

Data-Driven Approach for Dynamic Pricing for Decision Making Systems in Marketing and Finance

Petr Gladilin, Irek Saitov
ITMO University
Saint-Petersburg, Russia
peter.gladilin@gmail.com, xanilegendx@gmail.com

Abstract—Recently, the use of accumulated data in the trade and financial sector has become increasingly important to improve the quality of services provided. For successful business, decision-making systems in retail should combine various marketing strategies with analysis of available data. For suppliers of goods and services and banks acquirers the forecast of the success of outlets and customer's churn prediction are crucial, especially when price policy changes. We propose data-driven approach that evaluates clusters of the stores based on the success of each segment of suppliers and develop the model for assigning the optimal bank's acquiring rate within the shopping center taking into account different customer's loyalty rate accordingly to their success. In this study we analyze the data on consumer's activity inside large shopping centers on the basis of the transactional dataset provided by our partner bank. We highlight the segments of outlets based on their location, category of goods and services, average monthly turnover and the number of transactions and apply the constructed model to recommend the optimal rate. The proposed approach can be applied in the retail sector by large suppliers of goods and services and banks.

I. INTRODUCTION

As it states from World Payments Report 2018 in the last year, the volume of non-cash transactions almost doubled the number of cash withdrawal transactions from bank cards [1]. This indirectly confirms the statement, which analysts have been talking about for a long time that money becomes "electronic".

However, given the ever-increasing importance of electronic data for business in e-commerce, the significance of using in-depth knowledge of market features and the characteristics of its main players - customers, suppliers and competitors - also increases. In order to sale best business marketers should analyze what market segment has the highest willingness to buy and to pay and what specific price level will generate the highest market share, revenue and profits [2]. Detailed segmentation based on socioeconomics, value drivers, decision drivers and loyalty to pay more generates the most important data about the market [2], [4].

The modern approach to the differentiated pricing strategy involves wider use of the available information that could be collected from the retailers for the further assigning of different prices for the same goods and services in relation to different categories of customers or at different times [2], [3]. Such strategies allow suppliers to regulate demand throughout the year (weeks, days) and enable them to work with the most

profitable market segments. The era of big data challenges new tasks associated with efficient aggregation, processing and interpretation of large amounts of weakly formalized data. The main tasks closely related to the analysis of data in the socio-economic sphere are associated with predictive analytics and forecasting in retail, performing optimal marketing strategies, customer's churn prediction models etc. [2], [6].

One of the tasks of price discrimination is to set prices that will allow you to sell the product to those categories of customers who would not buy it at higher prices. And, at the same time, increase them for segments with high purchasing power. Using data-driven approach it is possible to analyze the dynamics of sales, comparing it with statistics accumulated over the previous time, and to take into account many factors in pricing. But the idea remains the same - to maximize company profits.

Most dynamic pricing models involve complex mathematical calculations and computer programming [5]. The basis of dynamic pricing is the use of information about various characteristics of customers (outlets) and related market conditions. Based on this information, the optimal interest rate for each client group selected using cluster analysis is established, which provides the maximum income taking into account the probabilities of an outflow of customers from target groups. If the margin is too high, they would go to the competitor, and in case of too low margin, they could pay more for the same offer.

The main objective of this study is to develop a data-based approach to optimize marketing strategies and bank acquiring systems by the example of shopping centers. The rest of the paper is organized as follows. Section II presents related work on methods which were applied to solve similar problems earlier. Section III gives formal description of the available data on which we made extrapolation of the data and further modeling, Section IV presents the mathematical formalization of the proposed dynamic loyalty model. Section V is for the methodic of experimental study and experimental results. And finally, Section VI presents conclusion and discussion.

II. RELATED WORK

The dependence of consumer sensitivity to price changes was discussed in [4], [6], [7] on the basis of historical data on relations between suppliers and consumers.

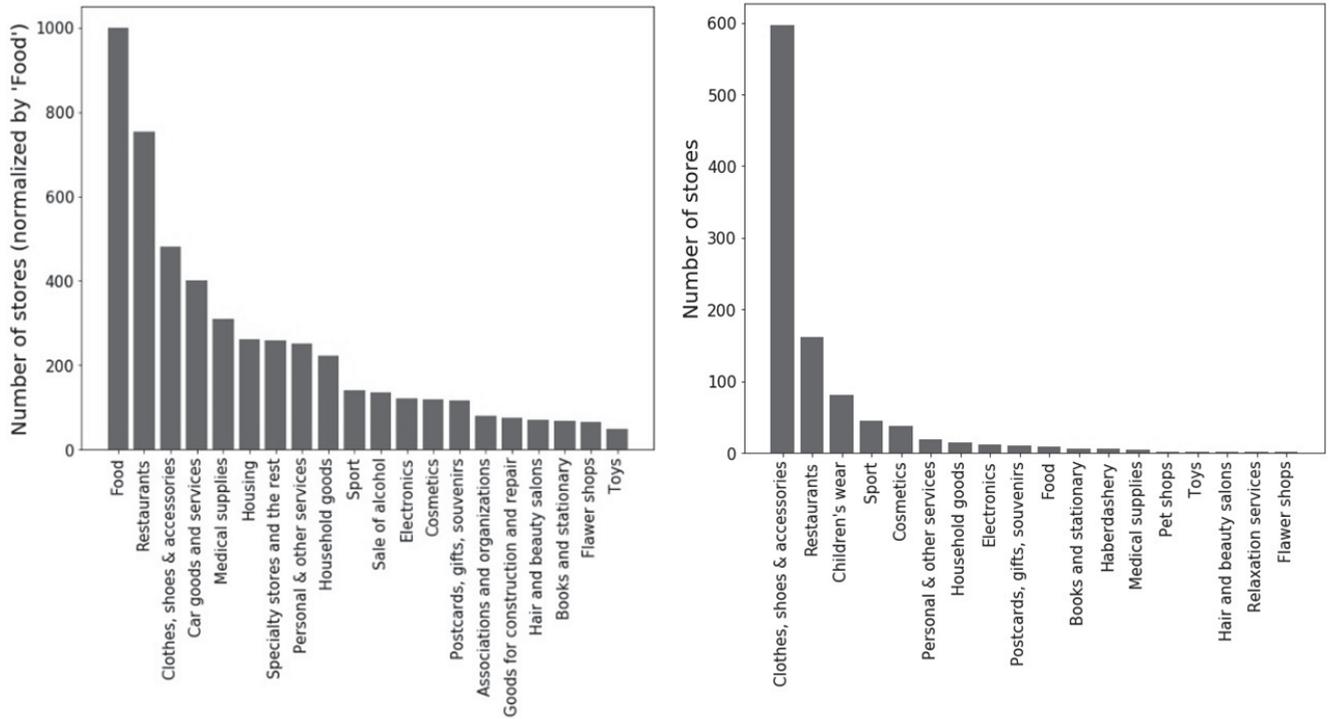


Fig. 1. Distribution of the number of the stores in different categories: *left*: in total for Northwest region of Russia, *right*: for 10 largest shopping centers in St. Petersburg

Main ideas of the customer’s satisfaction analytics can be summarized as follows: (i) the longer the relationship, the higher the propensity to use a certain product of the company; (ii) the longer the relationship, the less sensitivity to price changes for a certain product of the company; (iii) the greater the Relationship Breadth (RB, the amount of goods delivered to the customer), the higher the propensity to use a particular product of the company, but depending on the duration of the relationship; (iv) the greater the RB, the greater the sensitivity to price changes for a specific product. Nevertheless, historical data on customer interactions are often either inaccessible or insufficient to provide a comprehensive analysis and identify the degree of customer satisfaction. Providing special marketing research or special surveys is also often not applicable, therefore, it is necessary to use an available data to form the optimal marketing strategies.

The degree of customer satisfaction is directly related to the price elasticity of this customer segment [5]. Many studies investigating customer’s churn and the dependence of customer’s satisfaction on the existing pricing policy (including discussed above) used data accumulated over a large historical period [4], [7], [8]. This made it possible to investigate the response of various customers to the changes in pricing strategy or targeted offers, solving the classification or regression problem [6].

In this study we propose an approach dealing with customer churn and dynamic pricing strategy in the case where complete historical data on the interaction of retail outlets with banks (or any suppliers of goods and services) is not available. The

proposed model is based on aggregated transactions dataset provided by our bank-partner, supplemented by data from the web-sites of banks-acquirers. Such data in most cases is the only thing analysts can work with, therefore the development of such a decision-making model is a popular business task.

III. DATA AND MODELING

A. Data

1) Shopping centers characteristics

We use the dataset provided by our bank-partner (one of the largest regional banks in Russia) to extract the information about electronic purchases in 10 largest shopping malls $T^{(i)}$, $i=\{1, \dots, 10\}$ in Saint-Petersburg during the last 5 years. The binding of transactions to a specific shopping center was made on the basis of the addresses of the shopping center and the geo-location of the successful transaction. The data was extracted from the following 10 shopping centers: “Galeria”, “Europolis”, “Mercury”, “MEGA Dybenko”, “MEGA Parnas”, “Piterland”, “LETO”, “City Mall”, “Continent Zvezdnaya” and “Nevskiy Center”. Given information about the purchases in different stores $O^{(ij)}$, $i=\{1, \dots, 10\}$, $j=\{1, \dots, N(i)\}$ from the specific shopping center $T^{(i)}$ all the transactions were merged into 18 groups accordingly to the Merchant Category Codes (MCC), which were assigned to the category of each transaction.

Fig. 1 represents the distributions of the number of the stores in specific category summarizing the data: (i) in total for every transaction in the dataset (to the left) and (ii) only for the

extracted transaction in the largest shopping centers of Saint-Petersburg (to the right). One can see that the most popular (in terms of the number of stores) categories significantly differs in these two cases: in shopping centers, the most popular categories are “Clothing, shoes & accessories” stores, “Restaurants”, and “Children's wear”, while in the whole dataset markets mostly correspond to the “Food” category.

We added new values to the existing characteristics of transactions in shopping centers: the distance from the nearest metro station, reflecting the geo-availability of the outlet, the average number of transactions in a particular category in the store, and the average monthly transaction value in a given category.

For clustering and highlighting segments of retail categories in shopping centers, we have used data with extracted set of 62 “category in the mall” groups of retail stores in 10 shopping centers. Next, using implementation of the DBSCAN algorithm by scikit-learn Python library, 5 groups of stores were identified. 2D representation of these clusters made by the tSNE algorithm is shown in Fig. 3.

Taking into account marketing research described above (see discussions in [3], [4], [6], [7]), we have marked the segment with a high average monthly turnover and top categories (big black markers on the Fig. 3) as a segment with greater loyalty to increasing the rate of acquiring, and the remaining groups as a segment with reduced loyalty (“Others”).

B. MODELING

1) Calculation of acquiring rates

The rate of payment (authorization) consists of: interchange fee (commission re-turned to the bank that issued the card), commissions charged by International Payment Systems and the mark-up of the acquiring bank and. Thus the bank’s revenue is determined only by the last term. As the rate of payment and interchange fee can vary a lot depending on the bank’s marketing strategy and existing agreements between banks in this study we will evaluate full rate of payment which should pay the store to the bank-acquirer for each transaction.

Russian banks have different strategies for assigning the acquiring rate and very few of them take into account the categories of goods and services which are of the point of interest for the specific store. The most popular characteristic which affects the acquiring rate as the bank offer for retailers is the monthly turnover of the specific store. In the Fig. 2 we collect the open sources data from the Top-20 largest Russian bank’s web-sites about the dependence of the acquiring rates on the turnover of specific store. It can be seen that there is a clear tendency to a decrease in the rate with an increase in

$$P^{(i)}(r, k_S, p_S) = p_S + \frac{1}{k_S}(r - r_0^{(i)}) \quad (1)$$

monthly turnover.

2) Stores segmentation

To evaluate monthly turnover for the specific store $O^{(i)}$ in the shopping center $T^{(i)}$ we have taken into account information about the market share for our bank partner, and have multiplied the cumulative turnover value for the store in the specific category in our data by this coefficient. The result of evaluation shows that 96% of the outlets in the shopping centers have relatively large turnover (with > 300 000 rubles per month).

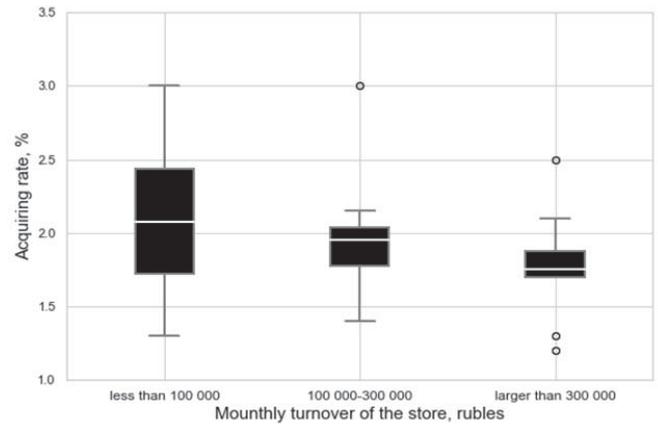


Fig. 2. The distribution of acquiring rates of 20 largest Russian banks for retail stores in dependence on its monthly turnover. Data was obtained from public sources and official websites of banks

The study [8] proposes an approach that relates tolerance to price change with an indicator of customer satisfaction. And since customer satisfaction in the case of store-bank interaction directly depends on sales, it is reasonable to associate the value of the acquiring rate with the turnover of the point of sale, which we observe in bank’s practice and in the experimental results of clustering in Section III.

2) The model of dynamic loyalty

We develop the mathematical model to evaluate the optimal rate for the specific segment of retailers in the shopping center. This model must meet the following conditions: (i) acquiring rate should be different for retail outlets from different network retail segments; (ii) the probability of leaving for the customer (point of sale) increases with an increase of the acquiring rate; (iii) the loyalty of retailer’s outlets to the increase of the acquiring rate differs for different segments.

In order to distinguish between the segment of loyalty stores (customers of the bank acquiring system) and the segment of customers who violently reacting to any changes in pricing policy we introduce the loyalty rate k_S . This rate should be greater than 0, and the case when k_S is close to 0 corresponds the situation when the segment is intolerant for the price price changes, and will almost certainly go away when it rises. Bank’s proceeds from the store $O^{(i)}$ from the shopping center $T^{(i)}$ will thus depend on the loyalty rate k_S , which corresponds to this store. The probability of the store’s churn $P^{(i)}$ for the specific loyalty segment S is:

where r_0 – current acquiring rate for the store $O^{(i)}$, r – chosen acquiring rate, p_S – is the initial probability of the outlet’s churn.

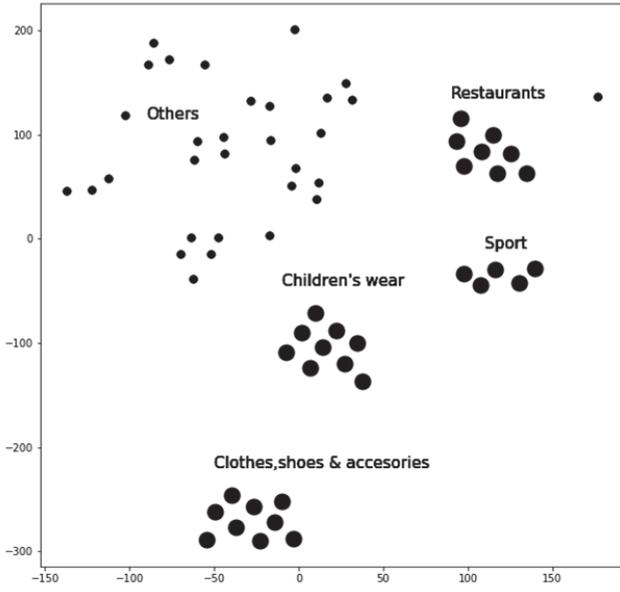


Fig. 3. Experimental result of DBSCAN clustering algorithm and further tSNE 2D projection of store's segments in ten shopping centers. Each point on the graph correspond to the "category in the mall" group and may include different number of the stores. The graph clearly distinguish the segments of the stores according mostly to their category and monthly turnover. On the graph, small markers indicate groups of outlets with a potentially low loyalty factor (low turnover and a small number of transactions), large markers indicate outlets with a potentially high loyalty factor (high turnover and a large number of transactions)

Thus, the optimal rate for a specific segment should maximize the bank income and can be evaluated as follows:

$$r_{opt}^{(i)}(k_S, p_S) = \arg \max_r (F_S^{(i)}(r, k_S, p_S)) \quad (2)$$

where r_{opt} – is the optimal acquiring rate for the segment S and $F_S^{(i)}(r, k_S, p_S)$ – is the sum of bank incomes from the customers of the segment S.

The maximally achievable value of the bank's revenue $F_{max}^{(i)}$ from the shopping center $T^{(i)}$ should be evaluated as follows:

$$F_{max}^{(i)} = \sum_S F_S^{(i)}(r_{opt}, k_S, p_S) \quad (3)$$

The model formulation (1) reflects the idea that even with no change in acquiring rate there is still the probability p_S of customer's churn and this value increases proportionally to the change of the acquiring rate r and depends on the loyalty rate k_S . We propose to split all the stores into two segments according to the loyalty rate which is ruled by success of the store, as we have done in Section III (A): "Segment 1" consists of non-loyalty customers (stores with low turnover and low frequencies of transactions) and "Segment 2" consists of successful, and therefore, potentially loyalty stores, which are

almost from the categories "Clothes, shoes & accessories", "Children's wear", "Restaurants" and "Sports".

Thus, we propose the following algorithm for evaluating the optimal acquiring rate for the specific segment S of customers reflecting main ideas of the described model of dynamic loyalty:

Algorithm 1 Algorithm for calculating the optimal acquiring rate for the customers segment S

Input: $r_0^{(i)}$ ← current acquiring rate for the category of stores,
 p_S ← initial probability of the outlet's churn,
 k_S ← loyalty rate,
 R ← array of rates,
 f ← income from one store in this category in the shopping center,
 n ← number of stores of this segment

For r **in** R :

$$F_S^{(i)} \leftarrow n \times r \times f$$

$$P^{(i)} \leftarrow p_S + \frac{1}{k_S} (r - r_0^{(i)})$$

For j **in** $\text{range}(0, n)$:

$$p \leftarrow \text{random}(0, 1)$$

If $p < P^{(i)}$:

$$F_S^{(i)} \leftarrow F_S^{(i)} - r \times f$$

$$r_{opt}^{(i)} \leftarrow \arg \max_r (F_S^{(i)})$$

Output: $r_{opt}^{(i)}$

Fig. 4 shows the dynamics of the normalized revenue depending on the increasing acquiring rate for the three basic cases: loyal customers, neutral customers and disaffected customers. Loyal customers are positive and loyal to raising prices - this category of customers is most suitable for applying the approach described here. The second category of neutral customers has practically no effect (up to a point) on revenue - for them, revenue growth is exactly offset by an outflow of customers due to price increases. And finally, the third category - disaffected or frustrated customers - for them, any increase in the rate will lead to their refusal of service and transfer to a competitor.

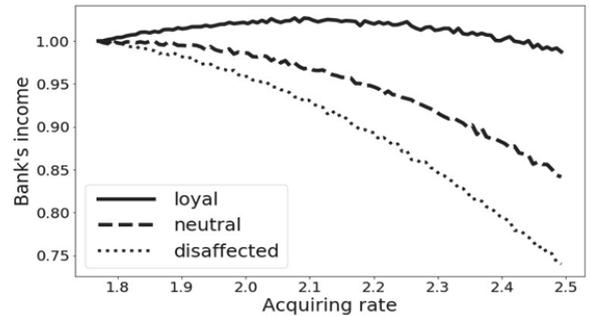


Fig.4. Revenue dynamics due to the acquiring system for three customer segments with different loyalty rate

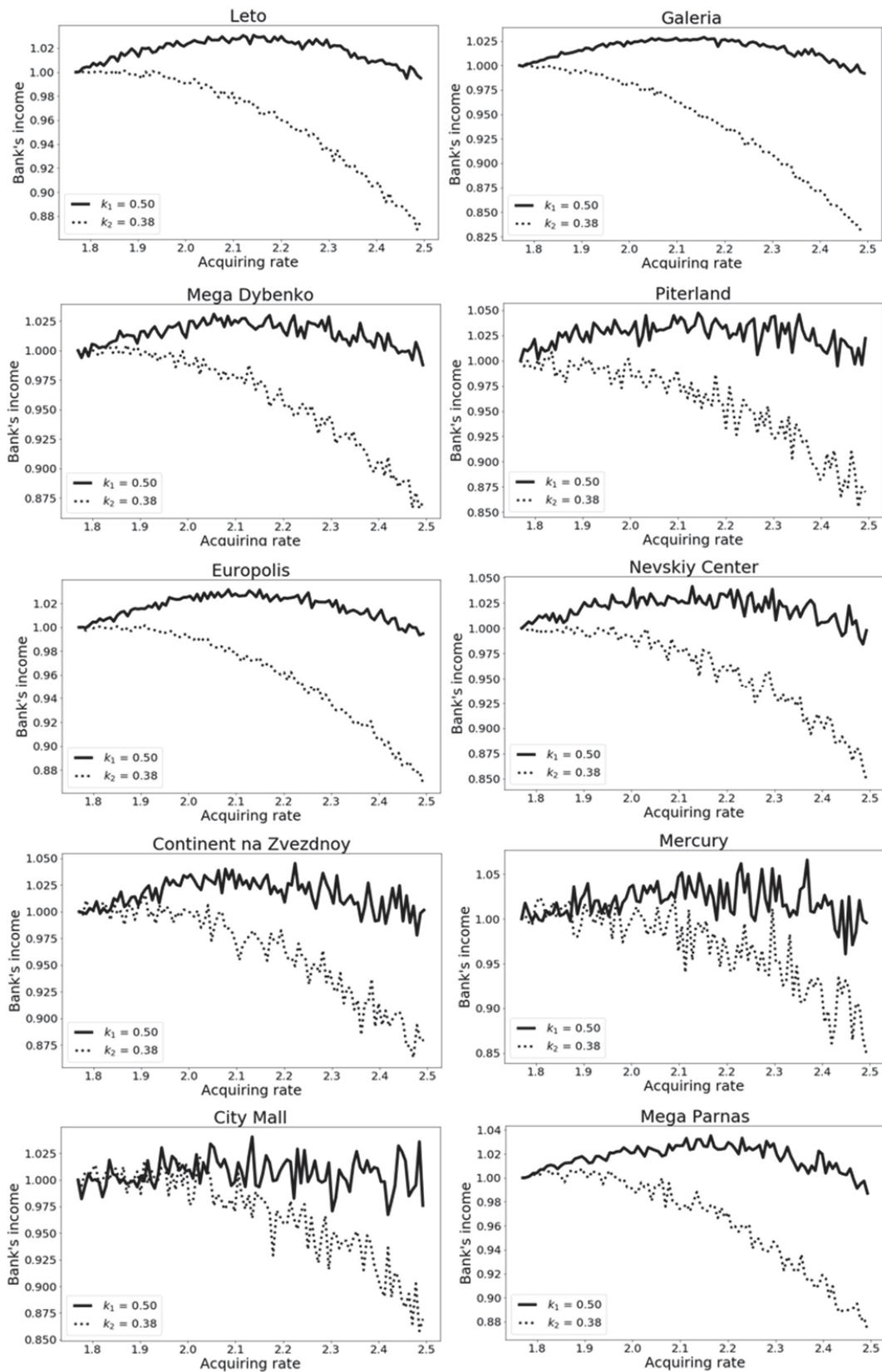


Fig. 5. Experimental results of the probabilistic model for bank's income (relative values) from the acquiring system implemented in 10 largest shopping centers in St. Petersburg, depending on the cost of acquiring for two segments of stores-customers with different loyalty rate to price changes. *Bold line* corresponds to the loyal segment of the stores and *dashed line* is for the non-loyal ones

IV. RESULTS AND DISCUSSION

We have simulated the change in the bank's income $F_S^{(i)}(r, k_S, p_S)$, from the acquiring system in the shopping center, subject to a price increase within the selected range from the two proposed in the Section III loyalty segments of the stores.

Fig. 4 represents experimental results of the modeling set for 10 shopping centers regarding increase of the acquiring rate starting from the average rate $r = 1.8\%$ for the stores with turnover greater than 300 000 rubles per month (see Fig. 3).

We have assigned $k_1 = 0.50$ and $k_2 = 0.38$ for the loyal and non-loyal segments respectively. The value $k_1 = 0.50$ is chosen in such a way that the model reflects the case when the maximum rate in the market is reached (for the selected range of values, it is $r_{\max} = 2.5\%$), the outflow of customers begin to suppress the increase in revenue from the rate increase. Such a choice of k_1 guarantees the presence of a local maximum of revenue corresponding to the optimal interest rate. The value $k_2 = 0.38$ for this data corresponds to the threshold for the model, which states that non-loyalty segment almost certainly leaves the supplier as the price increases. Such a behavior one can see on every chart on the Fig. 5 (dashed line). These coefficients are analogous to well-known price elasticity coefficients, taking into account possible different customer behavior [9]. A distinctive feature of our approach is that these coefficients are estimated not by special marketing research, but by analyzing the available data.

Following the results of the clustering analysis of various customer segments carried out in study [6] the probabilities of the initial outlet's churn p_S were equated to 0.05 for non-loyal ($s = 2$) and to 0.04 for the loyal segment ($s=1$). Ten charts in the Fig. 5 show experimental results of an average of 500 epochs of the evaluating bank income by Eq. 1 with random choice of the acquiring rate r within the range corresponding to the large turnover group. In every step of the rise of the acquiring rate r the bank income $F_S^{(i)}(r, k_S, p_S)$ tend to increase in accordance with Eq.1 but at the same time, the probability $P^{(i)}$ of customer outflow and loss of profit increases. Thus, depending on the loyalty rate, the dependence of the acquirer's bank's profit on the rate when it is raised will have a convex shape with a maximum corresponding to the optimal profit-maximizing acquiring rate $r_{opt}^{(i)}(k_S, p_S)$. At the same time, with a greatly increased or greatly reduced loyalty rate, the graph will degenerate, respectively, into a monotonically increasing or monotonically falling function.

The jitter on some charts is caused by a small number of categories of outlets represented in this shopping center, so the likely churn of one of the customer-stores greatly affects the total income in this case.

Experimental results on the Fig. 5 show that for the specific set of shopping centers (as "Galeria", "Leto", "Europolis", "Mega Dybenko") competent choice of acquiring rate can increase the revenue from interaction with the loyal customers segment by 2-3%, while the rate itself will change by five tenths of a percentage point. By analyzing such charts, bank analysts or marketers can evaluate the allowable range of price changes (interest rates) depending on the goals of the business.

V. CONCLUSION

In this study we propose an approach for stratified assignment of acquiring rates for the specific stores in shopping centers. The model of dynamic loyalty takes into account the data about the distribution of the bank acquiring rates for retail outlets with a certain turnover and calculates how much the bank can raise the rate, considering the prospective customer loyalty.

We have demonstrated that the weighted average turnover is the most important feature of the outlet for being distinguished from the others. However, average transaction numbers and the other characteristics of the store (like the distance from the underground station, average transaction rate for the mall, etc.) are not an important factor for the segmentation. This fact confirms the correctness of the banks approach to assign the rate of acquiring based on the turnover of the store.

This approach has natural limitations associated with the need for a sufficient amount of data on outlets for clustering and intelligent selection of a segment to determine the optimal bidding strategy. In addition, the data approach is applicable for sufficiently large network retailers and banks, for which there are a large number of customers, which is necessary from a statistical point of view for the adequate functioning of the model. The model described in Section III can be generalized to the case of a possible price reduction while taking into account the likely influx of new customers. This approach requires additional market research and analysis of related areas with similar product categories of outlets served by competitors to calculate the likelihood of their transition to service at a better price.

To increase the accuracy of this approach it seems appropriate further to collect additional information about the state of the business, customer reviews, financial reporting in the public domain, etc. Results discussed in this study may be useful in bank Research & Development Departments in order to improve the services and increase the revenue of the bank as well as auditors engaged in business analytics and planning the optimal development strategy for the customer companies. But it is necessary as well to take into account the price of profits increase due to this approach - in terms of data mining cost, software developing and so on.

The proposed approach is applicable not only for adjusting the optimal interest rate for acquiring a bank, but also for network retailers with a large customer base and outlets with accumulated data on attendance and sales for each of them. The method of clustering and subsequent pricing of the product, maximizing profits from this group of customers, taking into account the possible outflow, can be effectively used in marketing analytics and building an optimal business strategy for a network retailer.

ACKNOWLEDGMENT

This work financially supported by Ministry of Education and Science of the Russian Federation, Agreement #14.575.21.0165 (26/09/2017). Unique Identification RFMEFI57517X0165.

REFERENCES

- [1] Official website of The World Payments Report, Web: <https://worldpaymentsreport.com/>
- [2] D.B. Orekhov, "Economic model of the shopping mall: marketing aspects. Marketing problems, logistics", *Problems of the modern economy*, vol.4(52), 2014, pp. 181-184. (in russian)
- [3] K.O. Bochenina, A.V. Boukhanovsky and J.A. Holyst, "Financial information technology decision support systems based on big data", *Control Engineering*, vol. 3(81), 2019, pp. 20-24. (in russian)
- [4] A. Tamaddoni, S. Stakhovych and M. Ewing, "Comparing Churn Prediction Techniques and Assessing Their Performance: A Contingent Perspective", *Journal of Service Research*, vol.19(2), Nov.2016, pp. 123-141.
- [5] P. Kotler, *Marketing Management*, sixth edition, New Jersey, Prentice-Hall Inc., 1988.
- [6] A.C. Bahnsen, D. Aouada and B. Ottersen, "A novel cost-sensitive framework for customer churn predictive modeling", *Decision Analytics*, vol.2, June 2015, pp. 1-15.
- [7] E.W. Anderson, "Customer Satisfaction and Price Tolerance", *Marketing Letters*, vol.7, July 1996, pp. 265-273.
- [8] J. Dawels, "The Effect of Service Price Increases on Customer Retention", *Journal of Service Research*, vol.11, Feb.2009 pp. 232-245.
- [9] R.H. Frank, *Microeconomics and Behavior*, seventh edition, New York, McGraw-Hill, 2008.
- [10] T. H. A. Bijmolt, H. van Heerde and R. Pieters, "New Empirical Generalizations on the Determinants of Price Elasticity", *Journal of Marketing Research*, vol.42, Apr.2005, pp. 141-156.