NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

Elena Pokryshevskaya, Evgeny Antipov

IMPORTANCE-PERFORMANCE ANALYSIS FOR INTERNET STORES: A SYSTEM BASED ON PUBLICLY AVAILABLE PANEL DATA

BASIC RESEARCH PROGRAM

WORKING PAPERS

SERIES: MANAGEMENT
WP BRP 08/MAN/2013

Elena B. Pokryshevskaya[1], *Evgeny A. Antipov*[2]

IMPORTANCE-PERFORMANCE ANALYSIS FOR INTERNET STORES: A SYSTEM BASED ON PUBLICLY AVAILABLE PANEL DATA[3]

We demonstrate how publicly available ratings of Internet stores can be used to perform actionable importance-performance analysis (IPA). We use a panel of 422 non-food online merchants, which allows us to account for store-specific effects, and therefore obtain better estimates of each attribute's influence on purchase intention. Our approach is economical, as merchants do not have to spend money on surveys that firms typically have to conduct for IPA. We have enhanced traditional importance-performance analysis by deriving attribute importance indirectly using the Shapley Value decomposition of a first-difference regression's R^2. We have also modified the standard "importance-performance" graph by adding the information about the directions of changes in attribute ratings between the two periods, which made it easier to point out strategic mistakes. We demonstrate the diagnostic power of our approach by doing the actual importance-performance analysis for two of the stores. The potential of enhancing the expert system using other methods and other types of publicly available data is discussed.

JEL Classification: L81, M31.
Keywords: Shapley value, importance-performance analysis, panel data, e-commerce.

---

[1] National Research University Higher School of Economics (Saint-Petersburg, Russia). Department of Economics. Senior Lecturer; E-mail: epokryshevskaya@hse.ru
[2] National Research University Higher School of Economics (Saint-Petersburg, Russia). Department of Economics. Senior Lecturer; E-mail: eantipov@hse.ru
[3] This study was carried out within "The National Research University Higher School of Economics' Academic Fund Program in 2012-2013, research grant No. 11-01-0193".

1. Introduction

The identification of key service attributes is crucial for quality improvement. From a strategic point of view the natural areas for improvement are related to those attributes that are both important and with which customers are dissatisfied (Martilla & James 1977). In order to identify such important but low-performing attributes importance-performance analysis (IPA) is widely used (Deng et al. 2008; Matzler et al. 2004; Guadagnolo 1985; Deng 2007; Deng & Pei 2009). Usually it is based on customer satisfaction survey data, which is then plotted on a graph, where the horizontal axis represents the importance of every attribute and the vertical axis shows the performance of an attribute. The chart is divided into 4 quadrants:

1. Quadrant 1 (high importance, high performance): attributes from this quadrant are the key drivers of satisfaction. It is recommended to maintain the satisfaction with these attributes at a high level.
2. Quadrant 2 (low importance, high performance): attributes from this quadrant are given too much attention. It is recommended to avoid spending too many resources on their development and if possible, reduce the expenditures on them.
3. Quadrant 3 (low importance, low performance): attributes from this quadrant are minor shortcomings, improving which is not of high priority. It is recommended to delay their improvement up until some point of time in the future.
4. Quadrant 4 (high importance, low performance): attributes from this quadrant require immediate improvement.

Therefore, IPA complements descriptive satisfaction survey information and helps managers to make decisions about the allocation of scarce resources in a way that maximizes customer satisfaction or some other overall performance measure.

In this paper we use IPA framework to develop a system that allows conducting strategic quadrant analysis for Internet stores based on publicly available data. Our study has several features that distinguish it from other research:

1. We use a theoretically sound Shapley value approach to measure attribute importance. As a major limitation, IPA leads to different conclusions depending on how an attribute's importance is figured (Tontini & Silveira 2007). In most past studies direct importance ratings were used (Murdy & Pike 2012; O'Neill & Palmer 2004; Duke & Persia 1996; Smith & Costello 2009; Martin 1995). However, empirical comparisons suggest that derived importance provides better differentiation of attribute importance values (Fontenot et al. 2007; Van Ryzin & Immerwahr 2007). Pokryshevskaya and Antipov (2013) compared nine methods for measuring indirect importance and came to a conclusion that the Shapley value approach for R^2

decomposition can be recommended as a theoretically sound method, and the only one which is among the Top-5 methods, according to both stability and specificity.

2. We use panel data, which enables us to apply a first-difference estimator, which is a sound technique for causal inference. Most existing importance-performance studies use cross-sectional data, whereas with panel data we are able to control for store-specific fixed effects and therefore more properly measure the influence of each attribute on purchase intention. Moreover, IPA graph becomes more informative when shifts from one period to the other are depicted on it.

3. We use publicly available data. Most customer satisfaction and loyalty studies analyze a single company, using proprietary survey data. Instead, we gather publicly available data on a sample of Internet stores, which makes it possible to conduct the analysis with virtually any frequency without spending money on expensive surveys.

2. Data

Data was gathered in September 2010 and November 2012 from www.Bizrate.com. The panel dataset consists of average grades, given by customers to 422 nonfood Internet stores. All grades are on a 10-point scale with 10 being "excellent." We use a relatively strict requirement for the sample size, based on which the average grade is calculated – no less than 250 respondents during the last 3 months. This ensures that the average grade reflects the opinion of a reasonably large number of people.

The dependent variable (Y_i) is a repeat purchase intention characterized by the average agreement with the statement *Would shop here again* for store *i*. Attributes (X_{ij}, $j=1,..,13$) can be classified into three groups (Pokryshevskaya & Antipov 2012):

- Post-order service and satisfaction with claims (*On-time delivery*, *Order tracking*, *Customer support*, *Availability of product you wanted*, *Product met expectations*)
- Internet store usability and design (*Ease of finding what you are looking for*, *Overall look and design of site*, *Clarity of product information*, *Selection of products*)
- Economic considerations *(Shipping charges*, *Charges stated clearly before order submission*, *Variety of shipping options*, *Prices relative to other online merchants)*

The changes in mean attribute ratings are presented in Table 1. On average, companies managed to improve the performance of pre-order attributes (i.e. usability, selection of products, pricing, shipping charges and shipping options). At the same time, mean repeat purchase intention did not change significantly according to Wilcoxon test.

Table 1. Average changes in attribute performance between September 2010 and November 2012 (attributes, which ratings experienced a significant change are typed in boldface).

	Change (points)				Change (%)				Wilcoxon test p-value
	Mean	Standard Deviation	Minimum	Maximum	Mean	Standard Deviation	Minimum	Maximum	
Would shop here again	0,03	0,43	-1,90	2,40	0,53	5,21	-20,00	34,80	0,315
Ease of finding what you are looking for	0,03	0,21	-0,90	0,80	0,41	2,38	-9,40	9,40	0,000
Selection of products	0,05	0,19	-0,70	0,60	0,61	2,13	-7,90	7,50	0,000
Clarity of product information	0,04	0,21	-1,00	0,80	0,51	2,44	-11,40	9,80	0,000
Prices relative to other online merchants	0,11	0,25	-0,80	1,50	1,39	3,03	-9,10	21,70	0,000
Overall look and design of site	0,07	0,22	-0,90	1,10	0,87	2,56	-9,50	14,10	0,000
Shipping charges	0,12	0,47	-2,50	2,50	1,66	6,40	-26,60	43,90	0,000
Variety of shipping options	0,09	0,28	-1,00	1,30	1,15	3,50	-12,00	18,30	0,000
Charges stated clearly before order submission	0,05	0,22	-0,70	0,90	0,58	2,53	-8,00	11,10	0,000
Availability of product you wanted	0,06	0,40	-1,60	2,60	0,81	4,99	-18,40	44,80	0,005
Order tracking	0,05	0,44	-1,70	1,90	0,75	5,23	-21,80	28,40	0,003
On-time delivery	0,06	0,43	-2,30	2,00	0,84	5,17	-27,10	31,70	0,001
Product met expectations	0,01	0,39	-1,60	2,00	0,27	4,66	-17,20	27,00	0,861
Customer support	0,04	0,53	-1,80	2,60	0,69	6,69	-19,70	40,60	0,227

3. Methodology

If we want to conduct IPA for a particular store, we need to measure its attributes' performance and their importance. While the performance of every attribute (its mean rating) is directly available in the dataset, measuring importance is less straightforward.

We suggest estimating two regression models in first differences (but with the intercept that captures a possible change in the overall level of loyalty over time):

- **First absolute differences:** $\Delta Y_i = \alpha_0 + \alpha_1 \Delta X_{1i} + ... + \alpha_{13} \Delta X_{13i} + \varepsilon_i$ (1)
- **First percentage differences:** $\Delta(\%) Y_i = \beta_0 + \beta_1 \Delta(\%) X_{1i} + ... + \beta_{13} \Delta(\%) X_{13i} + u_i$ (2)

Estimating two models is useful for robustness check. Besides, depending on a situation and a practitioner's expert judgments, either model may be the preferred choice. If we had data from more than two time periods and wanted to use all the information, we could have used the so-called within estimator (for example, by including firm-specific individual effects using dummy variables), which is more efficient than the first-differences estimator when more than two periods are considered (Cameron & Trivedi 2005).

We will also compare first difference estimators with a **cross-sectional OLS regression based on the second (i.e. the latest available) period data** to see whether results based on panel data and one-period cross-sectional data differ significantly or not. This is an important issue, because sometimes an express study is needed, in which case there is no time to gather a panel dataset. The specification of the above-mentioned cross-sectional regression is as follows:

$Y_i^{t=2} = \gamma_0 + \gamma_1 X_{1i}^{t=2} + ... + \gamma_{13} X_{13i}^{t=2} + \nu_i$ (3)

Attribute importance was measured using the Shapley value decomposition of R^2, which recently came to econometrics from game theory. Therefore the Shapley value approach is theoretically grounded. In a regression, explanatory variables correspond to players and the coefficient of determination – to the total payoff of the coalition consisting of all the players (Lipovetsky & Conklin 2001). The Shapley value for a particular attribute equals to the average gain in R^2, which is observed when this variable is included in the model, across all possible sequences of independent variables inclusion into the model. For example, if there are n regressors, the number of combinations considered equals $\sum_{k=1}^{n} \frac{n!}{k!(n-k)!}$.

In addition, the sum of the Shapley values for all attributes equals R^2. That is why this approach is appropriate for the decomposition of R^2 into components. Percentage contributions obtained using the Shapley value approach can then be used in IPA, where they are confronted to attribute ratings.

4. The Shapley value decomposition

The estimation results for equations (1), (2) and (3) are presented in Table 2.

Table 2. Parameter estimates of regression models

(Dependent variable: *Would shop here again*)

	First point-differences		First percentage differences		Cross-sectional OLS regression based on the second period data	
	beta	robust std. error	beta	robust std. error	beta	robust std. error
Ease of finding what you are looking for	0.074	(0.127)	0.094	(0.137)	-0.057	(0.075)
Selection of products	0.055	(0.126)	0.071	(0.137)	0.074	(0.065)
Clarity of product information	0.022	(0.104)	0.010	(0.108)	0.071	(0.063)
Prices relative to other online merchants	0.085	(0.080)	0.117	(0.081)	0.027	(0.038)
Overall look and design of site	-0.072	(0.121)	-0.081	(0.130)	0.018	(0.076)
Shipping charges	-0.035	(0.033)	-0.030	(0.030)	-0.022	(0.016)

Variety of shipping options	-0.059	(0.071)	-0.089	(0.071)	0.065^{*}	(0.031)
Charges stated clearly before order submission	0.033	(0.087)	0.029	(0.089)	0.051	(0.066)
Availability of product you wanted	0.023	(0.068)	0.051	(0.067)	-0.104^{*}	(0.051)
Order tracking	0.117	(0.076)	0.117	(0.081)	-0.055	(0.057)
On-time delivery	0.265^{***}	(0.075)	0.254^{**}	(0.080)	0.340^{***}	(0.062)
Product met expectations	0.474^{***}	(0.059)	0.466^{***}	(0.058)	0.485^{***}	(0.050)
Customer support	0.164^{***}	(0.044)	0.163^{***}	(0.049)	0.199^{***}	(0.042)
Constant	-0.007	(0.014)	-0.097	(0.164)	-0.798	(0.469)
Observations	420		420		420	
R^2	0.746		0.746		0.840	
Adjusted R^2	0.738		0.738		0.834	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The Shapley value decomposition of each regression's R-squared is given in Table 3. In our study the number of attributes n is equal to 13, therefore, the number of combinations of regressors that should be considered to conduct each Shapley value decomposition equals $\sum_{k=1}^{13} \frac{13!}{k!(13-k)!} = 8191$. The procedure is relatively easy to program and even the computing power of a typical modern laptop is sufficient to obtain the results in a few minutes (it took us 88 seconds on Intel Core i7-3612QM CPU @ 2.10GHz with 8 Gb RAM).

Table 3. Percentage contributions of attributes based on the Shapley value decomposition (Dependent variable: *Would shop here again*)

Attribute	Importance, % or R^2 contributed by an attribute		
	First point differences	First percentage differences	Cross-sectional OLS regression based on the second period data
Ease of finding what you are looking for	0,46	0,49	1,48
Selection of products	0,88	0,92	2,8
Clarity of product information	0,92	1,01	3,1
Prices relative to other online merchants	1,12	1,31	2,12
Overall look and design of site	0,22	0,24	1,38
Shipping charges	0,27	0,23	0,82
Variety of shipping options	0,35	0,37	2,36
Charges stated clearly before order submission	0,41	0,47	2,26
Availability of product you wanted	13,86	14,35	9,4
Order tracking	14,24	14,31	12,09
On-time delivery	19,76	19,31	16,16
Product met expectations	29,58	28,91	26,95
Customer support	17,92	18,07	19,1

According to both first difference models, about 95% of repeat purchase intention growth is explained by post-order attributes, i.e. *On-time delivery*, *Order tracking*, *Customer support*, *Availability of product you wanted*, *Product met expectations*. The results based on point differences and percentages differences are essentially the same (Spearman correlation coefficient between the Shapley values based on model (1) and those based on model (2) is 0.989). Although other factors might be important for attracting new customers, they do not contribute to increasing retention rates. The results of the cross-sectional OLS regression based on the second period data are generally consistent with the results of the first difference estimation (Spearman correlation between the Shapley values from model (3) and any of the two other models is 0.901), but the importance of post-order attributes is estimated to be just 84%. We will use the first point difference estimates in our further analysis.

5. Importance-performance analysis

We have arbitrarily picked two Internet stores from our sample and conducted a brief importance-performance analysis of them. The names of the stores are not relevant for our illustrative example, which is why we do not disclose them.

Store A's repeat purchase intention mean score increased from 8.3 up to 9.3[4], despite the fact that the store has relatively low and decreasing ratings for *Variety of shipping options* and *Shipping charges*. This illustrates the idea that these attributes are among the unimportant ones: repeat purchase intentions are not very sensitive to changes in these characteristics. The importance-performance analysis for store A is summarized in Figure 1. Thanks to using panel data we are able to show (using different markers) how each attribute's rating changed between the two periods. The store's existing strategy can be maintained: the rating decreased only for unimportant attributes, none of the important variables have low performance. However, some usability ratings (*Selection of products, Clarity of product information* and *Charges stated clearly before order submission*) are already high enough for the store to revise their expenditures on improving the corresponding features, as well as to "educate" people about these advantages.

[4] Since only aggregate mean scores were available to us, we are not able to claim that the difference is statistically significant, but the fact that mean scores for all stores that are considered in Section 5 are based on responses obtained from over 500 shoppers makes all point estimates sufficiently reliable.

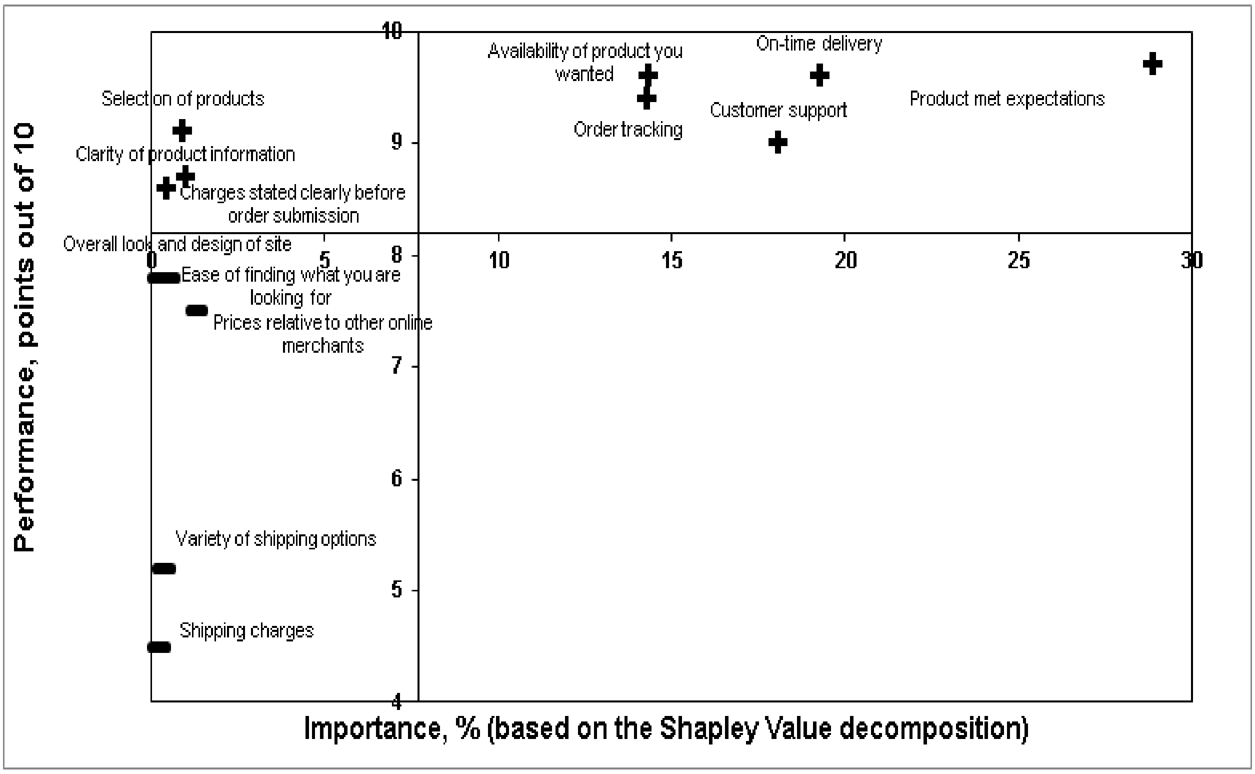

Figure 1. The importance-performance map for Store A (increased ratings are represented by "+", decreased ratings – by "-", unchanged ratings – by "●")

Store B's repeat purchase intention mean score decreased from 8.2 to 7.2, despite the fact that the store has relatively high and increasing ratings for *Ease of finding what you are looking for* and *Selection of products* and persistently high score for *Clarity of product information*. The importance-performance analysis for store B is summarized in Figure 2. None of the important attributes experience an improvement between the two periods. However, the satisfaction with most attributes representing Internet store usability and design and Economic considerations is high enough to advertise them as this store's features. So the store's strategy should be revised if they want to increase retention: specifically, improving important but underperforming attributes *Order tracking*, *Customer Support* and *On-time delivery* is of high priority.

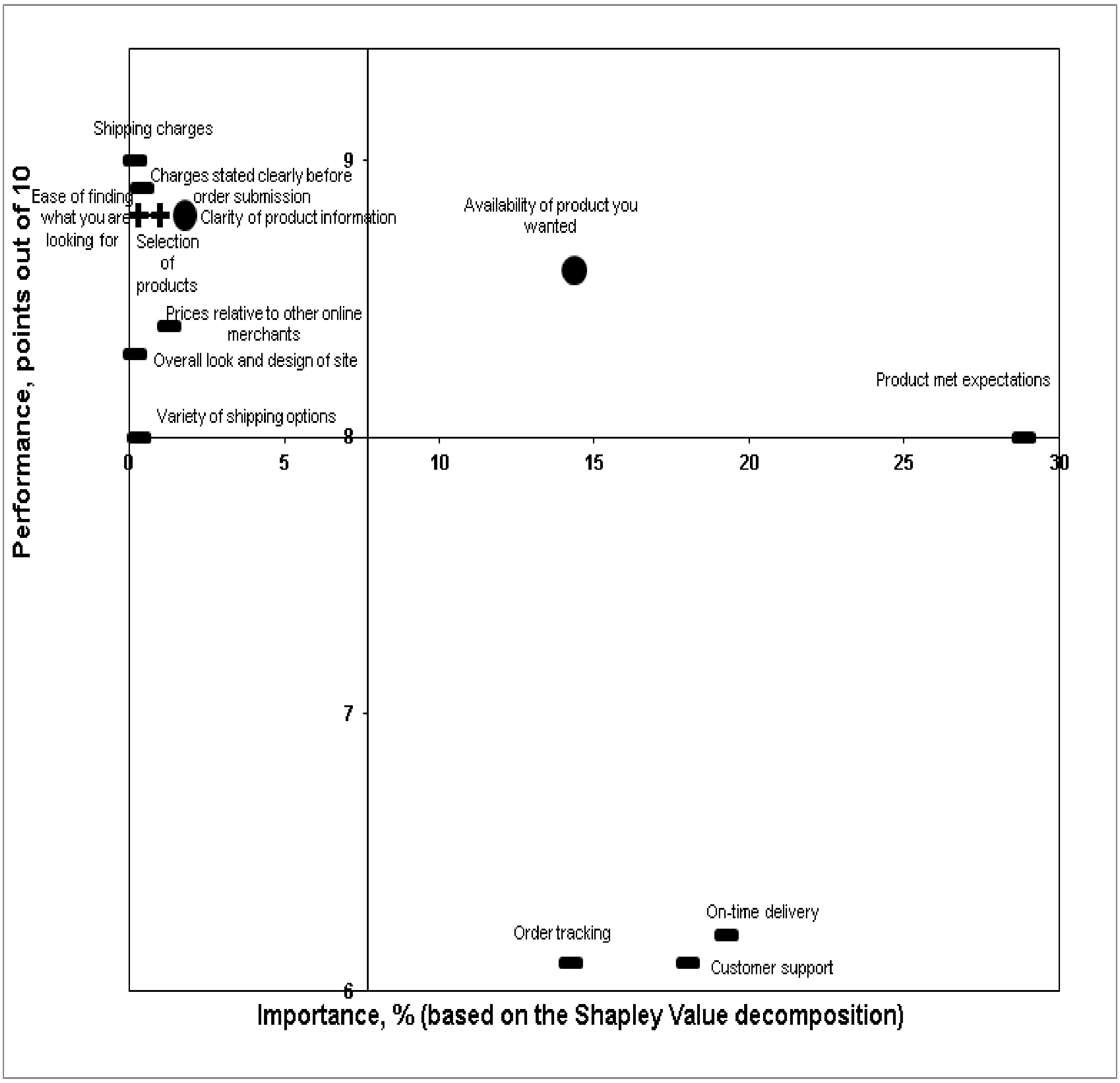

Figure 2. The Importance-performance map for Store B (increased ratings are represented by "+", decreased ratings – by "-", unchanged ratings – by "●")

6. Conclusion

Data from BizRate.com and other websites with store ratings can be easily collected using inexpensive web scraping software products or self-developed parsers. We have shown that such publicly available data provides an opportunity for managers to track the performance of their stores, as well as their competitors' stores, on annual, monthly or even weekly basis, i.e. to collect rich panel data. The use of longitudinal data allows wiping out time invariant (i.e. store-specific) omitted variables using the repeated observations over time.

With merchant ratings at hand, marketers are not limited to descriptive analysis only. With the help of Importance-performance analysis, backed up by a fixed-effects estimator and

Shapley value decomposition of overall performance, they can assess the “health status” of every service attribute and place them into one of the four categories: strengths, weaknesses, attributes which the company should bring customers’ attention to, and attributes, improvement of which is not of a high priority.

One of the main results is that post-order service and satisfaction with claims is much more important than usability, design and economic considerations (price, cost and speed of delivery). For instance, if the usability is great and continuously improved, it cannot compensate for the lack of satisfaction with claims or for the late delivery. This conclusion is especially obvious when indirect importance analysis is based on the results of the first-difference estimation, but cross-sectional analysis gave similar percentage contributions of attributes. In the first difference regression almost 95% of the explained variance is attributed to *Availability of product you wanted*, *Order tracking*, *On-time delivery*, *Product met expectations*, and *Customer support*. While we recommend accounting for unobserved firm-specific features using panel data, we should admit that in our case, the problem of unobserved heterogeneity is not very strong, which is why the cross-sectional regression based solely on the second period data only slightly underestimated the importance of post-order attributes.

It should be noted that it is reasonable to apply the proposed expert system when marketers are concerned with an insufficient level of loyalty, i.e. a low repeat purchase rate. Other type of research is necessary to work out, what should be done to attract more new clients. Another limitation of our study is that we assume that attribute importance is constant across different stores. Although we limit our sample to relatively large non-food stores, there still may be heterogeneous groups within the sample. To account for this heterogeneity, a researcher may use latent class models (e.g. finite mixture models) to reveal several types of stores, which have different key drivers of satisfaction. However, since the total number of US Internet stores, for which Bizrate ratings are available, exceeds 4000, an analyst can easily find a sufficiently large subset of stores for tracking and conducting importance-performance analysis. In addition, the information system described in this paper can be easily enhanced with micro-level data (i.e. with ratings posted by individuals rather than aggregated ratings). Availability of micro-data will make the sample size large enough for the analysis even at the level of a single store or a relatively small group of competing stores.

References

Cameron, A.C. & Trivedi, P.K., 2005. *Microeconometrics: Methods and Applications*, Cambridge University Press.

Deng, W, Chen, W. & Pei, W, 2008. Back-propagation neural network based importance–performance analysis for determining critical service attributes. *Expert Systems with Applications*, 34(2), pp.1115–1125.

Deng, W.-J. & Pei, Wen, 2009. Fuzzy neural based importance-performance analysis for determining critical service attributes. *Expert Systems with Applications*, 36(2), pp.3774–3784.

Deng, Weijaw, 2007. Using a revised importance–performance analysis approach: The case of Taiwanese hot springs tourism. *Tourism Management*, 28(5), pp.1274–1284.

Duke, C.R. & Persia, M.A., 1996. Performance-importance analysis of escorted tour evaluations. *Journal of Travel & Tourism Marketing*, 5(3), pp.207–223.

Fontenot, G. et al., 2007. Techniques for determining importance: Balancing scientific method and subjectivity. *Journal of Targeting, Measurement and Analysis for Marketing*, 15(3), pp.170–180.

Guadagnolo, F., 1985. The importance-performance analysis: An evaluation and marketing tool. *Journal of Park and Recreation Administration*, 3(2), pp.13–22.

Lipovetsky, S. & Conklin, M., 2001. Analysis of regression in game theory approach. *Applied Stochastic Models in Business and Industry*, 17(4), pp.319–330.

Martilla, J. & James, J., 1977. Importance-performance analysis. *Journal of Marketing*, 41(1), pp.77–79.

Martin, D.W., 1995. An importance/performance analysis of service providers' perception of quality service in the hotel industry. *Journal of Hospitality & Leisure Marketing*, 3(1), pp.5–17.

Matzler, K. et al., 2004. The asymmetric relationship between attribute-level performance and overall customer satisfaction: a reconsideration of the importance–performance analysis. *Industrial Marketing Management*, 33(4), pp.271–277.

Murdy, S. & Pike, S., 2012. Perceptions of visitor relationship marketing opportunities by destination marketers: An importance-performance analysis. *Tourism Management*, 33(5), pp.1281–1285.

O'Neill, M.A. & Palmer, A., 2004. Importance-performance analysis: a useful tool for directing continuous quality improvement in higher education. *Quality Assurance in Education*, 12(1), pp.39–52.

Pokryshevskaya, E.B. & Antipov, E.A., 2013. The Comparison of Methods Used to Measure the Importance of Service Attributes. *Working Paper available at SSRN: http://ssrn.com/abstract=2215680*.

Pokryshevskaya, E.B. & Antipov, E.A., 2012. The strategic analysis of online customers' repeat purchase intentions. *Journal of Targeting, Measurement and Analysis for Marketing*, 20(3), pp.203–211.

Van Ryzin, G.G. & Immerwahr, S., 2007. Importance-performance analysis of citizen satisfaction surveys. *Public Administration*, 85(1), pp.215–226.

Smith, S. & Costello, C., 2009. Culinary tourism: Satisfaction with a culinary event utilizing importance-performance grid analysis. *Journal of Vacation Marketing*, 15(2), pp.99–110.

Tontini, G. & Silveira, A., 2007. Identification of satisfaction attributes using competitive analysis of the improvement gap. *International Journal of Operations & Production Management*, 27(5), pp.482–500.

Elena B. Pokryshevskaya
National Research University Higher School of Economics (Saint-Petersburg, Russia). Department of Economics. Senior Lecturer;
E-mail: epokryshevskaya@hse.ru