FUNDAÇÃO GETULIO VARGAS

ESCOLA de PÓS-GRADUAÇÃO em ECONOMIA

MANUELA HUNGERBÜHLER LOPES

USING CONSUMER SEARCH COST AND LOYALTY TO EXPLAIN DISPERSION IN BANKING FEES

Rio de Janeiro
2011

MANUELA HUNGERBÜHLER LOPES

USING CONSUMER SEARCH COST AND LOYALTY TO EXPLAIN DISPERSION IN BANKING FEES

Dissertação para obtenção do grau de mestre apresentada à Escola de Pós-Graduação em Economia

Área de concentração: Organização Industrial

Orientador: Luis Henrique Bertolino Braido

Rio de Janeiro
2011

Ficha catalográfica elaborada pela Biblioteca Mario Henrique Simonsen/FGV

Lopes, Manuela Hungerbühler
Using consumer search cost and loyalt to explain dispersion in banking fees / Manuela Hungerbühler Lopes. – 2011.
32 f.

Dissertação (mestrado) - Fundação Getulio Vargas, Escola de Pós-Graduação em Economia.
Orientador: Luis Henrique Bertolino Braido.
Inclui bibliografia.

1. Organização industrial. 2. Preços. 3. Bancos – Tarifas. I. Braido, Luis H. B. II. Fundação Getulio Vargas. Escola de Pós- Graduação em Economia. III. Título.

CDD – 338.52

Dedico este trabalho aos meus heróis, Teresa e Rogério (meus queridos pais), exemplos de amor, bondade, força, perseverança e dedicação. Devo tudo a eles.

AGRADECIMENTOS

À Fundação Getúlio Vargas.

Ao meu orientador.

Aos meus colegas do mestrado.

Aos meus amigos.

À minha família.

RESUMO

Pode-se observar uma considerável dispersão entre os preços que diferentes bancos comerciais no Brasil cobram por um mesmo pacote homogêneo de serviços—dispersão esta que é sustentada ao longo do tempo. Em uma tentativa de replicar esta observação empírica, foi desenvolvido um simples modelo que lança mão do arcabouço da literatura de custos de procura (*search costs*) e que baseia-se também na lealdade por parte dos consumidores. Em seguida, dados de preços referentes ao setor bancário brasileiro são aplicados ao modelo desenvolvido e alguns exercícios empíricos são então realizados. Esses exercícios permitem que: (i) os custos de procura incorridos pelos consumidores sejam estimados, ao fixar-se os valores dos demais parâmetros e (ii) as correspondentes perdas de peso-morto que surgem como consequência dos custos de procura incorridos pelos consumidores sejam também estimadas. Quando apenas 80% da população é livre para buscar por bancos que cobrem menores tarifas, à taxa de juros mensal de 0,5%, o valor estimado do custo de procura médio incorrido pelos consumidores chega a 1805,80 BRL, sendo a correspondente perda de peso-morto média na ordem de 233,71 BRL por consumidor.

Palavras-chave: Custos de procura. Lealdade de consumidores. Dispersão de preços. Tarifas bancárias.

ABSTRACT

Prices of a homogenous package of services present considerable dispersion across banks. A simple model that uses consumer search costs and loyalty is devolped in order to reproduce this empirical observation. Using data for the brazilian banking sector, empirical exercises are performed to allow for: (i) the estimation of consumer search costs by fixing the values of other parameters; and (ii) the estimation of the corresponding deadweightlosses imposed by costly search. When only 80% of the population is free to engage in search activity, at a 0.5% monthly interest rate, the upper limit of the support of the search cost distribution is found to be 1805.80 BRL. In this case, the corresponding estimate of the deadweightloss imposed on an average searcher is 233.71 BRL.

Keywords: Search costs. Consumer loyalty. Price dispersion. Banking Fees.

Contents

1 Introduction

According to the Brazilian Geography and Estatistic Institute's (IBGE) 2008-2009 household budget survey (POF), in 2008 approximately 5% of families' expenses were on bank charges. It is therefore no wonder that the problem of banking fees has recently been receiving greater attention from the general public, as well as the public authorities. In fact, from 2008 onwards, the Brazilian Monetary Counsil (CMN), by means of a series of resolutions, has been increasing regulation in the banking sector with the intention of disciplining and bringing more transparency to the levying of tariffs.

Most brazilian commercial banks offer their clients a menu of "ready-made" packages to choose from— each of them including determined quantities of a certain variety of services. The contents of these packages usually vary from bank to bank, which makes any comparison between them somewhat difficult. The closest we come to a homogenous product is the *standardized package*,[1] which was created by the CMN and must be offered by every commercial bank in Brazil. In the real world, it is true that a product or service offered by one bank will not be exactly like that very same product or service offered by another. Differences in quality, such as the time clients expect to stay in line at the bank, or how easy it is to find ATMs and branches around town, do play a role when it comes to choosing where to open an account and are likely to be partially responsible for price discrepancies among banks. However, we will set choice of quality and possible consumer discrimination aside and seek for other factors thay may determine banking fees.

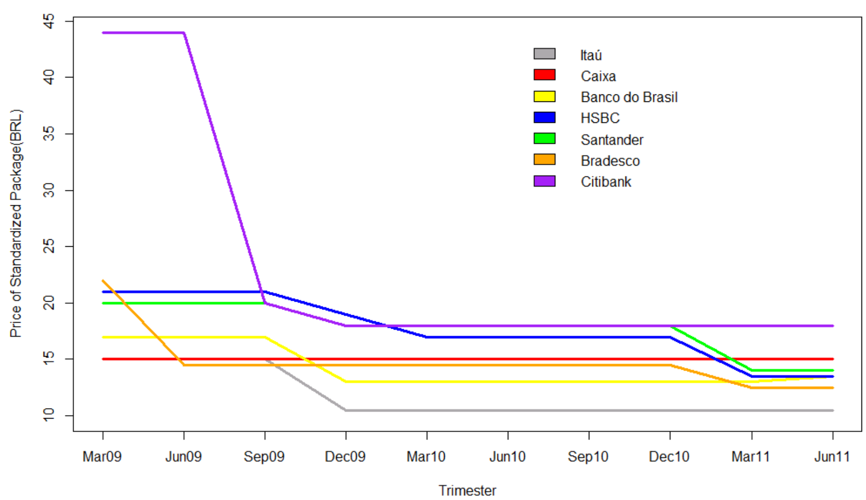
Since we are assuming the standardized package is a homogenous product, we will henceforth focus on it. We used the Brazilian Central Bank's (BACEN) website to collect quarterly data on prices of the standardized package, as well as market-shares and number of branches, for seven of the currently largest commercial banks in Brazil— Itaú, Caixa Econômica Federal, Banco do Brasil, HSBC, Santander, Bradesco and Citibank— from March 2009 to June 2011. Table 1 shows the prices being charged for the standardized package in June 2011 by these banks. We notice considerable price dispersion across banks: the highest price is 1.7 times the lowest one and the difference between them is 7.50 BRL. Looking at Graph 1, which shows how the prices for the standardized

[1] *Pacote Padronizado Pessoa Física.* It includes, for every month: 8 cash withdrawals; 4 bank statements regarding the current month and 2 regarding the past month; 4 bank transfers.

Table 1: Prices for the Standardized Package in June 2011 (Source: BACEN)

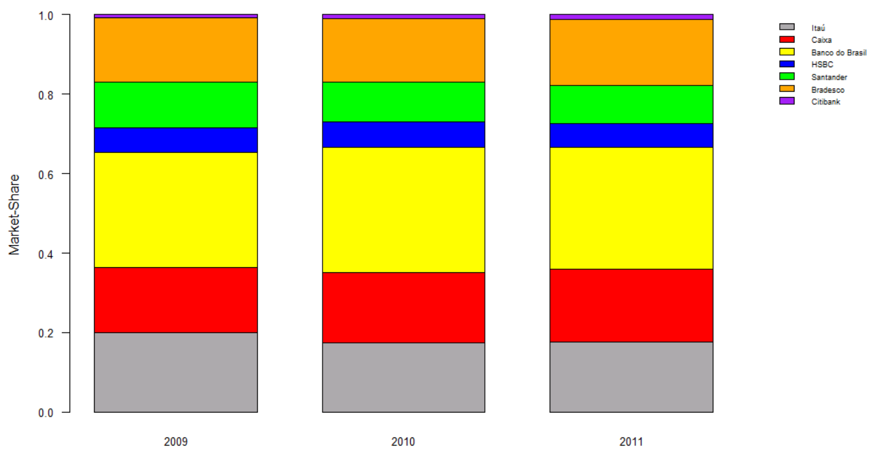
Bank	**Price (BRL)**
Itaú	10.50
Bradesco	12.50
Banco do Brasil	13.50
HSBC	13.50
Santander	14.00
Caixa Econômica Federal	15.00
Citibank	18.00
Mean	13.86
Maximum	18.00
Minimum	10.50
Maximum - Minimum	7.50
Maximum/Minimum	1.71

package have evolved from 2009 to 2011, one could say that most prices are left unchanged for reasonably long periods of time, especially from mid-2009 onwards.



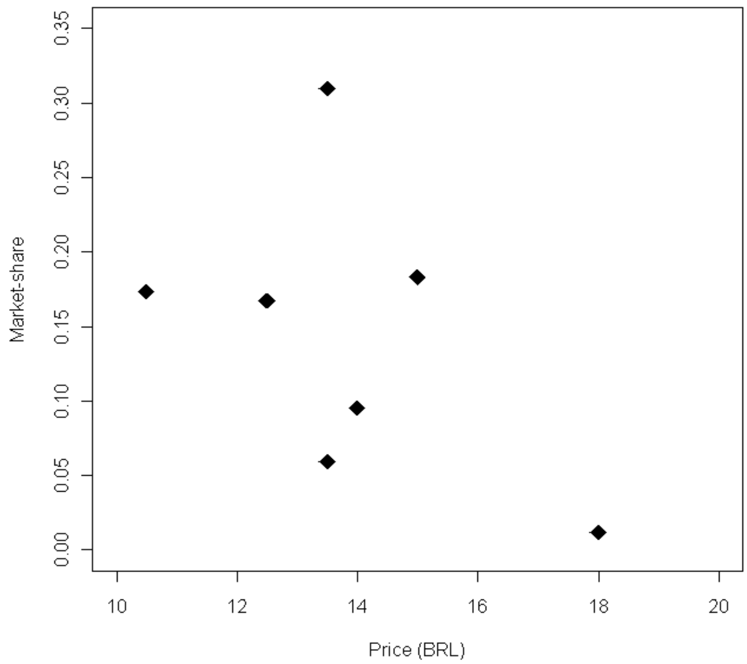
Graph 1: Prices for the standardized package for 2009-2011

Graph 2 shows that the market-shares[2] for the banks being considered have remained roughly unchanged since 2009, indicating perhaps that they are in some sort of stable equilibrium that does not seem to have been disrupted. Finally, Graphs 3 and 4 plot, respectively, market-share (calculated from the number of checking accounts) against the price of the standardized package, and price against number of branches, for June 2011, for the seven banks considered.

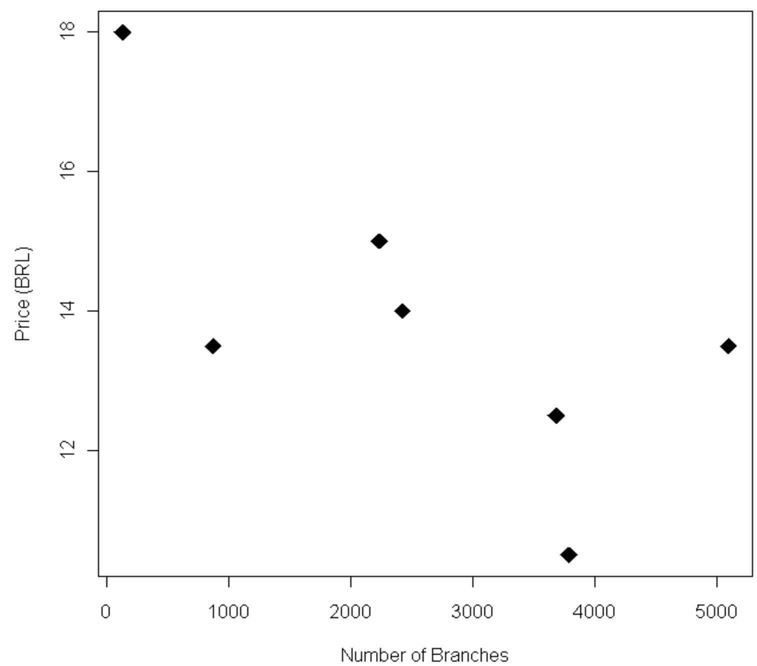


Graph 2: Market-shares for 2009-2011

[2] In this case we use deposit market-shares, defined as the amount on deposit at a particular bank divided by the total amount on deposit at all banks.



Graph 3: Market-share vs. standardized package price (Jun/11)



Graph 4: Standardized package price vs. no. of branches (Jun/11)

The above data has motivated us to search for a model of the checking account market that produces price dispersion in a pure-strategy equilibrium. Differences in the quality of the services provided by different banks, or the existence of market frictions, like consumer switching and search costs may all be forces acting towards this price dispersion. Indeed, it seems plausible to assume that search and switching costs permeate the banking market. When changing banks, a consumer is faced with many costs such as having to visit the banks personally to terminate the old contract and sign a new one, providing documents, memorizing new pin numbers, etc. Similarly, the search for banks may also be costly in the sense that consumers must inquire about prices, which usually entails also having to personally go to a bank branch. These barriers consumers must face when collecting information on prices and services or when changing banks may be so time-consuming or psychologically stressful that they represent significant costs, which affect the equilibrium in the banking market,

generating the price dispersion observed in the data. Literature on consumer search and switching costs are quite vast, but remain mostly independent from each other. Hoping to maintain our model's tractability, we chose to focus on only one of them: search costs. Doing so enabled us to make use of the Carlson and McAfee (1983) framework, which was especially appealing to us due to its results. Indeed, whereas in most switching or search costs models price dispersion typically arises as a result of mixed strategy equilibria,[3] in Carlson and McAfee it is possible to obtain price dispersion from a pure-strategy equilibrium, which seems to be in line with what is empirically observed regarding banking fees.[4]

Another interesting aspect of the banking market is that some consumers are forced to keep accounts with specific banks. For example, many workers, like public servants, for instance, have their salaries deposited into accounts on banks of their employers' choice.[5] We found it interesting to incorporate this aspect into our model, seeing as it might mean that many banks already enter the market with a guaranteed *"minimum market-share"* of consumers that is exogenously determined. We will call these the shares of *loyal customers*.[6] Hence, in addition to search costs that reduce the extensiveness of consumers' search activity across banks, we propose a slight addition to the Carlson and McAfee framework by introducing these shares of loyal customers. This will further weaken the incentive for banks to cut prices. We allow these exogenous shares of loyal customers to differ among banks, providing additional support for a pure-strategy equilibrium exhibiting price dispersion.

In Section 2 we develop the model that uses consumer search cost and loyalty to produce equilibrium price dispersion. In Section 3, we apply the data for June 2011 to the model and perform some empirical exercises in which we fix different parameters and estimate others, including the upper limit of the distribution of consumer search costs for the banking industry. Having done that, we then use these results to obtain the corresponding estimates for the deadweight loss an average searcher would expect to incur due to search costs. Assuming only 80%

[3] See Stahl (1989) and Varian (1980).

[4] See Dahlby and West (1986) for an example of an empircal paper that also uses the Carlson and McAfee framework.

[5] Attempting to reduce this problem, the CMN has in fact imposed measures to ensure that private sector workers be able to choose the bank and account in which to receive their salaries. These changes are very recent and not commonly used yet. Also, public servants and social security beneficiaries have been excluded from these measures until at least the end of 2011.

[6] See De los Santos, Hortaçsu and Wildenbeest (2009) for a model that combines costly consumer search with consumer loyalty.

of the population is free to engage in search activity, at a 0.5% monthly interest rate, we estimate the upper limit of the support of the search cost distribution to be 1805.80 BRL. Moreover, we find that, in this case, the corresponding deadweightloss imposed on an average searcher is 233.71 BRL.

2 The Model

The economy has multiple agents and banks. There is a total of S consumers who each demand one unit of the homogenous package of services offered by the total of N banks. By assumption, any given bank k has an exogenous fraction of consumers, namely $0 < \lambda_k < 1$, who are loyal to it.[7] These loyal consumers engage in no search activity and will demand the services offered by the bank to which they are loyal. Assume

$$0 < \lambda \equiv \sum_{k=1}^{N} \lambda_k < 1$$

meaning that only a fraction $(1 - \lambda)$ of consumers will search for the lowest prices.

A consumer i faces a cost of s_i for every bank he visits (except for the first sample, which we will assume to be free). This individual marginal search cost value can be regarded as a realization of a random variable s with continuous distribution function $H(s)$ and density function $h(s) = H'(s)$. So, consumers are only heterogenous with respect to search costs. We will assume that consumers search following a sequential, reservation-price strategy, which will be later described, and that search costs are uniformally distributed. More especifically, we have that:

$$H(s) = \begin{cases} \frac{s}{b} & \text{for } 0 \leq s \leq b \\ 1 & \text{for } s > b \\ 0 & \text{otherwise} \end{cases}$$

$$h(s) = \begin{cases} \frac{1}{b} & \text{for } 0 \leq s \leq b \\ 0 & \text{otherwise} \end{cases}$$

for some $0 < b < \infty$.

To simplify matters, we assume that, given the banks' pricing decision, consumers are able to observe the cumulative distribution function of prices, which implies that they know what expenditures they can be faced with, but cannot tell which bank is associated with which. After they have set their prices, the N banks can be rearranged and then labelled according to the price they charge for the package of services, in an increasing order, such that:

$$e_1 \leq e_2 \leq \ldots \leq e_N$$

[7] The share of loyal customers may differ from bank to bank.

Where e_j is the j-lowest price in the market and the bank that charges e_j for the package of services is now labelled "bank j". That is, after their price-setting decisions, banks are labelled so that bank 1 is the one that offers consumers the best deal, bank 2 the second-best, and so on, until bank N, which charges them the highest price.

In terms of sampling, we assume that every bank is drawn with equal probability and that there is replacement.[8] This means that consumers enter the market facing the following perceived distribution of total expenditure with bank services:

$$f(e) = \frac{1}{N} \quad \forall e \in \{e_1, ..., e_N\}$$

Let $\mathbf{e} \equiv (e_1, e_2, ..., e_N) \in \mathrm{R}^N_{++}$ be the vector of the prices available in the market and $0 < \delta < 1$ consumers' time discount factor. Suppose a consumer is faced with a certain bank j's price e_j. Then we define $d_j(\mathbf{e})$ as the expected gain from searching for a lower price. Hence:

$$\begin{aligned} d_j(\mathbf{e}) &= \sum_{k=1}^{j-1} \sum_{t=0}^{\infty} \delta^t (e_j - e_k) f(e_k) \\ &= \frac{1}{1-\delta} \sum_{k=1}^{j-1} \frac{e_j - e_k}{N} \end{aligned} \tag{1}$$

Note that we have assumed that once a consumer chooses a bank, provided that banks do not change their pricing strategies, he will remain with it throughout his lifetime. Therefore when he is computing the gain from searching for lower prices, he must consider that the difference in prices will be carried along every month, for the rest of his life— which we consider to be infinite. Also, observe that, for all $j \in \{1, ..., N-1\}$,

$$\begin{aligned} d_{j+1}(\mathbf{e}) - d_j(\mathbf{e}) &= \frac{1}{1-\delta} \sum_{k=1}^{j} \frac{e_{j+1} - e_k}{N} - \frac{1}{1-\delta} \sum_{k=1}^{j-1} \frac{e_j - e_k}{N} \\ &= \frac{j(e_{j+1} - e_j)}{N(1-\delta)} \\ &\geq 0 \end{aligned} \tag{2}$$

That is, a consumer's gain from searching is non-decreasing in the ranking

[8] This assumption simplifies calculations and is perhaps not that unreasonable, if you consider that, in real life, after sampling one bank, consumers have only a very small gain of information regarding the other prices being charged by the rest of the banks.

position of the bank he just sampled. If $j = 1$, naturally, $d_1 = 0$, because the consumer has already sampled the best deal and will have no gain from searching further. These $d_1, d_2, ..., d_N$[9] will help us define the search rule for our model. We know that consumers only decide to carry on searching if the expected marginal gain from searching further exceeds its marginal cost. Take a particular consumer i, for instance, whose search cost is s_i. When he samples a bank j, he expects to benefit in d_j if he decides to search further for a lower-priced bank. Thus, he will only choose to do so if $s_i < d_j$. If not, he sticks to bank j and does not engage in any more search activity. More specifically, suppose s_i is such that $d_{j*} \leq s_i < d_{j*+1}$ for some $j * \epsilon \{1, ..., N-1\}$. Then he will search for banks until he is faced with a price of e_{j*} or less.[10] Since consumers differ in search costs, different consumers will have different stopping points. For instance, a consumer whose search cost exceeds or equals d_N will stick to the first bank it samples whereas one whose search cost is less than d_2 will search until he finds the cheapest bank.

2.1 Demand Functions

Given prices $\mathbf{e}$ and thus the ranking implied by $e_1 \leq e_2 \leq ... \leq e_N$, we are able to determine the demand fuctions faced by the banks. Let us begin with the more expensive bank. In addition to its loyal customers, when sampled, the bank ranked in the Nth position will only be able to capture consumers whose search costs exceed or equal d_N. This happens with probability $H(b) - H(d_N)$. Since banks are equally likely to be sampled, the probability that the lowest ranked bank is sampled by a consumer who will stick to it is $\frac{1}{N}(H(b) - H(d_N))$. Bank N's expected market-share is therefore:

$$\begin{aligned} q_N(\mathbf{e}) &= \lambda_N + \frac{1-\lambda}{N} \cdot (H(b) - H(d_N)) \\ &= \lambda_N + \frac{1-\lambda}{N} \cdot \left(\frac{b - d_N}{b} \right) \end{aligned}$$

As for the $(N-1)th$ bank, besides having its loyal costumers and a $\frac{1}{N}$ share of the higher search cost customers, it also shares with the other $N-2$ banks ranked above it those customers whose search costs lie within the $[d_{N-1}, d_N)$

[9] From this point on, to simplify notation, we shall refer to $d_j(\mathbf{e})$ as d_j.

[10] In other words, a consumer whose search cost lies within an interval $[d_{j*}, d_{j*+1})$ has a reservation price of e_{j*}, meaning that he will sample banks until he is faced with a charge of e_{j*} or less.

interval. Therefore its expected market-share will be:

$$\begin{aligned} q_{N-1}(\mathbf{e}) &= \lambda_{N-1} + \frac{1-\lambda}{N} \cdot (H(b) - H(d_N)) + \frac{1-\lambda}{N-1} \cdot (H(d_N) - H(d_{N-1})) \\ &= \lambda_{N-1} + \frac{1-\lambda}{b} \cdot \left(\frac{b-d_N}{N} + \frac{d_N - d_{N-1}}{N-1} \right) \end{aligned}$$

So, given $\mathbf{e}$, a certain bank j's expected market-share is given by:

$$q_j(\mathbf{e}) = \begin{cases} \lambda_j + \left(\frac{1-\lambda}{b}\right) \left(\frac{b-d_N}{N} + \sum\limits_{k=j}^{N-1} \frac{d_{k+1}-d_k}{k} \right) & \text{if } j = 1, 2, ..., N-1 \\ \lambda_N + \left(\frac{1-\lambda}{b}\right)\left(\frac{b-d_N}{N}\right) & \text{for } j = N \end{cases} \tag{3}$$

where $d_1 = 0$. Using (2) on the above equation, we can rewrite the market-shares so that they are expressed in terms of the amounts charged by the banks, and their mean, $\overline{e} \equiv \sum\limits_{j=1}^{N} \frac{e_j}{N}$, and obtain:

$$q_j = \lambda_j + \frac{1-\lambda}{N} + \frac{1-\lambda}{Nb(1-\delta)} (\bar{e} - e_j) \tag{4}$$

2.2 Banks

There are N banks competing on prices for the consumers who are free to engage on actual search. We define $c_j \geq 0$ as bank j's constant marginal cost of acquiring another customer. We allow banks to differ in the proportion of customers loyal to them, as well as in their marginal costs.

Given these assumptions, a bank js' profits are as follows:

$$\pi_j(\mathbf{e}) = Sq_j(\mathbf{e}) \cdot e_j(\mathbf{e}) - Sq_j(\mathbf{e})\, c_j \tag{5}$$

2.3 Equilibrium

Banks know the distribution of search costs, as well as the search rule that consumers follow. Based on this knowledge, every bank j chooses to charge accountholders the amount e_j that maximizes expected profit, according to the following problem:

$$\max_{e_j \in \mathbf{R}} \pi_j(\mathbf{e})$$

The first order condition (FOC) is that:

$$\frac{\partial \pi_j(\mathbf{e})}{\partial e_j} = 0$$

That is:[11]

$$Sq_j + S\frac{\partial q_j}{\partial e_j}(e_j - c_j) = 0$$

$$\Leftrightarrow q_j + \frac{\partial q_j}{\partial e_j}(e_j - c_j) = 0 \tag{6}$$

From (1) we have that

$$\frac{\partial d_j}{\partial e_i} = \begin{cases} \frac{(j-1)}{N(1-\delta)} & \text{if } i = j \\ -\frac{1}{N(1-\delta)} & \text{if } i < j \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

Therefore, from (3) we obtain, for every j, that:

$$\frac{\partial q_j}{\partial e_j} = -\frac{(1-\lambda)}{b(1-\delta)} \cdot \frac{(N-1)}{N^2} \tag{8}$$

Substituting (8) in (6), the FOC becomes:

$$q_j - \frac{(1-\lambda)}{b(1-\delta)} \cdot \frac{(N-1)}{N^2}(e_j - c_j) = 0 \tag{9}$$

Substituting (4) into (9) yields, for every j:

$$\lambda_j + \frac{1-\lambda}{N} + \frac{1-\lambda}{Nb(1-\delta)}(\bar{e} - e_j) - \frac{(1-\lambda)}{b(1-\delta)} \cdot \frac{(N-1)}{N^2}(e_j - c_j) = 0$$

After some algebraic manipulation we find that:

$$e_j = \frac{1}{2(N-1)}\sum_{i \neq j} e_i + \frac{bN^2(1-\delta)}{2(N-1)(1-\lambda)}\lambda_j + \frac{bN(1-\delta)}{2(N-1)} + \frac{1}{2}c_j \tag{10}$$

[11] From here on, we supress the argument $\mathbf{e}$ and refer to $\pi_j(\mathbf{e}), q_j(\mathbf{e})$ and $e_j(\mathbf{e})$ as π_j, q_j and e_j, respectively.

For $i \neq j$:

$$e_i = \frac{1}{2(N-1)} \sum_{m \neq i,j} e_m + \frac{1}{2(N-1)} e_j + \frac{bN^2(1-\delta)}{2(N-1)(1-\lambda)} \lambda_i + \frac{bN(1-\delta)}{2(N-1)} + \frac{1}{2} c_i$$

Summing over $i \neq j$:

$$\sum_{i \neq j} e_i = \frac{N-2}{2(N-1)} \sum_{i \neq j} e_i + \frac{1}{2} e_j + \frac{bN^2(1-\delta)}{2(N-1)(1-\lambda)} \sum_{i \neq j} \lambda_i + \frac{bN(1-\delta)}{2} + \frac{1}{2} \sum_{i \neq j} c_i \tag{11}$$

Thus, substituting (11) in (10) yields:

$$e_j = \frac{bN^2(1-\delta)}{(2N-1)(1-\lambda)} \left(\lambda_j + \frac{\lambda}{N-1} \right) + \frac{bN(1-\delta)}{N-1} + \frac{N}{2N-1} \bar{c} + \frac{N-1}{2N-1} c_j \tag{12}$$

where we define the *average marginal cost* $\bar{c} \equiv \sum_{j=1}^{N} \frac{c_j}{N}$.

We were able to express the total equilibrium amounts charged by banks for their services solely in terms of the model's parameters. From the above expression, we can calculate:

$$\bar{e} = \frac{bN(1-\delta)}{(N-1)(1-\lambda)} + \bar{c} \tag{13}$$

We can now obtain the equilibrium market-shares by substituting (12) and (13) in (4):

$$q_j = \frac{1}{N} + \frac{N-1}{2N-1} \left(\lambda_j - \frac{\lambda}{N} \right) + \frac{(N-1)(1-\lambda)}{bN(2N-1)(1-\delta)} (\bar{c} - c_j) \tag{14}$$

Observe that those banks operating on lower marginal costs are able to set relatively lower prices than those facing a higher marginal cost. This reflects on their market-shares: the lower a bank's marginal cost relative to the average, the greater its market-share. Equilibrium prices are increasing in b, the parameter of the uniform search cost distribution, and decreasing in the total proportion of searchers in the economy, $(1-\lambda)$, as one would expect. Also, all other things kept constant, the larger the λ_j, the greater the amount charged by bank j. This makes sense, since the banks are faced with a trade-off between exploiting their loyal customer base and setting lower prices to compete for searchers. The greater the proportion of loyal customers a bank is endowed with, the smaller its incentive to lower prices and compete with the other banks for searchers. Simi-

larly, a larger b implies that potential searchers will generally be less inclined to engage in actual search and banks need not compete so "fiercely" for them. As consumers' discount factor, δ, approaches 1 (i.e., consumers value the following month as much as they do the present), the discounted sum of price differences between banks increases and they are relatively more inclined to search. Consequently, as $\delta \longrightarrow 1$, the first two terms in the right-hand side of equation (12) tend to zero— the fact that searchers are more inclined to search increases the banks' incentive to compete in price and their share of loyal customers becomes reletively less important in their pricing decision. In terms of market-shares, since the banks with relatively lower marginal costs are also those who can offer lower prices and better compete for searchers, as $\delta \longrightarrow 1$, differences across banks' marginal costs become relatively more relevant in determining differences across market-shares.

3 Empirical Exercises

As it is, the model cannot be identified, because there are only two moment conditions (equations (12) and (14)), but sixteen unknown paramenters to be estimated: b, δ, c_j and λ_j for $j \in \{1, ..., 7\}$. Ideally, we would need quite a large number of instrumental variables (at least seven) to make the model identifiable. Unfortunately, we are unable find suitable choices of instruments from the variables available at BACEN's website. So in the attempt of working around this problem, we have performed instead the following empirical exercises:

Firstly, we determine that a bank's share of loyal customers is proportional to its participation in the total number of branches, $a_j \equiv \frac{\text{bank } j\text{'s number of branches}}{\text{total number of branches}}$

$$\lambda_j = \beta a_j$$

Behind this is the conjecture that a bank's share of loyal customers might consist of consumers who simply choose to open their accounts at the bank nearest to them.[12] That being the case, then a bank with a relatively greater number of branches is more likely to be chosen by this kind of consumer and we would expect β to be positive. It seems somewhat reasonable to assume that, in the short run, a bank's number of branches is fixed, which allows us to treat it as an exogenous variable. Note that if we sum over λ_j for the total of banks, we find that $\lambda = \beta$. So, in fact,

$$\lambda_j = \lambda a_j$$

Additionally, we also made the simplifying assumption that the marginal cost of offering the standardized package is some value c that is the same for every bank. Hence we are faced with the following system of equations:

$$\begin{aligned} e_j &= \frac{bN^2(1-\delta)\lambda}{(2N-1)(1-\lambda)}\left(a_j + \frac{1}{N-1}\right) + \frac{bN(1-\delta)}{N-1} + c \qquad (15) \\ q_j &= \frac{1}{N} + \frac{N-1}{2N-1}\left(a_j - \frac{1}{N}\right)\lambda \end{aligned}$$

We collected data from BACEN's website on the seven largest commercial banks in Brazil: Itaú, Caixa Econômica Federal, Banco do Brasil, HSBC, San-

[12] This formulation is quite a bit restrictive and would perhaps require that at least a constant be added to a bank's share of loyal consumers. However, this would unfortunately impose more parameters to be estimated, which is precisely what we intended to avoid.

tander, Bradesco, Citibank for June 2011.[13] In order to estimate b and λ, we fixed values for δ and c that ranged from 0.980 to 0.995 in 0.005 intervals and from 0 BRL to 10 BRL in intervals of 2 BRL, respectively.[14] The system was estimated by the *Generalized Method of Moments (GMM)*, as described in Hansen (1982), based on the following vector of moment conditions:

$$E\left[\begin{array}{c} e_j-\frac{bN^2(1-\delta)\lambda}{(2N-1)(1-\lambda)}\left(a_j+\frac{1}{N-1}\right)-\frac{bN(1-\lambda)}{N-1}-c \\ q_j-\frac{1}{N}-\frac{N-1}{2N-1}\left(a_j-\frac{1}{N}\right)\lambda \end{array}\right]=\mathbf{0}.$$

The sample counterpart of this is:

$$\bar{g}\left(b,\lambda\right)\equiv\left[\begin{array}{c} \frac{1}{N-1}\sum_{j=1}^{N-1}\left(e_j-\frac{bN^2(1-\delta)\lambda}{(2N-1)(1-\lambda)}\left(a_j+\frac{1}{N-1}\right)-\frac{bN(1-\delta)}{N-1}-c\right) \\ \frac{1}{N-1}\sum_{j=1}^{N-1}\left(q_j-\frac{1}{N}-\frac{N-1}{2N-1}\left(a_j-\frac{1}{N}\right)\lambda\right) \end{array}\right]=\mathbf{0}.$$

It is made as close to zero as possible by minimizing the quadratic function $\bar{g}\left(b,\lambda\right)'W\bar{g}\left(b,\gamma\right)$, where W is a positive definite symmetric matrix of weights. The GMM depends only on moment conditions and there is thus no need to make assumptions and approximations for the distributions of the variables. This distribution-free feature of the GMM is especially appealing, because the error terms in the model are not very well specified— there are several "sources of error", such as: the approximation of λ_j by λa_j, omitting other factors that also explain customer fidelity; the fact that market-shares are calculated from the total number of checking accounts of each bank instead of the number of standardized packages sold, which would be more accurate but is not made available by BACEN; prices that do not take into account discounts given out in real life and marginal costs that are simplified to be the same for every bank.

The results obtained from our first attempt at estimating the model by GMM with fixed values for δ and c are ilustrated in Table 2. The estimated values for b are disappointingly low in comparison with what would be the discounted sums of price differences between the banks. Additionally, by (15), we would expect b to increase with δ and decrease with c. However, the estimated values for b do not seem to exhibit much sensitivity towards variations to the set values

[13] We eliminated the equations for the last bank, Citibank, so that the system would not be singluar, since the market-shares add up to 1, by construction, if all the banks' equations are included.

[14] We chose values for δ that were compatible with monthly interest rates of 0.5% to 2% and marginal costs that were positive but did not exceed the lowest price for the standardized package.

Table 2: Results for first empirical exercise (p-values in parentheses).

	δ							
	0.980		0.985		0.990		0.995	
c	**b**	$\boldsymbol{\lambda}$	**b**	$\boldsymbol{\lambda}$	**b**	$\boldsymbol{\lambda}$	**b**	$\boldsymbol{\lambda}$
0	7.145	0.986	6.608	0.991	6.718	0.994	7.444	0.996
	(0.997)	(0.799)	(0.998)	(0.798)	(0.999)	(0.797)	(0.999)	(0.797)
2	7.178	0.984	7.650	0.987	6.728	0.992	7.017	0.996
	(0.997)	(0.800)	(0.995)	(0.799)	(0.998)	(0.798)	(0.999)	(0.797)
4	7.151	0.980	7.041	0.986	6.968	0.990	7.570	0.995
	(0.994)	(0.801)	(0.994)	(0.799)	(0.997)	(0.798)	(0.998)	(0.797)
6	6.925	0.976	6.886	0.982	7.388	0.987	7.090	0.994
	(0.993)	(0.802)	(0.996)	(0.801)	(0.996)	(0.799)	(0.998)	(0.798)
8	6.508	0.969	6.882	0.975	7.003	0.983	6.666	0.992
	(0.991)	(0.804)	(0.995)	(0.803)	(0.995)	(0.801)	(0.998)	(0.798)
10	1.695	0.987	6.343	0.962	6.412	0.975	7.100	0.986
	(0.999)	(0.800)	(0.989)	(0.806)	(0.993)	(0.803)	(0.996)	(0.801)

of these parameters. For this version of the model, we find all of the estimated values of λ to be practically equal to 1. This suggests that when we suppose that the only source of heterogeneity among banks is their shares of loyal customers and we approximate these by a linear function of their participation in the total number of branches, search costs do not seem to be what causes dispersion in banking fees. However, setting equal marginal costs across banks may be somewhat problematic, since it entails that prices and market-shares only differ with respect to λ_j. Since both variables are increasing in λ_j, then the larger a bank's share of loyal customers, the higher its equilibrium price and market-share. Consequently, there would be a necessarily positive relationship between prices and market-shares, meaning that the most expensive bank should also be the one with the greatest total amount of customers. However, this does not seem to be empirically true, as Graph 3 shows. So, it is perhaps likely that the marginal cost of supplying the standardized package does in fact differ across banks.

The other possibility we considered was fixing $\lambda_j = \frac{\lambda}{N}$, $\forall j$ and assuming

$$c_j = \exp(\gamma \tilde{a}_j) + \epsilon_j$$

where $\mathrm{E}(\epsilon_j) = 0\ \forall j$ and $\tilde{a}_j \equiv \frac{\text{bank } j\text{'s number of branches}}{1000}$. The moment conditions

Table 3: Results for second empirical exercise (p-values in parentheses).

	$\boldsymbol{\delta}$							
	0.980		0.985		0.990		0.995	
$\boldsymbol{\lambda}$	**b**	$\boldsymbol{\gamma}$	**b**	$\boldsymbol{\gamma}$	**b**	$\boldsymbol{\gamma}$	**b**	$\boldsymbol{\gamma}$
0.2	451.420	-43.176	601.900	-15.285	902.870	-66.270	1805.800	-152.280
	(0.001)	(0.997)	(0.001)	(0.999)	(0.001)	(0.996)	(0.001)	(0.994)
0.4	338.590	-24.765	451.420	-43.176	677.140	-29.682	1354.300	-78.779
	(0.001)	(0.998)	(0.001)	(0.997)	(0.001)	(0.997)	(0.001)	(0.995)
0.6	225.700	-67.375	300.970	-48.644	451.420	-43.176	902.870	-66.270
	(0.001)	(0.996)	(0.001)	(0.996)	(0.001)	(0.997)	(0.001)	(0.996)
0.8	112.880	-88.078	150.460	-43.337	225.700	-67.375	451.420	-43.176
	(0.001)	(0.995)	(0.001)	(0.999)	(0.001)	(0.996)	(0.001)	(0.997)

become:

$$E\left[\begin{array}{c} e_j - \frac{bN(1-\delta)}{(N-1)(1-\lambda)} - \sum_{i=1}^{N} \frac{\exp(\gamma \tilde{a}_i)}{2N-1} - \frac{(N-1)\exp(\gamma \tilde{a}_j)}{2N-1} \\ q_j - \frac{1}{N} - \frac{(N-1)(1-\lambda)}{bN(2N-1)(1-\delta)}\left(\sum_{i=1}^{N} \frac{\exp(\gamma \tilde{a}_i)}{N} - \exp\left(\gamma \tilde{a}_j\right)\right) \end{array}\right] = \mathbf{0}$$

We estimated values for b and γ, for λ ranging from 0.2 to 0.8 in 0.2 intervals and δ ranging from 0.980 to 0.995 in 0.005 intervals. The results for this empirical exercise are shown in table 3. Note that the estimate for the search cost distribution parameter b varies considerably, in response to changes in the values fixed for δ and λ. Under this second formulation of the model, b assumes much larger values (from 100 BRL to 1800 BRL, roughly), which are now also statistically significant. As expected, for a same value of λ, the estimate for b increases as we increase δ, since larger search costs are needed to explain the same amount of price dispersion if consumers' valuation of the future increases. On the other hand, keeping δ fixed, when we increase λ the estimated value for b falls, as searchers become relatively less abundant than loyal consumers and consequentely search costs become less relevant in explaining the banks' price-setting decisions. The fact that estimates of b are sensitive to variations to the value in which λ is fixed perhaps indicates that consumer loyalty, as well as search costs, must also be explaining dispersion in banking fees. In that case, it should also be estimated.

Even though the estimated values of γ are not so relevant to our model, it is interesting to note that $\gamma < 0$. Intuitively, it should be comparatively less costly for banks with a large quantity of branches to accomodate one extra client.

Hence, when $\gamma < 0$, marginal cost is decreasing in the bank's total number of branches, consequently implying a negative relationship between price and number of branches, since lower marginal costs result in lower prices. If we refer to Graph 4 we note that this seems to be compatible with what is empirically observed.

Comparing the results obtained for the two empirical exercises described above, we come to the following conclusions: firstly, according to our empirical observations, we should have heterogeneity in banks' marginal costs, thus making the second formulation of the model seemingly more "compatible" with what we observe. However, this does not mean that consumer loyalty should be discarded as a factor influencing banks' pricing decisions. Indeed, as we had conjectured when elaborating the theoretical model in the previous section, there seems to be evidence that both search costs and consumer loyalty could combinedly be causing the observed price dispersion.

3.1 Estimating Consumer Welfare Loss with Search

The model yields an equilibrium in which costly search occurs. Since the costs consumers incur in terms of search is simply lost and not appropriated by any other agent in the economy, deadweight loss is imposed on them— the socially efficient outcome would have all banks charging the same amount for their services, and no search taking place. The quantity of banks a consumer expects to sample depends on the realization of his search cost. One whose search cost is greater than or equal to d_N will not search at all. Similarly, one whose search cost is smaller than d_N but not less than d_{N-1} will only search for another bank if he samples the Nth bank. Another whose search cost lies in the interval $[d_{N-2}, d_{N-1})$ will only search for lower prices if he samples the $(N-1)th$ or the Nth bank, and so on. Hence, consumers' decision to engage in search activity can be characterized the following way, according to their search costs:

$$\text{Probability of choosing the sampled bank} = \begin{cases} \frac{1}{N} & \text{if } 0 \leq s < d_2 \\ \frac{2}{N} & \text{if } d_2 \leq s < d_3 \\ \ldots & \\ \frac{N-1}{N} & \text{if } d_{N-1} \leq s < d_N \\ 1 & \text{if } s \geq d_N \end{cases}$$

Observe that it is as if the number of costly samples taken by a consumer followed a geometric distribution. So, the number of times consumers expect to

Table 4: Estimates for DWL for a 0.995 discount factor

λ	DWL (BRL)
0.2	233.71
0.4	311.63
0.6	467.44
0.8	934.92

engage in costly search is given by:[15]

$$\text{Expected number of costly samples } = \begin{cases} N-1 & \text{if } 0 \le s < d_2 \\ N-2 & \text{if } d_2 \le s < d_3 \\ ... & \\ 1 & \text{if } d_{N-1} \le s < d_N \\ 0 & \text{if } s \ge d_N \end{cases}$$

Finally, we arrive at the following expression for the total expected deadweightloss consumers face:

$$DWL = \frac{1}{2b} \sum_{j=1}^{N-1} (N-j)\left(d_{j+1}^2 - d_j^2\right) \tag{16}$$

In order to estimate deadweight loss imposed onto consumers by search costs within the banking industry, we substitute the estimation results for b obtained in the second empirical exercise and the observed prices into equation (16). As an example, we have fixed $\delta = 0.995$. The corresponding estimates for the deadweight loss, depending on the value of λ, are shown in table 4.

[15] Note that we do not consider the first sample, since it imposes no search cost on the consumer.

4 Conclusion

We propose a simple model that makes use of consumer search cost and loyalty, to generate equilibrium price dispersion in a market where banks compete for the supply of a homogenous package of services. Based on banks' June 2011 prices for the standardized package, as well as market-shares and number of branches, we try estimating the search cost distribution parameter for the brazilian banking market. In the absence of suitable (and numerous) instrumental variables, we are rather limited and can only perform empirical exercises, in which we estimate some parameters by having to fix others. Unfortunately, we are not able to let both marginal cost and proportion of loyal customers vary across banks at the same time. When fixing $\lambda_j = \frac{\lambda}{N}\ \forall N$, we obtain estimates for the search cost distribution parameter that range from 112.88 BRL to 1805.80 BRL, depending on the values set for λ and δ. We observe quite a large variability of b in response to variations in λ (for a given discount factor), perhaps indicating that the proportion of loyal customers does seem to play an important role in the brazilian banking sector and, like the distribution search cost parameter, it should also be estimated. Thus, in view of the model's non-identifiablity, the next step would naturally involve seeking suitable instruments. However, this task is far from trivial, especially because the moment conditions are not linear in some of the parameters, and we do not know exactly what is contained in the error terms. Nonetheless, even within the limitations of our empirical exercises, we can use the results to assess how well our proposed model fits the banking sector. For instance, if we fix the value of the discount factor in 0.995 and assume the proportion of searchers is 80%, we find that the search cost distribution parameter is 1805.80 BRL. The corresponding estimate for the deadweight loss incurred by an average searcher is 233.71 BRL. These values seem unrealistically high, indicating that there might be other market frictions, besides consumer loyalty and search costs, causing the empirically observed price dispersion across banking fees. Some immediate candidates are switching costs and differences in the quality of banks' services, both of which we had chosen not to model. It would be interesting to extend our model to include switching costs, for instance. However, this would probably add more parameters to the model and further aggravate the identifiability problem.

References

CARLSON, J. A.; MCAFEE, R. P. Discrete equilibrium price dispersion. **Journal of Political Economy**, v.91, pp.480-93, jun., 1983.

CHAUSSÉ, P. Computing generalized method of moments and generalized empirical likelihood with R. Working paper. Département des Sciences Économiques de L'Uqam. Available in:
<http://www.er.uqam.ca/nobel/k34115/images/gmm_with_R.pdf.>

DAHLBY, B.; WEST, D. S. Price dispersion in an automobile insurance market. **Journal of Political Economy**, v.94, no.2, pp.418-438, 1986.

SANTOS, B. de los, HORTAÇSU, A.; WILDENBEEST, M. R. Testing models of consumer search using data on web browsing and purchasing behavior. 2009 Working Paper. Kelley School of Business and University of Chicago. Available in: <http://www.kelley.iu.edu/mwildenb/searchtest.pdf.>

HANSEN, L. P. Large sample properties of generalized method of moments estimator. **Econometrica**, v.50, pp.1029-1054, jul., 1982.

STAHL, D. O. Oligopolistic pricing with sequential consumer search. **American Economic Review**, v.79, no.4, pp.700-712, sep., 1989.

VARIAN, H. R. A model of sales. **American Economic Review**, v.70, no.4, pp.651-659, sep., 1980.

FECOMERCIO. As novas regras das tarifas bancárias: correções ou distorções? 2009. Available in:
<http://www.sincomercioata.com.br/Legislacoes/Arquivos/cart_tarifas_fecom.pdf>

BANCO CENTRAL do BRASIL: database. Available in:
<http://www.bcb.gov.br/?SFN>

Appendix

In order to estimate the parameters using GMM, the *gmm* package in the statistical program R[16] was used. The commands are listed below:[17]

```
#########GMM commands
#####First version: equal marginal costs; lambda depends on %branches:
n<-7
#delta
des<-0.995
#1-delta
d<-1-des
# marginal cost
c<-2
#removing citibank from the sample
citimsag<-0.007022934
#banks' participation in total quantity of branches
msagnciti<-c(0.207944694, 0.122627016, 0.279545704, 0.047624273, 0.133106551, 0.202128827)
#prices
pnciti<-c(10.5, 15, 13.5, 13.5, 14, 12.5)
#market-shares
mscontasnciti<-c(0.300908397, 0.082339472, 0.340837936, 0.062725801, 0.015790777, 0.19288138)
#gathering data
dados<-cbind(mscontasnciti, pnciti, msagnciti)
# moment equtions
g2<-function(tet, x) {
m1<-((1/n)+((n-1)*tet[2]*(x[,3]-(1/n))/((2*n)-1))-x[,1])
m2<-(d*n*tet[1]/(n-1))+(d*(n^2)*tet[1]*tet[2]*(x[,3]+(1/(n-1)))/(((2*n)-1)*(1-tet[2])))+c-x[,2]
f<-cbind(m1, m2)
return(f)
}
#gradient matrix
dg2<-function(tet, x) {
G2<-matrix(c(0, d*((n/(n-1))+(tet[2]*(n^2)*(2-citimsag)/((n-1)*(1-tet[2])*((2*n)-1)))), ((1/n)-citimsag)/((2*n)-1),d*(tet[1]*(2-citimsag)*(n^2)/((n-1)*((2*n)-1)*((1-tet[2])^2)))), nrow=2, ncol=2)
return(G2)
}
```

[16] www.r-project.org/

[17] See Chaussé in URL http://www.er.uqam.ca/nobel/k34115/images/gmm_with_R.pdf

```
res2<-gmm(g2,dados,c(6,0.6),grad=dg2)
summary(res2)

#####Second version: constant lambda(j); c depends on exp #branches
# we will remove the equations for citibank
# number of branches/1000 citibank:
citimsag<-128/1000
# building up the data matrix:
# number of banks
n<-7
# market-shares
mscontasnciti<-c(0.300908397, 0.082339472, 0.340837936, 0.062725801, 0.015790777,
0.19288138)
# number of branches
agnciti<-c(3790,2235,5095,868,2426,3684)
# prices
pnciti<-c(10.5, 15, 13.5, 13.5, 14, 12.5)
# gathering data
dados2011nciti<-cbind(mscontasnciti, pnciti, agnciti/1000)
# discount factor
des<-0.99
d<-(1-des)
# lambda
lambda<-0.4
l<-1-lambda
# moment conditions:
g2<-function(tet, x) {
m1<-((1/n)+((n-1)*l*((sum(exp(tet[2]*x[,3]))/n)-exp(tet[2]*x[,3]))/(tet[1]*n*((2*n)-
1)*d))-x[,1])
m2<-(d*n*tet[1]/((n-1)*l))+(sum(exp(tet[2]*x[,3]))/((2*n)-1))+(exp(tet[2]*x[,3])*(n-
1)/((2*n)-1))-x[,2]
f<-cbind(m1, m2)
return(f)
}
# gradient matrix:
dg2<-function(tet, x) {
G2<-matrix(c(-l*((exp(tet[2]*citimsag))-(sum(exp(tet[2]*x[,3])/n)))/(d*(tet[1]^2)*n*((2*n)-
1)), n*d/(l*(n-1)), l*((citimsag*exp(tet[2]*citimsag))-(sum(x[,3]*exp(tet[2]*x[,3]))/n))/(tet[1]*d*n*((2*n)-
1)), ((2*sum(x[,3]*exp(tet[2]*x[,3])))-(citimsag*exp(tet[2]*citimsag)))/((2*n)-1)),
nrow=2, ncol=2)
return(G2)
}
res2<-gmm(g2,dadosnciti,c(600,-3),grad=dg2)
summary(res2)
```