each subsequent node, one of the following two possibilities occurs: (a) with probability $\lambda \in (0, 1)$, the insuree remains matched to the last broker one has contracted with; and (b) with probability $(1 - \lambda)$, the insuree is randomly matched to a new broker. The parameter λ accounts for the fact that some consumers exogenously enter and leave the market in every period.[3]

In every period, each insuree observes the offer from a matched broker and $I \geq 1$ offers from unmatched brokers. Insurees seek the smallest total cost and, in addition to the brokerage fee, they face a fixed cost $c > 0$ to switch from a matched to an unmatched broker. This is an important advantage of the matched broker.

This environment can be seen as if $I + 1$ brokers were setting prices to attract a single consumer. This switching-cost version of the Bertrand game is repeatedly and independently played for each consumer. We formally describe this as a recursive game, in the sense of Duffie et al. (1994, p. 771).

Recursive Game. Consider a game with $I+1$ brokers and a single consumer. Nature plays in the initial node and defines one broker to be matched to that consumer. Next, all brokers set their fees. The game ends for those who do not make a sale. For the winning broker, an identical game starts in the next period with probability λ. In this case, that broker is surely selected by nature to be matched to the consumer.

In each period, the space of actions is $Q \subset \mathbb{R}$. The brokers' payoffs depend on a state variable $z \in Z = \{m, u\}$ that describes the matched/unmatched status. Conditional on z, the payoffs do not depend on the previous history of the game.

A recursive strategy for each player $i \in \mathbb{I} = \{0, 1, ..., I\}$ is a map $\sigma_i : Z \to \Delta$, where Δ is the space of probability distributions on Q. We use the following notation throughout the paper: $\sigma_{-i} = (\sigma_{\hat{\imath}})_{\hat{\imath} \neq i}$ and $(\sigma_0, ..., \sigma_I) = (\sigma_i, \sigma_{-i})$.

Payoff Functions. For each broker i, define $v_{i,m}$ and $v_{i,u}$ as the expected life-time discounted payoff function when i faces a matched and unmatched insuree, respectively.

[3] Since the probability of the scenarios (a) and (b) add up to one, the number of consumers (just like the set of brokers) remains constant over time.

The function $v_i : Z \times Q \to \mathbb{R}$ is defined by:

$$v_{i,m}(q) = \prod_{\hat{\imath} \in \mathbb{I} \setminus \{i\}} \left(1 - \sigma_{\hat{\imath},u}(q - c)\right) \left(q + \lambda\beta \int v_{i,m}(\tilde{q})\, d\sigma_{i,m}(\tilde{q})\right), \tag{1}$$

and

$$v_{i,u}(q) = \left(1 - \sigma_{i^*,m}(q + c)\right) \prod_{\hat{\imath} \in \mathbb{I} \setminus \{i^*, i\}} \left(1 - \sigma_{\hat{\imath},u}(q)\right) \left(q + \lambda\beta \int v_{i,m}(\tilde{q})\, d\sigma_{i,m}(\tilde{q})\right), \tag{2}$$

where $\sigma_{i,m}$ and $\sigma_{i,u}$ are the recursive strategies of broker i when dealing with a matched and an unmatched insuree. Note that i^* represents the matched broker in equation (2).

To understand equation (1), notice that $\prod_{\hat{\imath} \in \mathbb{I} \setminus \{i\}} \left(1 - \sigma_{\hat{\imath},u}(q - c)\right)$ is the probability that the insuree accepts a fee q from the matched broker. In other words, it is the probability that all fees set by the I unmatched brokers are greater than $q - c$. Moreover, the term $\lambda\beta \int v_{i,m}(q)\, d\sigma_{i,m}(\tilde{q})$ represents broker i's expected discounted continuation value of making a sale today. (Recall that, after a purchase, the consumer becomes matched to the broker with probability λ.)

Analogously, to understand equation (2), notice that the probability that the insuree accepts a fee q from the unmatched broker is $\left(1 - \sigma_{i^*,m}(q + c)\right) \prod_{\hat{\imath} \in \mathbb{I} \setminus \{i^*, i\}} \left(1 - \sigma_{\hat{\imath},u}(q)\right)$. In other words, this is the probability that q is smaller than the matched broker's fee plus the switching cost as well as smaller than all other $I - 1$ unmatched offers. Like in equation (1), the expected discounted continuation value of making a sale is also $\lambda\beta \int v_{i,m}(q)\, d\sigma_{i,m}(\tilde{q})$.

Recursive Nash Equilibrium. A recursive Nash equilibrium consists of a strategy profile $(\sigma_0, ..., \sigma_I)$ such that, for each $i \in \mathbb{I}$:

(i) There exists a unique value function v_i' associated with (σ_i', σ_{-i}), defined by (1) and (2), for any possible recursive strategy $\sigma_i' : Z \to \Delta$.

(ii) There is no deviation $\sigma_i' : Z \to \Delta$ such that $E_{\sigma_i'}\left[v_{i,z_i}'(q)\right] > E_{\sigma_i}\left[v_{i,z_i}(q)\right]$, where v_i' is the corresponding value function associated with (σ_i', σ_{-i}) and z_i describes player i's current status in $Z = \{m, u\}$.

It is important to stress that, unlike Duffie at al. (1994), the space of action Q need not be compact. As a consequence, the payoff function may not be well-defined for all strategies in Δ. This is why the payoff functions are part of our equilibrium concept. Propositions A1 and A2 in the Appendix show two alternative sets of conditions for the existence and uniqueness of a value function satisfying (1) and (2). When computing an equilibrium in the next section, we will focus on candidates satisfying the conditions presented in Proposition A2.

Remark 1. If $c = 0$ were allowed, this game would become a switching-cost version of the standard Bertrand model. In that case, all brokers setting $q = 0$ would be a pure-strategy equilibrium. Notice however that no pure strategy equilibrium exists when $c > 0$ and $q \in \mathbb{R}$.

3.1 Deriving a Symmetric Recursive Equilibrium

Let $Q = \mathbb{R}$ and define a symmetric strategy $\sigma_i = \sigma$ (identical for all i) as follows: (i) $\lim_{q \to \infty} (1 - \sigma_u(q))^{I-1} q < \infty$; and (ii) for each $z \in \{m, u\}$, σ_z is an atomless probability distribution such that $\sigma_z(a) = 0$, for some $a > -\infty$.

Given this strategy profile, the expected life-time discounted payoffs associated to each action $q \in [a, +\infty)$ of the matched (v_m) and unmatched (v_u) brokers are:

$$v_m(q) = (1 - \sigma_u(q - c))^I \left(q + \lambda\beta \int v_m(\tilde{q})\, d\sigma_m(\tilde{q}) \right), \tag{3}$$

and

$$v_u(q) = (1 - \sigma_m(q + c))(1 - \sigma_u(q))^{I-1} \left(q + \lambda\beta \int v_m(\tilde{q})\, d\sigma_m(\tilde{q}) \right). \tag{4}$$

It follows from Proposition A2 in the appendix that v_m and v_u are uniquely well-defined in $[a, +\infty)$. Notice that the term $(1 - \sigma_u(q - c))^I$ stands for the probability that the insuree accepts a fee q from the matched broker; $(1 - \sigma_m(q + c))(1 - \sigma_u(q))^{I-1}$ is the probability that the insuree accepts a fee q from a given unmatched broker; and $\lambda\beta \int v_m(\tilde{q})\, d\sigma_m(\tilde{q})$ represents the broker's expected discounted continuation value of making a sale.

Unmatched-Broker Strategy. Let us now explicitly derive the distribution σ_u. Set arbitrary values for $a \in \mathbb{R}$ and $\bar{v}_m > 0$, and define $\sigma_u(\cdot)$ such that:

$$(1 - \sigma_u(q - c))^I (q + \lambda\beta\bar{v}_m) = \bar{v}_m, \ \forall q \in [a + c, +\infty). \tag{5}$$

Solving for $\sigma_u(q)$:

$$\sigma_u(q) = 1 - \left[\frac{\bar{v}_m}{q + c + \lambda\beta\bar{v}_m} \right]^{1/I}, \ \forall q \geq a. \tag{6}$$

Notice that $\lim_{q\to\infty} \sigma_u(q) = 1$. Moreover, for any $a \in \mathbb{R}$, one has $\sigma_u(a) = 0$ if and only if:

$$\bar{v}_m = \frac{c+a}{(1-\lambda\beta)}. \tag{7}$$

One can then write:

$$\sigma_u(q) = \begin{cases} 0, & \forall q < a; \\ 1 - \left[\frac{c+a}{(1-\lambda\beta)q+c+\lambda\beta a}\right]^{1/I}, & \forall q \geq a. \end{cases} \tag{8}$$

Matched-Broker Strategy. Using an analogous reasoning, define $\bar{v}_u > 0$ and let σ_m be such that:

$$(1-\sigma_m(q+c))\left[\frac{\bar{v}_m}{q+c+\lambda\beta\bar{v}_m}\right]^{(I-1)/I}(q+\lambda\beta\bar{v}_m) = \bar{v}_u,\ \forall q \in [a,\infty). \tag{9}$$

Solving for $\sigma_m(q)$:

$$\sigma_m(q) = 1 - \frac{\bar{v}_u}{q-c+\lambda\beta\bar{v}_m}\left[\frac{q+\lambda\beta\bar{v}_m}{\bar{v}_m}\right]^{(I-1)/I},\ \forall q \geq a+c. \tag{10}$$

Notice that $\lim_{q\to\infty} \sigma_m(q) = 1$, for every $I \geq 1$. Moreover, from (7), one has $\left[\frac{a+c+\lambda\beta\bar{v}_m}{\bar{v}_m}\right] = 1$. Therefore, $\sigma_m(a+c) = 0$ if and only if:

$$\bar{v}_u = a + \lambda\beta\bar{v}_m = \frac{\lambda\beta c + a}{(1-\lambda\beta)}. \tag{11}$$

Therefore, σ_m is written as follows:

$$\sigma_m(q) = \begin{cases} 0, & \forall q < a+c; \\ 1 - \frac{(\lambda\beta c+a)\left[\lambda\beta+\frac{1-\lambda\beta}{c+a}q\right]^{(I-1)/I}}{(1-\lambda\beta)(q-c)+\lambda\beta(c+a)}, & \forall q \geq a+c. \end{cases} \tag{12}$$

Symmetric Equilibrium. We always have $\bar{v}_m > \bar{v}_u$, since $c > 0$. Moreover, $\bar{v}_u > 0$ if and only if:

$$a > -\lambda\beta c. \tag{13}$$

We also have by construction that:

$$v_m(q) \leq \bar{v}_m,\ \forall q < a+c; \tag{14}$$

$$v_m(q) = \bar{v}_m,\ \forall q \geq a+c; \tag{15}$$

and

$$v_u(q) \leq \bar{v}_u,\ \forall q < a; \tag{16}$$

$$v_u(q) = \bar{v}_u,\ \forall q \geq a. \tag{17}$$

Given the unmatched brokers' strategies, the function v_m is uniquely defined and leaves the matched broker indifferent to any action in the support of σ_m, namely $(a + c, \infty)$. Moreover, given that all other players draw fees from σ_m and σ_u, the function v_u is also uniquely defined and leaves any unmatched broker indifferent to any action in $\operatorname{supp} \sigma_u = (a, \infty)$.

Therefore, the payoff functions that are uniquely defined by equations (3) and (4) joint with the strategies defined by (8) and (12) describe a recursive equilibrium for any given $(I, \lambda\beta, c, a)$ satisfying the positive-profit condition (13). To see this, suppose by contradiction that they are not a recursive equilibrium. In that case, there should exist a strategy outside $\operatorname{supp} \sigma_z$ that yields an expected payoff higher than $\bar{v}_{i,z}$ for some $z \in Z = \{m, u\}$, which is not possible given conditions (14) through (17).

Comments on the Symmetric Equilibrium. The parameters of the model are $(I, \lambda\beta, c)$. For each parameter value and for any $a > -\lambda\beta c$, the payoffs (v_m, v_u) and strategies (σ_m, σ_u) defined by (3), (4), (8), and (12) characterize a symmetric recursive equilibrium. Therefore, for any vector of parameters $(I, \lambda\beta, c)$, there exists a continuum of equilibria indexed by $a > -\lambda\beta c$.

It follows from (7) that $a + c = \bar{v}_m (1 - \lambda\beta) > 0$, and then $\sigma_m(0) = 0$. Therefore, only unmatched brokers set negative fees. This occurs when $a < 0$. Moreover, equations (7) and (11) define the equilibrium payoffs for matched and unmatched brokers. Since there is a continuum of $a > -\lambda\beta c$ that are consistent with a recursive equilibrium, there is also a continuum of possibilities for the equilibrium profits.

This feature of the model is related to the fact that we have imposed no reservation price for consumers. Like in static Bertrand models (see Baye and Morgan, 1999), the absence a reserve leads to this type of Folk Theorem in which any finite positive payoff can be achieved in equilibrium. From a Poperian point of view, this is very inconvenient since the model is too flexible to accommodate a vast set of empirical observations. However, from a statistical perspective, this flexibility is welcome and is used to estimate the parameters of the model in conjunction with a particular equilibrium (indexed by a) which most likely could generate the observed data.

4 Data-Generating Distribution

The data display fees $y_j \in \mathbb{R}_+$ that have been paid by the insurees rather than the offers generated by σ_m and σ_u. That is, the data generation distribution combines the mixed-strategies equilibrium σ_m and σ_u with the insurees' selection rules.

Each insuree in the sample accesses one offer q_0 drawn from σ_m and I independent offers drawn from σ_u. They accept the offer q_0 from their matched broker whenever $q_0 \leq \min(q_1, ..., q_I) + c$. Moreover, they prefer the offer $q_{\min} = \min(q_1, ..., q_I)$ whenever $q_0 > \min(q_1, ..., q_I) + c$. Therefore, each observation j in the sample is generated as follows:

$$q_0 \leq q_{\min} + c \Rightarrow y_j = q_0; \tag{18}$$

$$q_0 > q_{\min} + c \Rightarrow y_j = \max[0, q_{\min}]. \tag{19}$$

The distribution function for $q_{\min} = \min(q_1, ..., q_I)$ is given by:

$$H(q_{\min}) = 1 - (1 - \sigma_u(q_{\min}))^I. \tag{20}$$

Let $\Pr(\cdot)$ be a conditional probability measure—defined on some arbitrary measurable space $(\Omega, \mathcal{F})$—that represents the data generating distribution. Notice from (18) to (19) that, for any $y \geq 0$, one has:

$$\Pr[y_j \in (a, y)] = \Pr[A \cup B], \tag{21}$$

where $A = \{\omega \in \Omega : q_0 \leq q_{\min} + c \text{ and } q_0 \leq y\}$ and $B = \{\omega \in \Omega : q_0 > q_{\min} + c$ and $q_{\min} \leq y\}$. Since A and B are disjoint sets, one obtains:

$$\Pr[y_j \in (a, y)] = \int_a^y (1 - \sigma_m(\tilde{y} + c))\, dH(\tilde{y}), \text{ for } y < a + c; \tag{22}$$

and

$$\Pr[y_j \in (a, y)] = \int_a^y (1 - \sigma_m(\tilde{y} + c))\, dH(\tilde{y}) + \int_a^y (1 - H(\tilde{y} - c))\, dF_m(\tilde{y}), \text{ for } y \geq a + c. \tag{23}$$

We have interpreted negative fees as being nonpecuniary gifts and services that are offered to consumers and not observed by the econometrician. Thus, a mass of zero-brokerage fees is generated here as a censoring of the equilibrium strategy σ_u at the point $q = 0$. The probability of observing a zero fee is given by:

$$\Pr[y_j = 0] = \int_a^0 (1 - \sigma_m(\tilde{y} + c))\, dH(\tilde{y}). \tag{24}$$

Moreover, for $y > 0$, we use the Leibniz rule to find:

$$\Pr\left[y_j \in (0, y)\right] = \int_0^y p(\tilde{y})\, d\tilde{y}, \tag{25}$$

where the density function $p(y)$ is given by:

$$p(y) = (1 - \sigma_m (y + c)) \frac{\partial H(y)}{\partial y}, \text{ for } 0 < y < a + c; \tag{26}$$

and

$$p(y) = (1 - \sigma_m (y + c)) \frac{\partial H(y)}{\partial y} + (1 - H(y - c)) \frac{\partial \sigma_m (y)}{\partial y}, \text{ for } y \geq a + c. \tag{27}$$

5 Maximum Likelihood Estimation

We use the maximum likelihood procedure for censored data to estimate our model (see Meeker and Escobar, 1994). Let $\mathbb{J} = \{1, ..., J\}$ the sample-index set, and define $\mathbb{J}_0 = \{j \in \mathbb{J} : y_j = 0\}$. Thus the likelihood function is given by:

$$L(I, \lambda\beta, c, a) = \prod_{j \in \mathbb{J}_0} \Pr\left[y_j = 0\right] \prod_{j \in \mathbb{J} \backslash \mathbb{J}_0} p(y_j). \tag{28}$$

We define the following grids for the structural parameters: $I \in \{1, 2, ..., 20\}$; $\lambda\beta \in \{0.01, 0.02, ..., 0.99\}$; $c \in \{0.01, 0.02, ..., 500\}$; and $a = -\alpha\lambda\beta c$, where $\alpha \in \{0.01, 0.02, ..., 0.99\}$. We then use the *Wolfram Mathematica 7* software to compute and select the highest value of (28), given our data on brokerage fees, for all possible parameter combinations.[4] *Mathematica* is a powerful tool which is necessary here because there is no closed formula for the integral in (24) that defines $\Pr\left[y_j = 0\right]$. This is solved by means of generalized hypergeometric functions.

Table 3 presents estimation results. Column (a) displays the estimates for the unrestricted model. Results suggest that: (i) insurees typically access two offers before purchasing an insurance policy; (ii) they face a switching cost equivalent to 142 BRL (i.e., about US$ 71 in purchasing power parity); and (iii) the brokers' expected lifetime discounted profit is about 269 BRL when trading to a matched consumer, and about 127 BRL when trading to an unmatched consumer. The estimated value for $\lambda\beta$ is 0.61, but the model does not allow us to detach the discount factor from the probability λ.

Columns (b) and (c) present restricted estimations for $\lambda\beta = 0.90$ and for $I = 2$. Log-likelihood ratio tests indicate that those two restricted models are statistically different from the unrestricted model. Interestingly, the switching cost and brokerage

[4] Statistical results in other parts of this paper were computed in *STATA 10*.

profitability do not vary much across those different models. The obtained values are economically reasonable for the city of Sao Paulo.

[Table 3]

Next, Figure 3 graphs the density function for the parameters of the unrestricted estimation against the empirical histogram. For scale reasons, we graph only positive brokerage fees. The empirical and estimated mass at $q = 0$ are, respectively, 15.73% and 10.73%. Notice that the density curve combines two hyperbolic functions obtained from σ_m and H.

[Figure 3]

5.1 Econometric Comments

The likelihood function is not continuous when the parameters $\lambda\beta$, c, and a take arbitrary real values. However, when the parameter vector $(I, \lambda\beta, c, a)$ is evaluated in a finite set, then the likelihood function is continuous. Any function with a finite domain is continuous. In this case, standard results from M-estimators guarantee consistency and asymptotic normality of our maximum-likelihood estimator. We do not explore these issues in further detail in this paper, we simply want to emphasize that consistency and asymptotic log-likelihood tests in this paper depend on the existence of an economic justification for the assumption that the model parameters are valued in a finite set.

A second important econometric issue is related to the identification of a single maximum value in the estimation process. We have computed likelihood values for all possible combinations of the parameters. Table 4 displays the parameters associated with the ten highest likelihood values. The likelihood values decrease monotonically with little variation in the parameter values.

Next, Figures 4-7 present the likelihood values for different $(I, \lambda\beta, c, a)$. Each figure freezes three parameters at their optimal point and varies the fourth parameter around its optimal value. These figures illustrate the marginal effect of each parameter over the likelihood function around the optimal value.

[Table 4 & Figures 4-7]

5.2 Subsample with a Single Vehicle Model

We now perform the analysis from the last section on a subsample that refines our baseline sample to include only the most popular vehicle model in Brazil (Gol 1.0L).

The results are similar to the previous one, as can be seen in Table 5 and Figures 9-13. They confirm that potential heterogeneity across insurees is not the main element behind the large variation in brokerage fees as observed in the data.

[Table 5 & Figures 8-12]

6 Conclusion

Figure 3 summarizes most of our work. We start the analysis by noticing that data on brokerage fees exhibited wide dispersion which was not strongly correlated to observables. We propose a theory that accounts for the key economic forces behind the brokerage problem, then take that model to the data.

Our theory suggests that matched and unmatched brokers set fees randomly. Unmatched fees start at $a = -37.26$ and appear uncontested by matched offers up to the point $a + c = 104.2$. After that point, offers from matched and unmatched brokers appear together in the data, generating a second peak in the density function. The parameter a is directly related to the mass of policies with zero fees, while c is related to the second peak of the density function.

Based on this, we have been able to infer important structural parameters such as individual costs when switching to a new (unmatched) broker, and the brokers' profitability when dealing with old (matched) and new (unmatched) customers. Our dynamic model with switching costs replicates stylized facts that the existing models of price competition cannot generate.

A Appendix: Uniquely Well-Defined Value Functions

In this appendix, we present two alternative sets of conditions that guarantee that equations (1) and (2) define a unique value function v_i. We first consider the case where Q is a compact interval of the real line. Next, we analyze the case where $Q = \mathbb{R}$ and players' strategies are well behaved when q approaches $-\infty$ and $+\infty$.

Proposition A1. Fix $Q = [a, b] \subset \mathbb{R}$. There exists a unique measurable bounded function $v_{i,m} : [a, b] \to \mathbb{R}$ that satisfies condition (1), for any strategy profile $(\sigma_1, ..., \sigma_I)$.

Proof. Let $\mathcal{F}$ be the complete metric space of $\sigma_{i,m}$-measurable bounded functions mapping $[a, b]$ into $[a, b]$. Let us endow this space with sup norm and define $T : \mathcal{F} \to \mathcal{F}$

as follows:

$$Tv_{i,m}(q) = \prod_{\hat{\imath}\in\mathbb{I}\setminus\{i\}} (1 - \sigma_{\hat{\imath},u}(q-c)) \left(q + \lambda\beta \int v_{i,m}(\tilde{q})\, d\sigma_{i,m}(\tilde{q})\right). \tag{29}$$

Being that $q \in [a,b]$, $Tv_{i,m}$ defines a bounded function for any bounded $v_{i,m}$. Since $\lambda\beta < 1$, one has:

$$\sup_{q\in[a,b]} \left(\mid Tv'_{i,m} - Tv''_{i,m} \mid\right) < \lambda\beta \sup_{q\in[a,b]} \left(\mid v'_{i,m} - v''_{i,m} \mid\right), \ \forall \left(v'_{i,m}, v''_{i,m}\right) \in F \times F. \tag{30}$$

The result follows then from the contraction mapping theorem. ■

Notice that, when $Q = [a,b]$, the function $v_{i,m} : [a,b] \to \mathbb{R}$ is uniquely defined for each strategy profile $(\sigma_0, ..., \sigma_I)$. When extending the domain to $Q = (-\infty, +\infty)$, restrictions on strategies are necessary in order to have a well-defined $v_{i,m}$. An important case used in this paper is when a broker i faces strategies for which there exists $a > -\infty$ such that $\sigma_{\hat{\imath},z}(a) = 0$ and $\lim_{q\to+\infty} (1 - \sigma_{\hat{\imath},z}(q))\, q < +\infty$, for any $\hat{\imath} \neq i$ and $z \in \{m,u\}$.

Proposition A2. Let $q \in Q = \mathbb{R}$ and assume that for each player $\hat{\imath} \neq i$ and each state $z \in \{m,u\}$: (i) there exists $a > -\infty$ such that $\sigma_{\hat{\imath},z}(a) = 0$; and (ii) $\lim_{q\to+\infty} \prod_{\hat{\imath}\in\mathbb{I}\setminus\{i\}} (1 - \sigma_{\hat{\imath},u}(q-c))\, q < +\infty$. Then, any action $q_i < a$ is strictly dominated by $q_i = a$. Moreover, there exists a unique measurable bounded function $v_{i,m} : [a, +\infty) \to \mathbb{R}$ that satisfies condition (1), for any strategy profile $\sigma = (\sigma_1, ..., \sigma_I)$.

Proof. Any action $q_i < a$ is strictly dominated by $q_i = a$, since $q_{\hat{\imath}} > a$ almost surely, for all $\hat{\imath} \neq i$. Then, equation (1) defines a unique bounded function $v_{i,m} : [a, +\infty) \to \mathbb{R}$. The proof for that is analogous to the proof of Proposition 1. Define the operator T as in (29) and notice that T is a contraction map of modulus $\lambda\beta < 1$ that maps the space of $\sigma_{i,m}$-measurable bounded functions on $\mathbb{R}$, into itself. Notice that $Tv_{i,m}$ is $\sigma_{i,m}$-measurable and bounded for any $\sigma_{i,m}$-measurable bounded $v_{i,m}$, since $\lim_{q\to+\infty} \prod_{\hat{\imath}\in\mathbb{I}\setminus\{i\}} (1 - \sigma_{\hat{\imath},u}(q-c))\, q < +\infty$. Thus the result once again follows from the contraction mapping theorem. ■

To conclude this appendix, note that for each given $v_{i,m}(\cdot)$, the function $v_{i,u}(\cdot)$ is also uniquely defined by (2).

References

[1] Baye, Michael R. and John Morgan (1999). A folk theorem for one-shot Bertrand games. *Economics Letters*, 65 (1), 59–65.

[2] Braido, Luis H.B. (2009). Multiproduct price competition with heterogeneous consumers and nonconvex costs. *Journal of Mathematical Economics*, Forthcoming.

[3] Duffie, Darrell, John D. Geanakoplos, Andreu Mas-Colell, and Andrew M. McLennan (1994). Stationary Markov equilibria. *Econometrica* 62 (4), 745-781.

[4] Meeker, William Q., and Luis A. Escobar (1994). Maximum likelihood methods for fitting parametric statistical models to censored and truncated data. In Probabilistic and Statistical Methods in the Physical Sciences, J. Stanford and S. Vardeman (eds.), New York: Academic Press.

[5] Sharkey, Willian W. and Sibley, David S. (1993). A Bertrand model of pricing and entry. *Economics Letters*, 41 (2), 199-206.

[6] Stahl, Dale O. (1989). Olipolistic pricing with sequential consumer search. *American Economic Review*, 79 (4), 700-712.

[7] Varian, Hal R. (1980). A model of sales. *American Economic Review*, 70 (4), 651-659.

Table 1. SUSEP/AUTOSEG Auto-Insurance Data

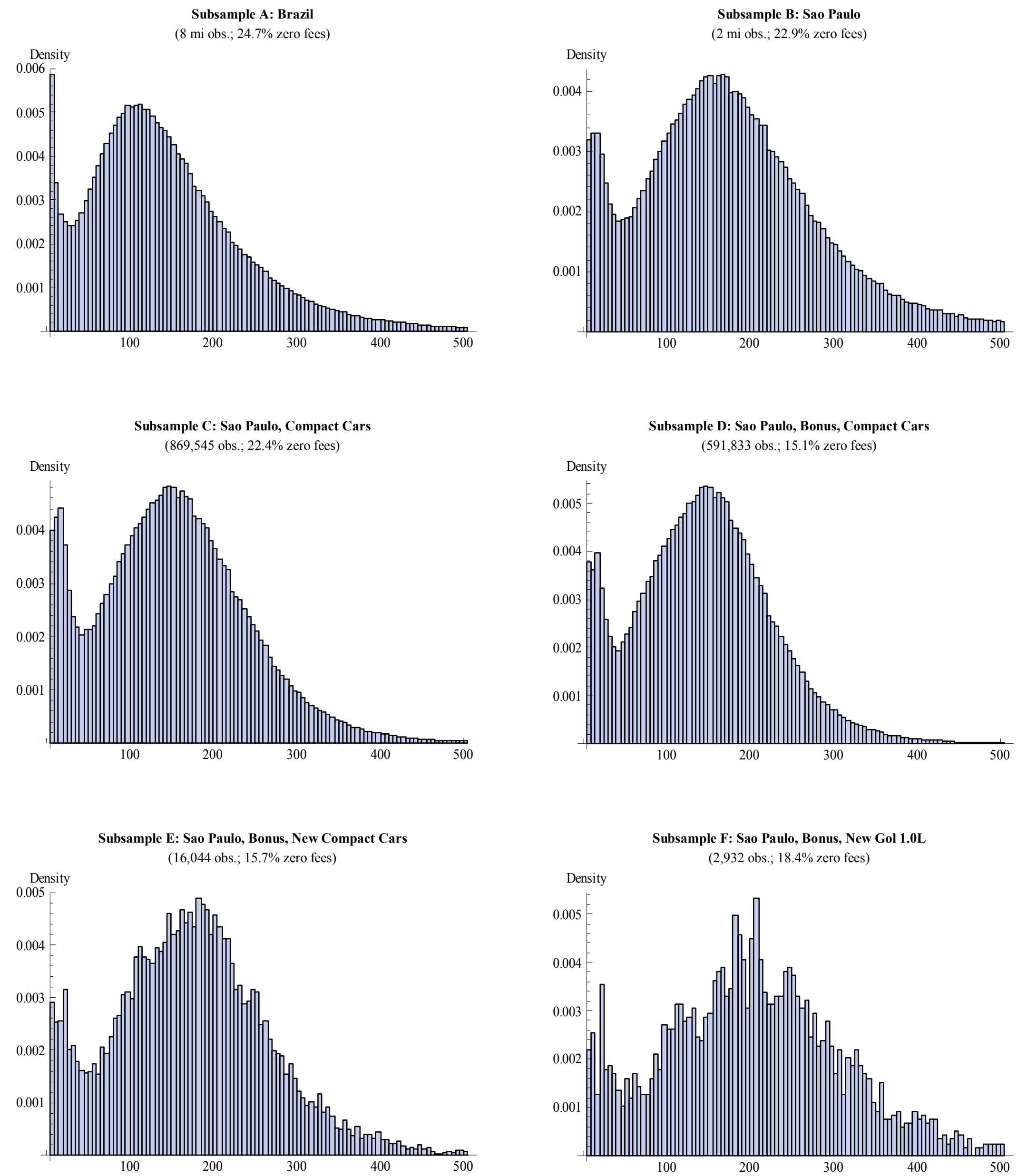
Variable	Description
Total Insurance Premium[(a)]	Total premium paid by the insuree to the insurance company (in BRL).
Broker's Percentage Fee (PERC_CORR)	Fraction of the total premium that is transferred from the insurance company to the broker.
Brokerage Fee	Product of the Broker's Percentage Fee and the Total Insurance Premium.
Vehicle Type (COD_MODELO)	Alpha-numerical variable describing the make, model, and power train of the insured vehicle. For the main estimation we restrict attention to nine models of compact cars with 1.0L (1,000cc) engines (namely, Celta, Clio, Corsa, Fiesta, Gol, Ka, Palio, Peugeot 206, and Uno).
Vehicle Model Year (ANO_MODELO)	Year of the vehicle model. When performing the estimations, we restrict attention to new vehicles (model year 2003).
Geographic Region (REGIAO)	Numerical variable ranging from 1 to 41 that describes the geographic region of the insured policy. For our estimations, we restrict the sample to the metropolitan region of Sao Paulo (11).
Coverage Type (COBERTURA)	Qualitative variable describing the type of insurance coverage for first-party vehicular damages. Code 1 is used for comprehensive coverage; 2 for coverage against fire and theft; 3 for fire only; 4 for write-off, collision, and theft; 9 for others. We restrict attention to contracts with comprehensive coverage (1).
Deductible Type (TIPO_FRANQ)	Qualitative variable describing the level of deductible for physical damages to the insured vehicle. Codes 1, 2 and 3 indicate low, regular, and high deductible levels, respectively. Code 9 indicates no deductible.
Physical Damage Coverage (IS_CASCO and IS_RCDM)	Maximum insurance coverage for first-party (IS_CASCO) and third-party (IS_RCDM) vehicular damages. Values are expressed in BRL.
Personal Injury Coverage (IS_APP and IS_RCDP)	Maximum insurance covered (in BRL) for personal injuries caused to passengers of the insured vehicle (IS_APP) and to passengers of third-party vehicles and pedestrians (IS_RCDP).[(b)]
Bonus Discount (PERC_BONUS)	Percentage discount over the premium for insurance against physical damages to the insured vehicle.[(c)] For the estimations, we restrict attention to contracts with positive bonus discount.
Gender (SEXO)	Qualitative variable describing the gender of the vehicle's primary driver, as declared in the contract. The code "M" is used for male; "F" for female; and "J" for corporate entities. We exclude policies code "J".
Date of Birth (DATA_NASC)	Primary driver's date of birth as stated in the contract. The value "00000000" is used for fleet vehicles (code "J").
Zip Code (CEP)	Zip code pertaining to the primary driver's address as stated in the contract. We use the first five digits of this eight-digit code.

(a) The sum of the following variables: PRE_CASCO, PRE_RCDM, PRE_APP, PRE_RCDP, and PRE_OUTROS.

(b) The coverage limits include medical expenses and death/disability benefits that extend the minimum coverage.

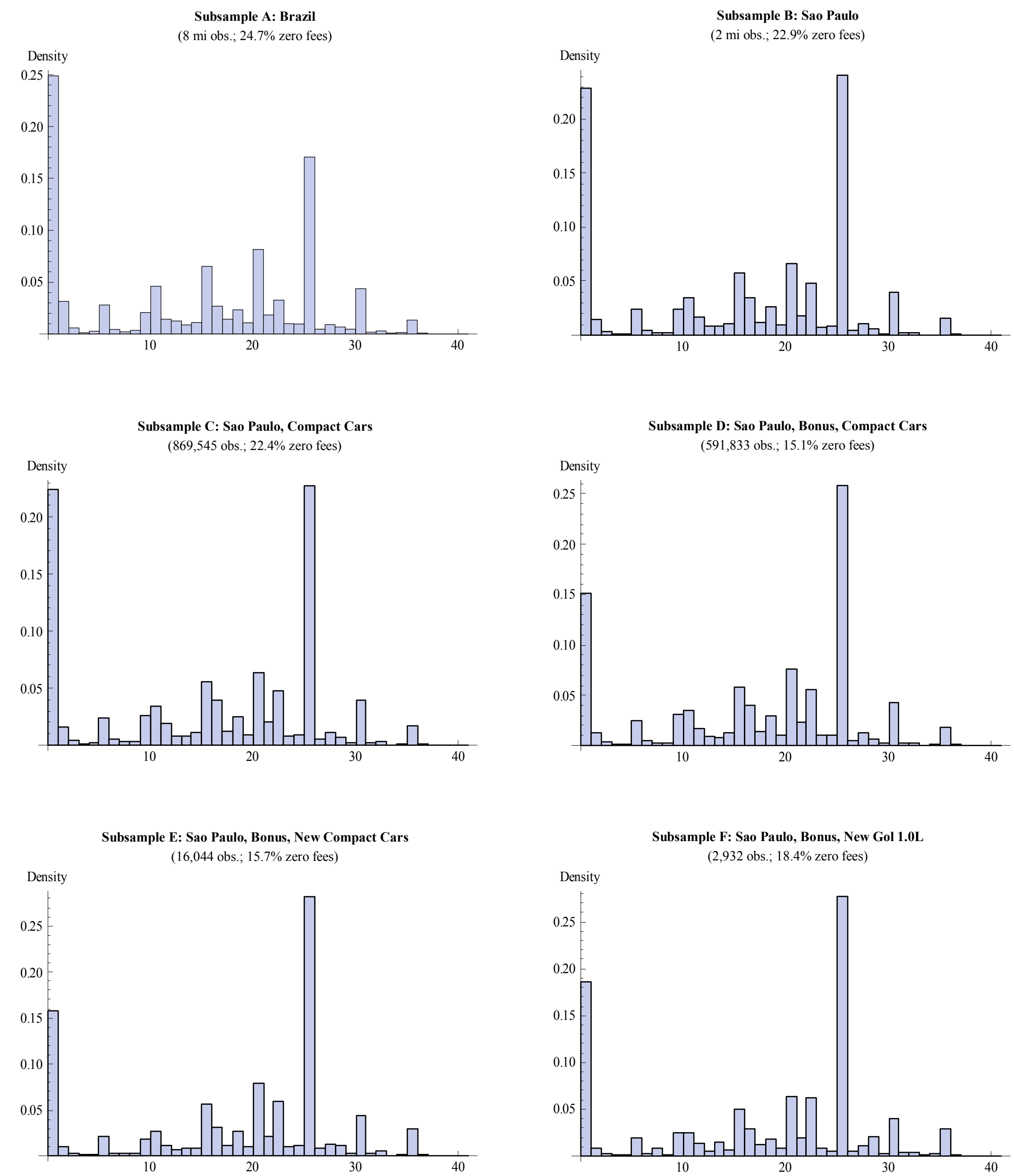
(c) This bonus depends on the insuree's previous claims record which is publicly accessible by all companies. Each agent starts with no bonus and becomes eligible for a 10 percent discount after one year without a claim. The discount bonus is increased by 5 percentage points after each additional year without a claim, peaking at around 40 percent, depending on the insurance company. In the case of a claim, the insuree's rating is reduced by one level. In case of multiple claims in a same year, the classification might be reduced by more than one level, depending on the insurance company.

Figure 1: Positive Brokerage Fees (in BRL)



For scale reasons, we only graph positive fees. Histograms' Width: 5 BRL.

Figure 2: Broker's Percentage Fees



Zero fees included. Histograms' Width: 1 Percentage Point.

Table 2. Reduced-Form OLS Regressions

	Zero-Brokerage Dummy	**Positive Fees (in BRL)**	**Positive Percentage Fees**
Primary Driver Gender (male = 1)	.0140	4.99*	-.5592*
Zero-Deductible Dummy	.1711	-135.72*	-11.7380*
Low-Deductible Dummy	.0018	20.62*	-.5931*
High-Deductible Dummy	.0371	-24.87	-1.1330
Maximum Coverage for Damages to the Insured Vehicle (BRL)	-2.16e-06	.0049*	9.93e-05*
Maximum Coverage for Damages to Third-Part Vehicles (BRL)	-2.45e-06*	.0010*	8.96e-05*
Maximum Coverage for Personal Injury to Passengers of Third-Part Vehicles and Pedestrians (BRL)	8.17e-07*	-.0002*	-2.75e-05*
Maximum Coverage for Personal Injury to Passengers of Insured Vehicle (BRL)	1.31e-07	.0001*	-1.57e-05*
Zip Code Fixed Effects	828 dummies	802 dummies	802 dummies
Vehicle Model Year Fixed Effects	8 dummies	8 dummies	8 dummies
Primary Driver Age Fixed Effects	72 dummies	72 dummies	72 dummies
Bonus-Category Fixed Effects	16 dummies	16 dummies	16 dummies
Sample Size	*16,044*	*13,521*	*13,521*
R^2	*9.67%*	*23.38%*	*16.96%*

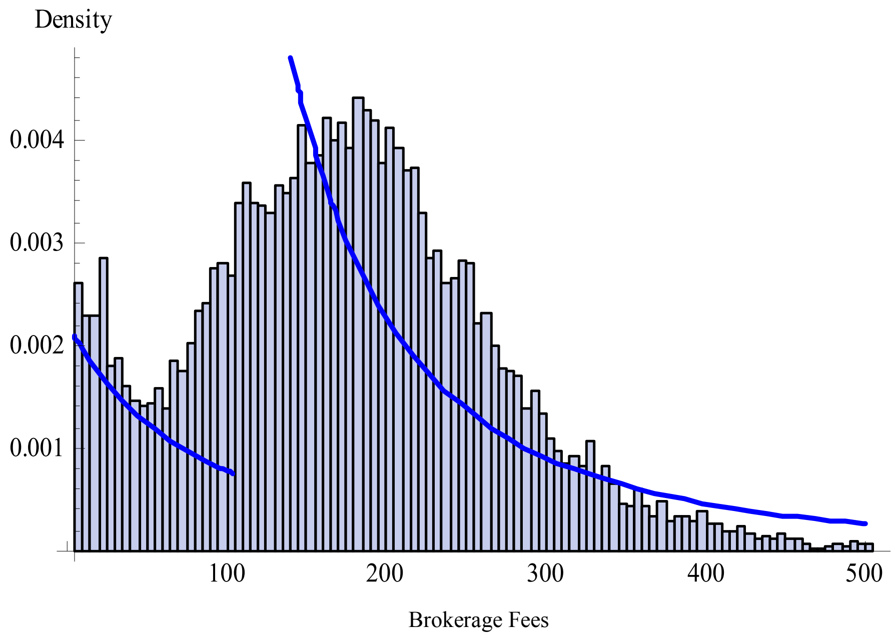
* Significant at 1%; Robust standard errors; All regressions include a constant term.

Table 3. Structural Maximum-Likelihood Estimation

Structural Parameters	Unrestricted Estimation (a)	Restricted Model: λβ = 0.90 (b)	Restricted Model: *I* = 2 (c)
Adjusted Discount Factor (λβ)	0.61	0.90	0.57
Number of Unmatched Offers (*I*)	1	1	2
Switching Cost (*c*), in BRL	142.06	92.18	142.04
Highest Nonpecuniary Gift (-*a*), in BRL	37.26	59.73	37.24
Sample Size	*16,044*	*16,044*	*16,044*
Likelihood	$1.4\times10^{-39,026}$	$1.7\times10^{-39,900}$	$1.0\times10^{-39,181}$
Log-Likelihood Ratio		*4,024.48**	*714.36**
Profitability			
Expected Life-Time Discounted Profit per Matched Customer ($\bar{v}_m$), in BRL	268.71	324.47	243.71
Expected Life-Time Discounted Profit per Unmatched Customer ($\bar{v}_u$), in BRL	126.65	232.29	101.67

* The 1% and 10% critical values for the log-likelihood ratio test are 6.64 and 2.71, respectively. Thus, the two restricted models are statistically different from the unrestricted model.

Figure 3. Empirical and Estimated Density Functions



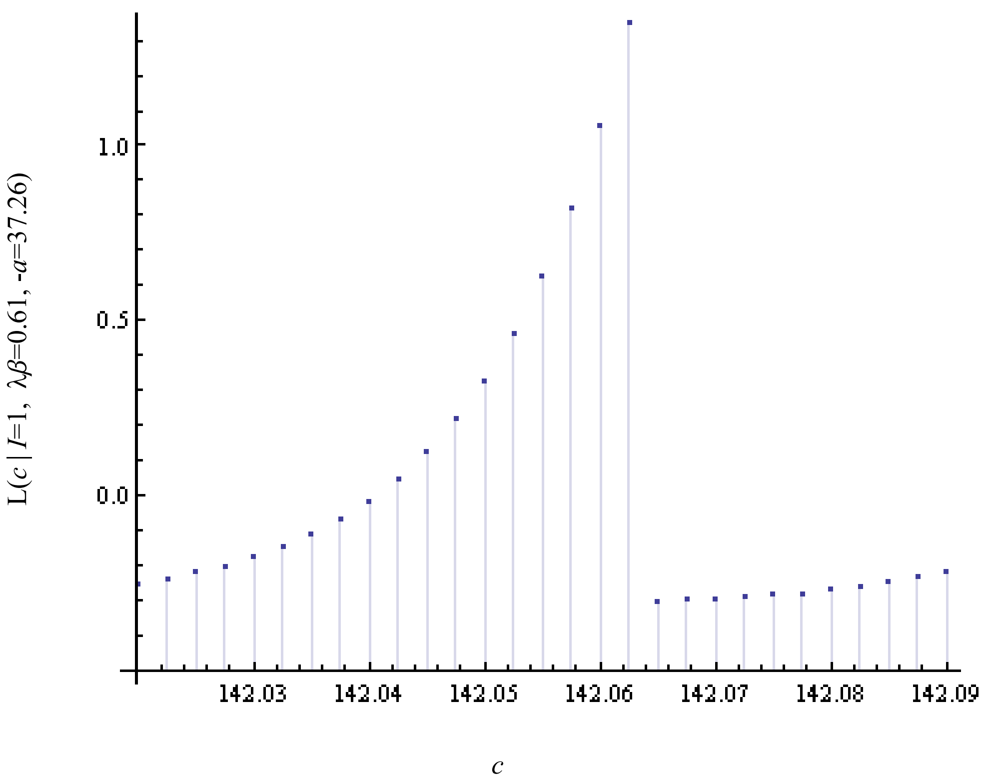
Vertical bars for the empirical histogram and the line for the estimated density function.
For scale reasons, we only graph positive brokerage fees. Histogram Width: 5 BRL.
Sample Size: 16,044; Empirical Fraction of Zero Fees: 15.73%; Estimated Fraction of Zero Fees: 10.73%.

Table 4. Global Maximum: Top-10 Likelihood Values

Likelihood ($\times 10^{-39,026}$)	***I***	**$\lambda\beta$**	***c***	***-a***
1.3803	1	0.61	142.06	37.26
0.6470	1	0.61	142.05	37.26
0.3033	1	0.61	142.04	37.26
0.1421	1	0.61	142.03	37.25
0.1008	1	0.61	142.09	37.27
0.0666	1	0.61	142.02	37.25
0.0473	1	0.61	142.08	37.27
0.0312	1	0.61	142.01	37.25
0.0221	1	0.61	142.07	37.26
0.0205	1	0.61	141.94	37.23

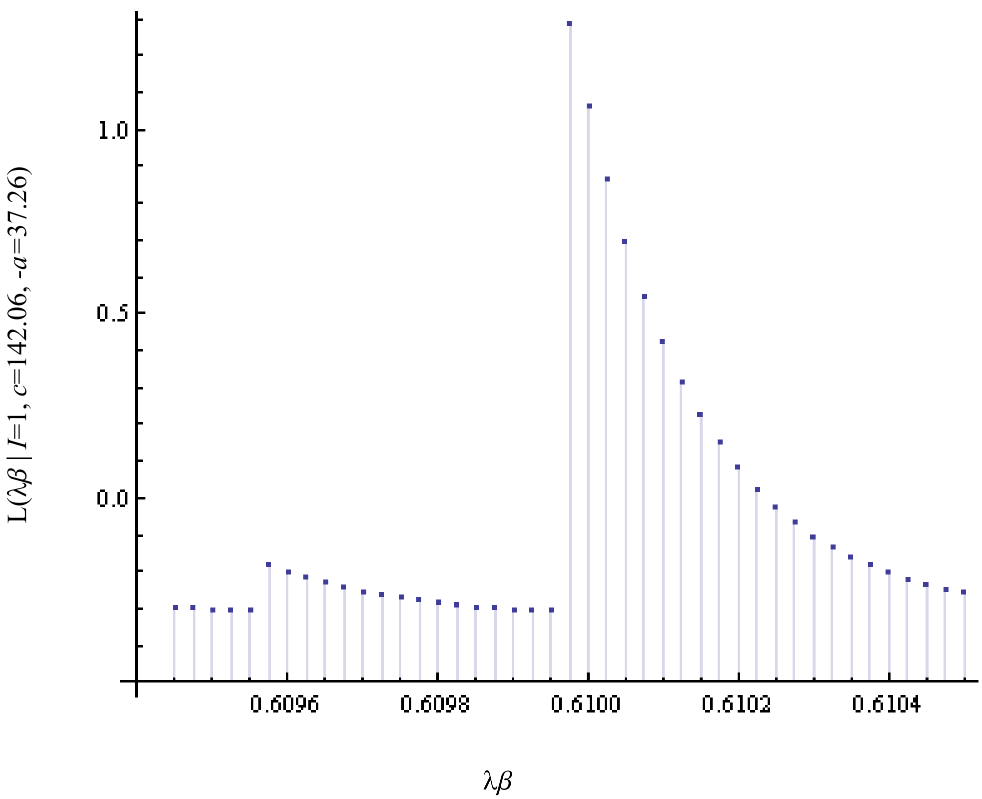
Sample Size: 16,044

Figure 4. Local Maximum: Switching Cost (c)



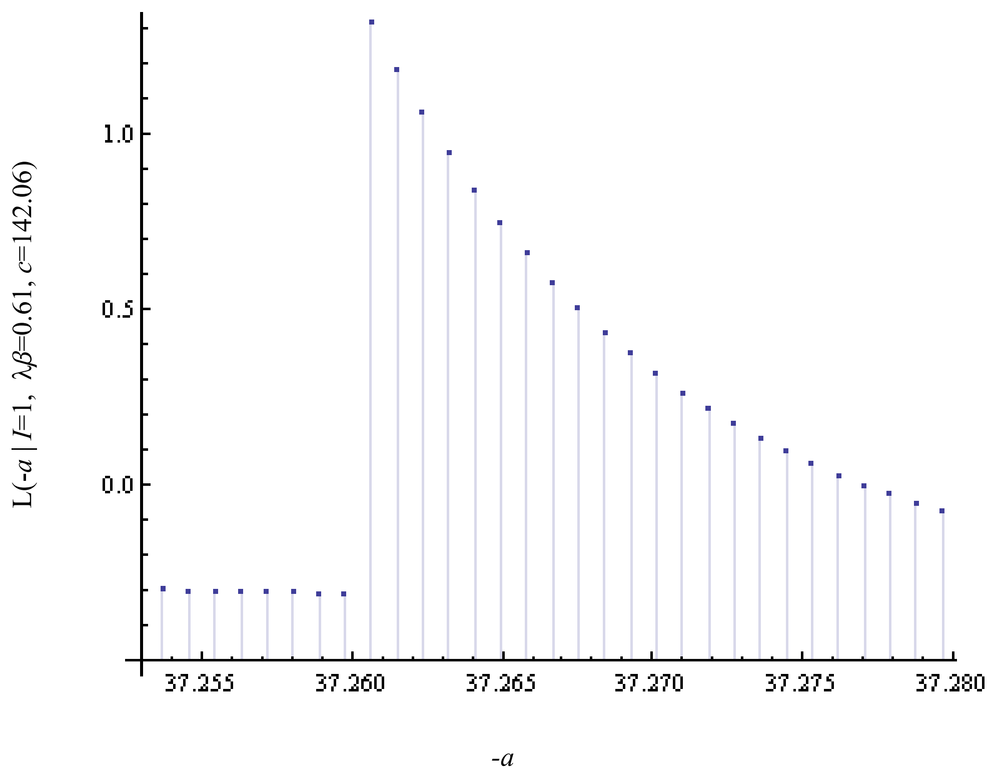
Likelihood ($\times 10^{-39,026}$) for different switching costs when $I = 1$, $\lambda\beta = 0.61$, and $-a = 37.26$.
Global Maximum: $1.3803\times 10^{-39,026}$; Sample Size: 16,044.

Figure 5. Local Maximum: Adjusted-Discount Factors ($\lambda\beta$)



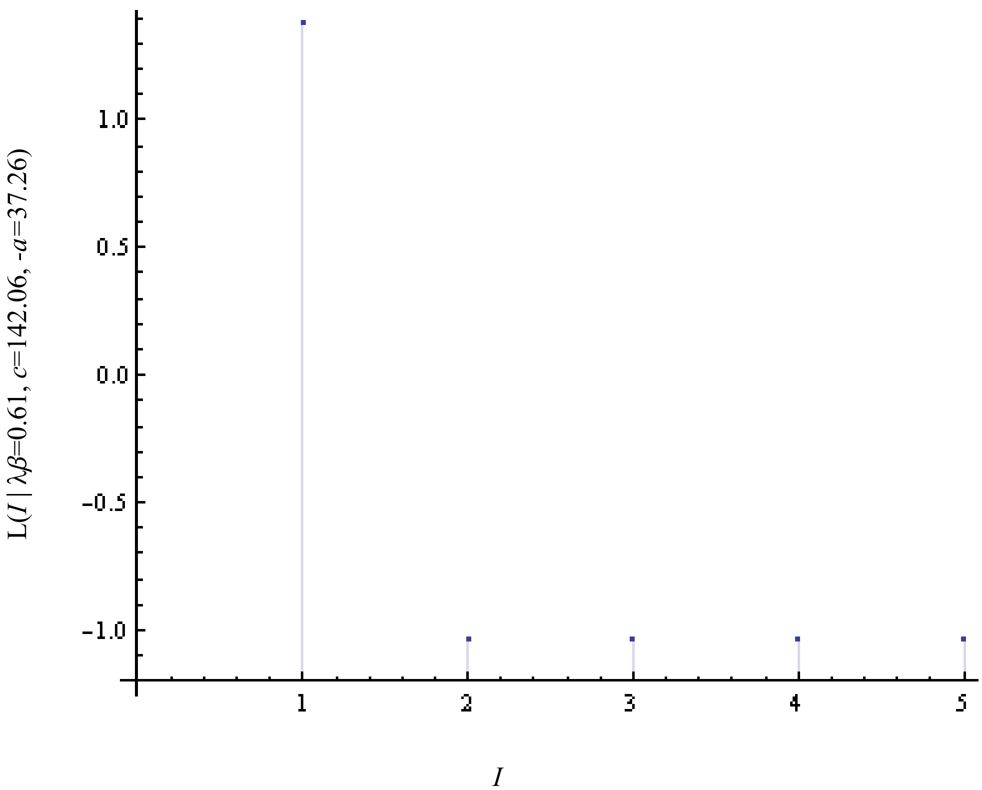
Likelihood ($\times 10^{-39,026}$) for different adjusted-discount factors when I = 1, c = 142.06 and $-a$ = 37.26.
Global Maximum: $1.3803 \times 10^{-39,026}$; Sample Size: 16,044.

Figure 6. Local Maximum: Highest Gift (-*a*)



Likelihood ($\times 10^{-39,026}$) for different values of -*a* when *I*=1, $\lambda\beta$=0.61, and *c*=142.06.
Global Maximum: $1.3803 \times 10^{-39,026}$; Sample Size: 16,044.

Figure 7. Local Maximum: Unmatched Offers (*I*)



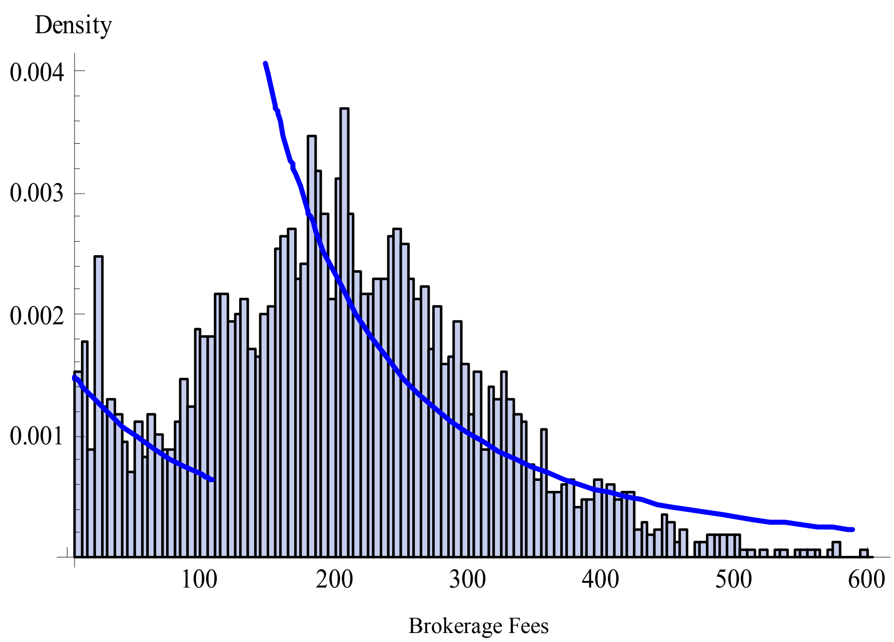
Likelihood ($\times 10^{-39{,}026}$) for different values of *I* when $\lambda\beta$=0.61, *c*=142.06, and $-a$ =37.26.
Global Maximum: $1.3803\times 10^{-39{,}026}$; Sample Size: 16,044.

Table 5. Structural Estimation: Single-Model Subsample

Structural Parameters	Unrestricted Estimation (a)	Restricted Model: $\lambda\beta = 0.90$ (b)	Restricted Model: I = 2 (c)
Adjusted Discount Factor ($\lambda\beta$)	0.67	0.90	0.63
Number of Unmatched Offers (*I*)	1	1	2
Switching Cost (*c*), in BRL	170.40	144.03	168.00
Highest Nonpecuniary Gift (-*a*), in BRL	61.65	102.41	59.27
Sample Size	*2,932*	*2,932*	*2,932*
Likelihood	$2.4\times10^{-7,179}$	$4.5\times10^{-7,324}$	$1.7\times10^{-7,211}$
Log-Likelihood Ratio		*666.47**	*148.10**
Profitability			
Expected Life-Time Discounted Profit per Matched Customer ($\overline{v}_m$), in BRL	329.54	416.25	293.86
Expected Life-Time Discounted Profit per Unmatched Customer ($\overline{v}_u$), in BRL	159.14	272.22	125.86

* The 1% and 10% critical values for the log-likelihood ratio test are, respectively, 6.64 and 2.71. Thus, the two restricted models are statistically different from the unrestricted model.

Figure 8. Empirical and Estimated Density Functions: Single-Model Subsample



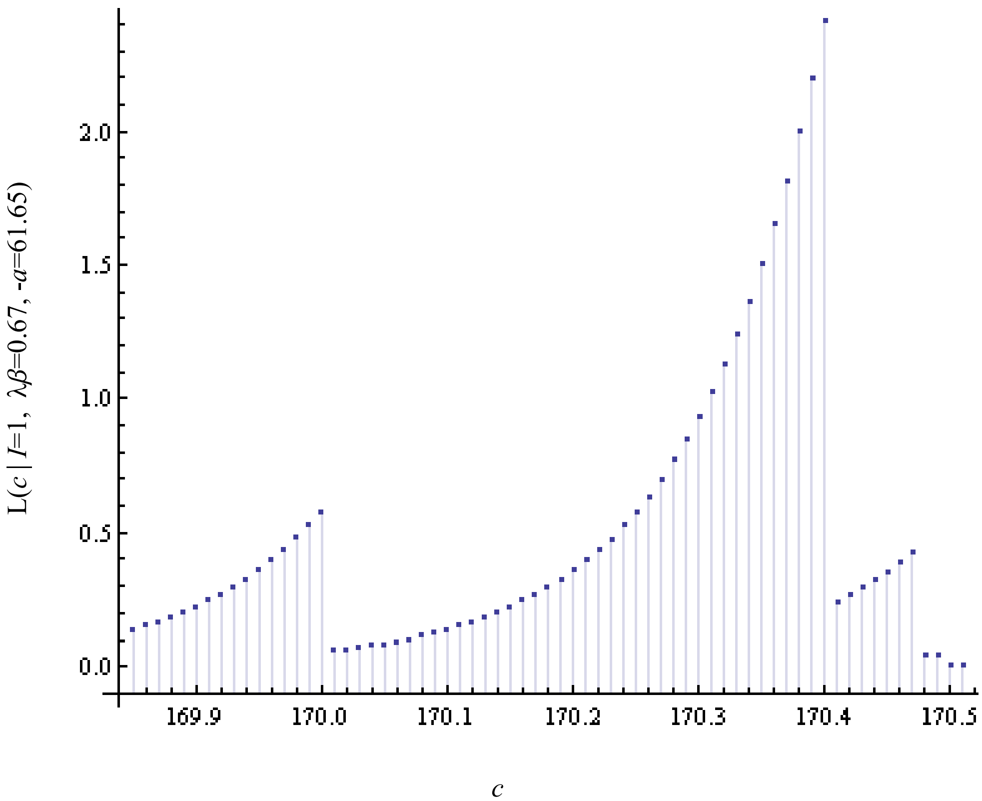
Vertical bars for the empirical histogram and the line for the estimated density function.
For scale reasons, we only graph positive brokerage fees. Histogram Width: 5 BRL.
Sample Size: 2,932; Empirical Fraction of Zero Fees: 18.42%; Estimated Fraction of Zero Fees: 13.44%.

Table 6. Global Maximum: Top-10 Likelihood Values – Single Model Subsample

Likelihood ($\times 10^{-7,179}$)	***I***	**$\lambda\beta$**	***c***	***-a***
2.41	1	.67	170.40	61.65
2.19	1	.67	170.39	61.65
1.99	1	.67	170.38	61.64
1.81	1	.67	170.37	61.64
1.65	1	.67	170.36	61.64
1.50	1	.67	170.35	61.63
1.36	1	.67	170.34	61.63
1.24	1	.67	170.33	61.63
1.12	1	.67	170.32	61.62
1.02	1	.67	170.31	61.62

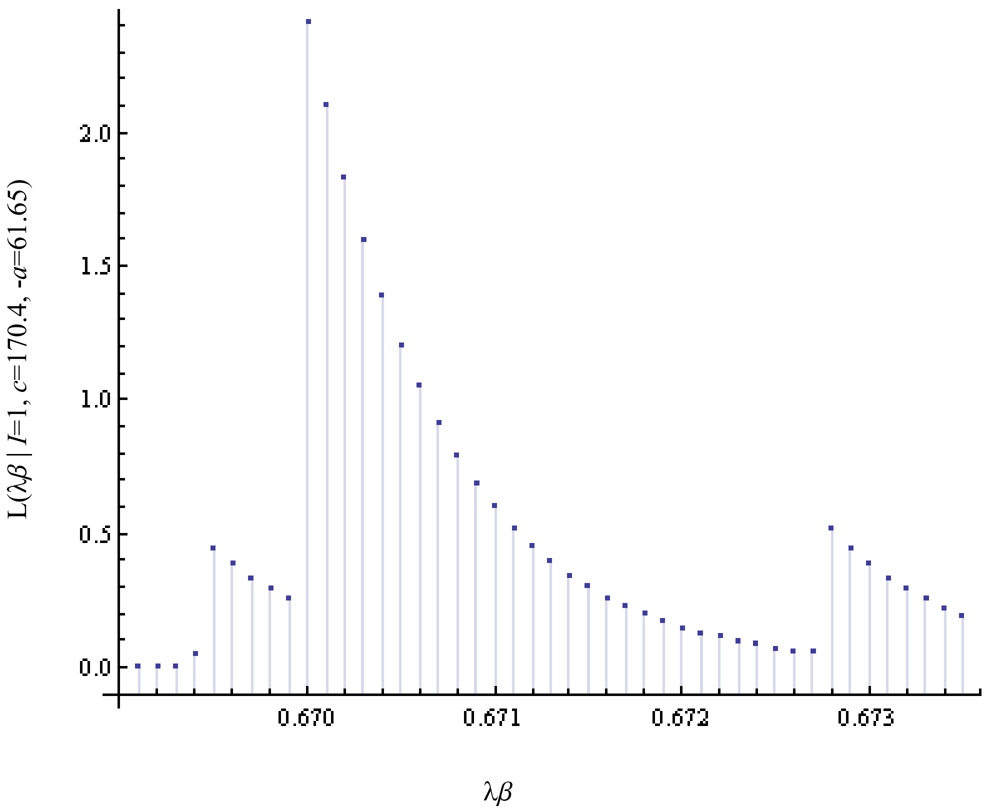
Sample Size: 2,932.

Figure 9. Local Maximum: Switching Cost (*c*) – Single-Model Subsample



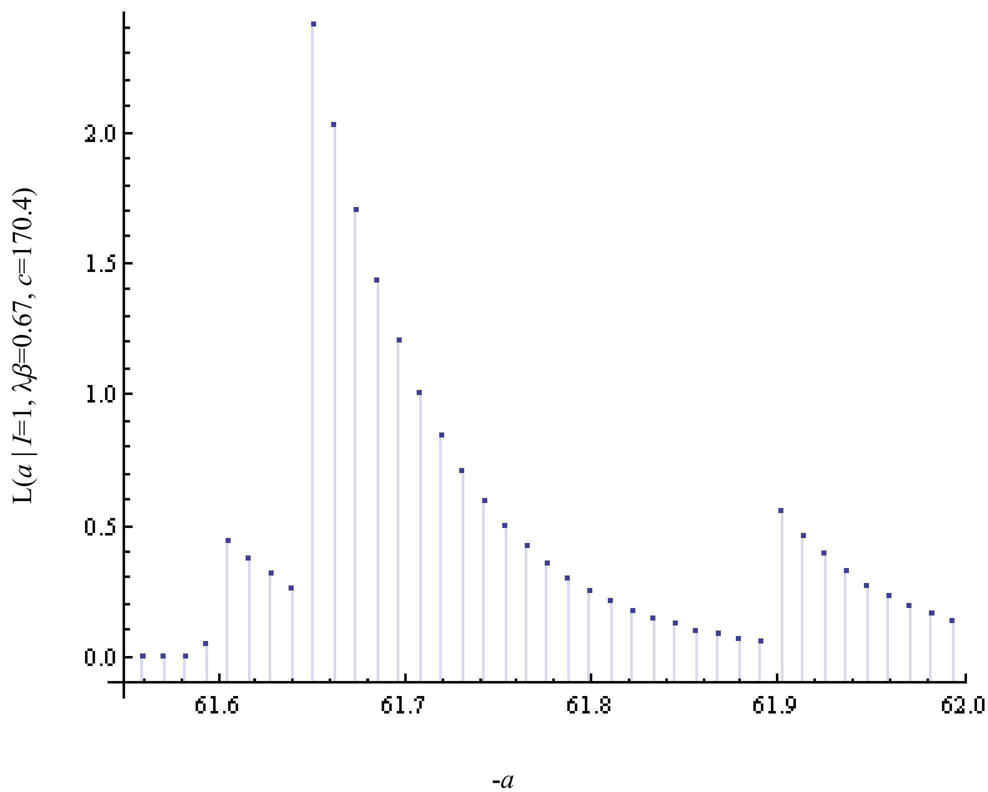
Likelihood ($\times 10^{-7,179}$) for different switching costs when $I = 1$, $\lambda\beta = 0.67$, and $-a = 61.65$.
Global Maximum: $2.4\times 10^{-7,179}$; Sample Size: 2,932.

Figure 10. Local Maximum: Adjusted Discount Factor ($\lambda\beta$) – Single-Model Subsample



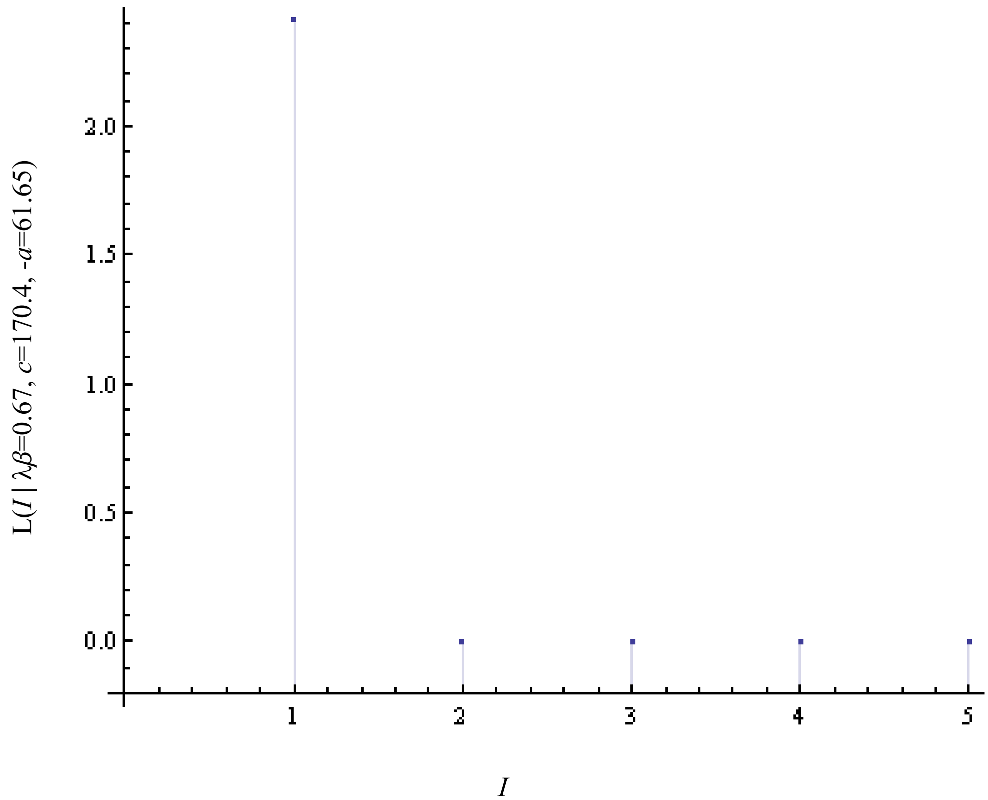
Likelihood ($\times 10^{-7,179}$) for different adjusted-discount factors when I = 1, c = 170.4, and $-a$ = 61.65.
Global Maximum: $2.4\times 10^{-7,179}$; Sample Size: 2,932.

Figure 11. Local Maximum: Highest Gift (-*a*) – Single-Model Subsample



Likelihood ($\times 10^{-7,179}$) for different values of -*a* when I = 1, $\lambda\beta$ = 0.67, and c = 170.4.
Global Maximum: $2.4 \times 10^{-7,179}$; Sample Size: 2,932.

Figure 12. Local Maximum: Unmatched Offers (*I*) – Single-Model Subsample



Likelihood ($\times 10^{-7,179}$) for different values of I when $\lambda\beta = 0.67$, $c = 170.4$, and $-a = 61.65$.
Global Maximum: $2.4\times 10^{-7,179}$; Sample Size: 2,932.