Report on **Bachelor Thesis**

Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague

Student:	Theodor Petřík
Advisor:	Mgr. Kateřina Chadimová
Title of the thesis:	Impact of free customer credits on purchases: Case of a major discount portal

OVERALL ASSESSMENT (provided in English, Czech, or Slovak):

Please provide your assessment of each of the following four categories, summary and suggested questions for the discussion. The minimum length of the report is 300 words.

The goal of this thesis is to evaluate the efficiency of a loyalty program implemented at one of the Czech discount portals. The author has access to unique data of all customers and all purchases registered at the portal over the period of eight years, but uses just a small subset of these data in his analysis and collapses the individual data into daily sales at the firm level. The resulting aggregate dataset is analyzed using time series modelling tools in the scope exceeding the Bachelor-level curriculum at IES. The author concludes that the loyalty program does not affect sales of the analyzed discount portal.

Contribution

The main contribution of the thesis is given by the unique dataset that is used. The author analyzes effectiveness of a specific loyalty program that has not been analyzed before. Moreover, the author compares two time-series modeling methods: SARIMAX and a Facebook-provided procedure Prophlet. Comparison of these two methods in this specific setting is another contibution of the thesis. The application oh this thesis results are probably limitted to the specific company analyzed here.

Methods

As written above, the author uses two methods. SARIMAX is an extention of the well-known ARIMA that accounts for periodicity and allows for inclusion of explanatory variables.Prophlet is a Facebook-provided tool for time series data analysis. The choice of SARIMAX is a natural choice given the characteristics of the data used. The choice of Prophlet is proably given by the willingnss to compare traditional time-series modeling with a quasi-automated analytical tool as provided by Facebook. What I miss in the thesis is a more detailed comparision of these tools and a discussion of why their results differ. Is it possible to set Prophlet parameters such that its performance improves?

Some points need to be raised when it comes to the methodological part of the thesis:

First, the authors restriction of the dataset is not correct and most probably introduces a significant bias to the analysis. Instead of analysing the behavior of all customers of the discount portal (or all customers who made at least one purchase), the author restricts the dataset to purchases made by „active users", defined as users „who made at least ten purchases in the last four years of the collection period (from May 2014 - May 2018), and at least one purchase executed during the first and last three months of the observed period". This means that only returning users are included in the dataset and, consequently, the analysis in not able to detect the effect of the loyalty program that consists of flat distributiuon of credits over all regustered users (i.e. not only „active users"). To put it differently – a loyalty program works if a user who visits the discount portal once to make a specific purchase returns to it because he/she is motivated by the credits assigned and would not have returned if no credits were given. To identify this effect one has to observe a user who makes one purchase (a user who registers to the discount portal) and receives some credits and compare his/her behavior with a similar user who does not receive any credits. Observing the behavior of users who do not return to the portal is crutial, as these might be the users who did not receive any credits or received them at wrong times, etc.

Report on **Bachelor Thesis**

Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague

Student:	Theodor Petřík
Advisor:	Mgr. Kateřina Chadimová
Title of the thesis:	Impact of free customer credits on purchases: Case of a major discount portal

The author actually does not work with individual customer data, but aggregates weekly purchases by all „active users" to create an aggegate firm-level time series. While this approach disreggards a lot of information (e.g. when working with individual consumers one could apply panel data methods) is still can be valuable. Nevertheless, not counting other than „active user's" purchases into total sales introduces a huge bias to the analysis. When only loyal customers are included it is not possible to identify what made them loyal.

Further, there are some sample selection issues. For example, the author drops purchases made by customers who did not reveal their gender. Whie this has to be done when disagregating the results by gender, data generated by those consumers could be incuded in the aggregate analysis. Thise who reveal more information to the company might also beheve differently than those who reveal fewer information. For example, one can be motivated by a tergeted loyalty campaign to reveal his/her personal information and this response is not captured when the dataset is limited.

Finally, it is not clear from the text how forecasting is performed in case of sales and credit use. Are future (expected) values of the explanatory variables used or is forecasting done solely based on the information available up to date? Note that especialy credits used strongly depends on credits assigned and the behavior of the latter (which is totally under the control of the discount portal) was not stable over time. It might be better to include future (expected) values of credit assigned to predict sales and credits spent. One might use different scenarios – for example a scenaio where credits assigned follow the previous trend, where they level off at ast year's value, or where they go down – to formulate different predictions.

Literature

While I am not familiar with the literature on the topic of loyalty programs analysis, the literature review chapter seems complete to me. I only miss a clear positioning of the thesis within the cited literature.

Manuscript form

The manuscript is written in a clear and concise language. Its format is correct. The only minor coment concerns tables' descriptions and choice of variable names which are not always clear and one has to browse in the text to understand the information presented in form of tables.

Summary and suggested questions for the discussion during the defense

Overall, the thesis is of a good quality and is devinitely worth granting a bachelor's degree. I suggest that during the defence the author discusses the isues discussed above under the methods section.

Report on **Bachelor Thesis**

Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague

Student:	Theodor Petřík
Advisor:	Mgr. Kateřina Chadimová
Title of the thesis:	Impact of free customer credits on purchases: Case of a major discount portal

SUMMARY OF POINTS AWARDED (for details, see below):

CATEGORY		POINTS
Contribution	*(max. 30 points)*	25
Methods	*(max. 30 points)*	19
Literature	*(max. 20 points)*	19
Manuscript Form	*(max. 20 points)*	18
TOTAL POINTS	*(max. 100 points)*	**81**
GRADE	**(A – B – C – D – E – F)**	**B**

NAME OF THE REFEREE: **Barbara Pertold-Gebicka**

DATE OF EVALUATION: **27.5.2020**

_____
Referee Signature

EXPLANATION OF CATEGORIES AND SCALE:

CONTRIBUTION: *The author presents original ideas on the topic demonstrating critical thinking and ability to draw conclusions based on the knowledge of relevant theory and empirics. There is a distinct value added of the thesis.*

METHODS: *The tools used are relevant to the research question being investigated, and adequate to the author's level of studies. The thesis topic is comprehensively analyzed.*

LITERATURE REVIEW: *The thesis demonstrates author's full understanding and command of recent literature. The author quotes relevant literature in a proper way.*

MANUSCRIPT FORM: *The thesis is well structured. The student uses appropriate language and style, including academic format for graphs and tables. The text effectively refers to graphs and tables and disposes with a complete bibliography.*

Overall grading:

TOTAL	GRADE
91 – 100	**A**
81 - 90	**B**
71 - 80	**C**
61 – 70	**D**
51 – 60	**E**
0 – 50	**F**