Being Top X on Platform as a Driver of Petrol Station Cutting Prices

1 Introduction

This paper aims to build on Wein's 2021 paper *Why Abandon the Paradise? Stations' Incentives to Reduce Gasoline Prices at First*, and search for alternative explanations on gas stations' drivers to reduce prices.

The German gas stations market is characterized by Edgeworth cycles – strong intraday price cycles. Gas prices of a local market increase and decrease multiple times during the day, shoot up in late night, then fails heavily back in the early morning for another intense day competition. Since new gas prices, required by authority, must be published with 5 mins of change, customers can access prices almost in real-time and determine which stations to visit; meanwhile, suppliers are also threatened by competitors in real-time and are forced to cut price in response. The competition of lowering prices is terminated until one of suppliers cannot suffer the loss and increase price. The rest follows. The period of high price is broken by one of suppliers who wants to undercut again. Therefore, "the war of attrition, jump and undercutting" goes on (Wein, 2021).

Wein considers the price pressure from the closest competitors increases the probability of stations initiating a price reduction cycle, and station-specific characteristics as fixed effects factors influencing the said probability. The price pressure is tested to be highly significant, but has limited economic impact, given most price reductions are within 2 ct. Some station-specific factors are stronger, but in total, the pattern is not recognizable.

Prior literature mainly focuses on pressure exerted by competitors to reduce prices, mostly from the perspective of the supply side. For example, Eibelshäuser and Wilhelm considers the dynamic competition environment set by stations' opening hours (competitive environment is closer to perfect competition during the day and shifts to oligopoly during the night) (2017). On top of this, I'd like to propose to consider the threat posed by all competitors collectively on incumbent's pre-formulated competitive strategy given their knowledge on customer behavior. Instead of considering incumbent initiating an immediate response toward a competitor's price reduction, consider gradual changes on the incumbent's ranking in the entire local market.

Ranking is a crucial part of marketing because a good ranking represents devotion from consumers' limited attention span. Manufacturers pay supermarkets to have products displayed at customers' eye level; retailer pay for posters shown at bus stops. A good ranking has become even more important in the digital age. Websites pay search engines to be top of the displayed results; universities invest in events not necessarily linked to quality of education, hoping a better ranking will bring in more funding and more prominent students; shopping websites utilize algorithms to show the "recommended" list of products. Business, not-for-profit organizations, and governments pay to be "visible" towards targeted audience.

2 Hypothesis

A conjecture therefore can be formulated: petrol stations have direct or indirect knowledge on certain consumer behaviors, and therefore collectively apply a proximately constant competitive strategy on price adjustments to stay visible within the sight of consumers, such that resulting seemingly dynamic competing behaviors.

In this case, the gas prices and station information are gathered, sorted, and published on clever-tanken website and cell phone app. The website is the one of the 400th most websites in Germany and ranks third in the reference materials – public records and directories in Germany. Its audience include various age groups: 13.4% of age 18-24, 20.5% of age 25-34, 18.4% age 35-

44, 19.1% age 45-54, and 28.7% of age 55 or older (*clever-tanken.de (December 2022)*). Clever-tanken app has 31k reviews and ranks NO.5 in the navigation on app store (App store Germany, 2022). Thus, it is an influential platform and a good ranking on its website or app has business implications.

Based on the conjecture, a hypothesis from the perspective of ranking can be formulated: petrol stations adopt a "top X" strategy on online platforms that display and sort petrol prices, and it a station's ranking is outside of top X, it will reduce gas prices. This strategy is platform-specific and applies to intraday price fluctuations. Wein defines 10,918 price reductions events and finds puzzling that three quarter of them are reductions of only 1-2 ct (2021). This hypothesis could explain such behavior: the stations do not necessarily need to be the provider of the lowest petrol prices in the region, but only needs to the one of the providers that offer lower prices, because customers don't only consider the cheapest price but several cheapest prices due to station-specific factors in play. Such "top X" strategy can be easily implemented by stations owners and employees in operation. Further, customer inertia, a mindset that customers adopt by being consistent with prior purchases to minimize thinking and regret, hinders a station's ability to attract new customers and profit from being the lowest price provider (Henderson et al, 2021). Therefore, decreasing price drastically in the active day market is not rewarding; rather, showing multiple small price reductions over the day give stations opportunities to frequently be visible. The hypothesis formula is written as:

H1: price reduction event (0/1)

= $\beta 0$ + $\beta 1$ * ranking out of top X (0/1) + βi * controlling factors

A second hypothesis was presented on Dec 13, 2022, regarding time-specific factor that influences interday frequency of price cycles: gas stations adapt the frequency of price reduction cycles based on the consumer commute behaviors: during weekdays there are fewer cycles but concentrated on peak hours such as 8-10am; on weekends there are more cycles and more diffused throughout the day. A further hypothesis is that COVID lockdown forces people to commute less and travel in less peak hours, therefore there should be a lower percentage of price reduction events during the normal peak hours, and this relationship is causal. Searching for data representation of lockdown yields COVID-19 government response tracker. Researcher at Oxford collected systematic information on governments' policy response to tackle COVID-19, coded into 23 indicators such as school closures and travel restrictions, and aggregated into one index ranges 0-100. Yet the variation of index is rather non-significant during the chosen timeframe (min 64.35 max 76.85, 76.85 sustains for 18 days during Apr 15 – May 2, 2020). A year comparison could not be performed because the index is constant 75 in during Apr 15 – May 14, 2022. Therefore, the test of hypothesis two will not be performed.

3 Data selection

The city of Rosenheim is chosen as the targeted local market (around 61,000 in habitants) (*Rosenheim, Bavaria, Germany*). Geographically, the city takes a long pointy shape. To confidently include all relevant gas stations, I expanded the coverage to include the entire area centered at Rosenheim train station (Südtiroler Pl. 1, 83022) with a radius of 7km. The city of Rosenheim has an area of 37.22 square km (*Rosenheim, Bavaria, Germany*), equivalent to a circle with a radius of 6.1 km. Therefore, it is considered inclusive to round up the radius to 7km.

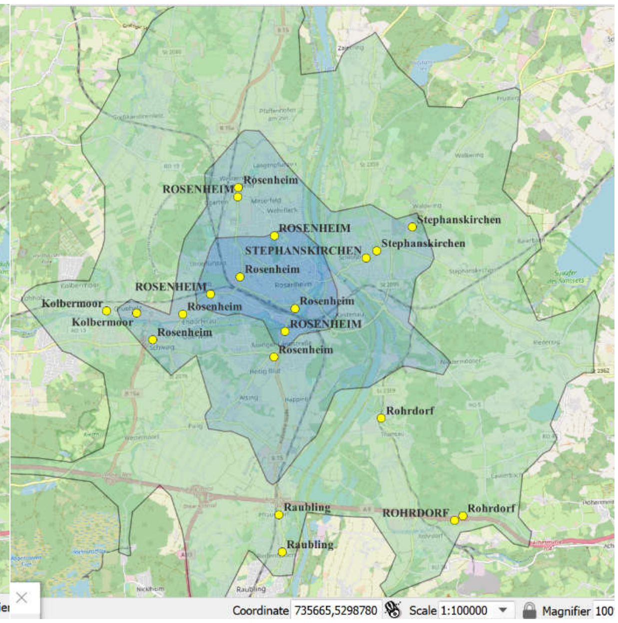

Figure 1. 20 selected stations within 7km radius centered at Rosenheim station. Stations are labeled by city.

Figure 2. 5, 10, 15 mins driving coverage starting from Rosenheim station. Stations are labeled by city.

.

This gives exactly 20 stations in April 2020 (including 12 stations in Rosenheim, 3 stations in Stephanskirchen, 3 in Rohrdorf, 2 in Raubling) (see Figure 1). The 20 stations are sounded by non-residential area (yellow part in Figure 1), therefore collectively form a local market. Figure 2 shows all selected stations are within 15 mins of driving time from the Rosenheim station. 5 stations are within 5 mins of driving time, 9 stations are between 5-10 mins of driving time, 6 stations are between 10-15 mins of driving time. Overall, majority of stations are located within 20 mins of driving time from each other (see Figure 3). The selected period for this report is April 15 – May 14, 2020.

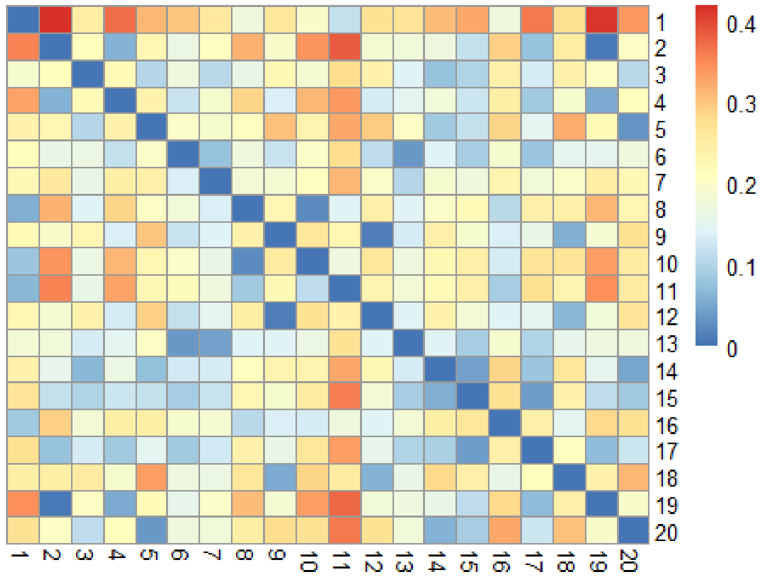

Figure 3. Time of driving (in hr) between selected petrol stations. Majority of stations are located within 0.3 hrs (20 mins) of driving time from each other. Data computed by OpenStreetMap QGIS.

Due to the unclarity of the default sorting algorithm on the clever-tanken app, I was forced to define a sorting method for the project and therefore could not replicate the order of display on the app. The app has four sorting methods: by standard method, by distances from closest to furthest, by operator brand name alphabetically from A to Z, and by gas prices from lowest to highest. The default sorting method is the standard method, yet the algorithm behind it is not perfectly clear – it seems to be a mixture of price, duration the current price (the earliness of the latest price change, the earlier the higher on ranking), name of operator brand, and distance from the starting point.

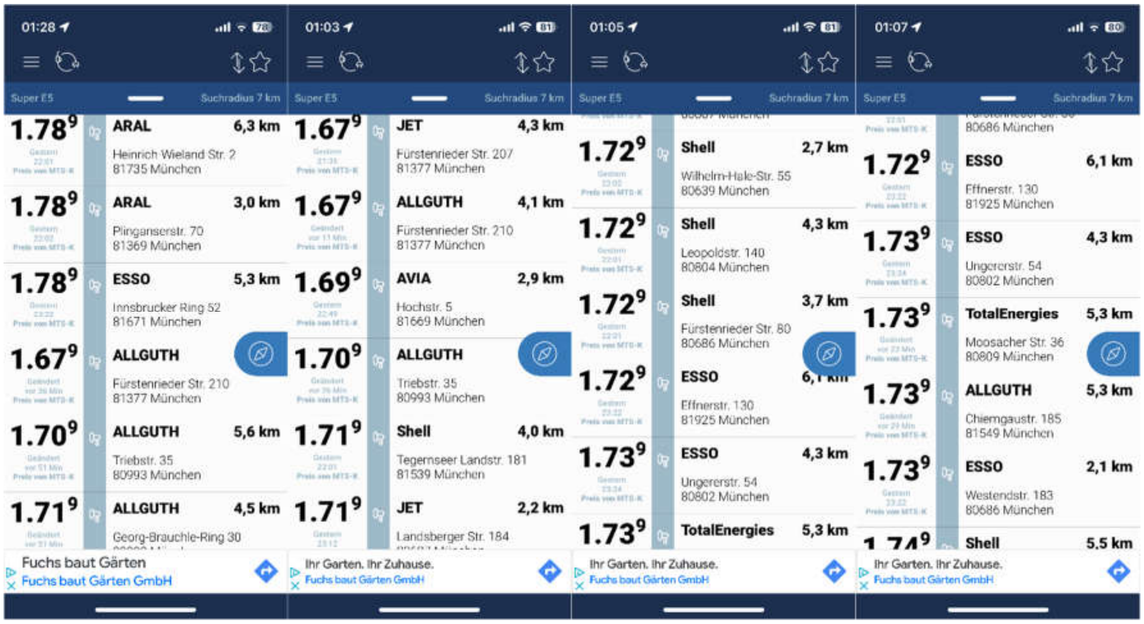

Figure 4. Partial snapshots of gas stations (using Munich station as the starting point) sorted by the standard method of clever-tanken app. The sorting algorithm behind standard method seems to predominantly relies on price, duration of the current price. Operator brand name and distance from starting point plays a minor factor.

To illustrate the unclarity, snapshots were taken using Munich city center as starting point due to populated gas stations in the surrounding area. 46 opening gas stations were listed, and the snapshots display partials of the list. From parts 1-3 in Figure 4, it is apparent that gas price is the most influential sorting factor. Multiple examples on stations ordering at the same price level suggest that the duration of the current price is the second strongest factor. The example of 1.739 ALLGUTH (5.3km) lies before 1.739 ESSO (5.3km) indicates alphabetical order of operator brand could be a factor. Similarly, distance from the starting point also takes a minor role: 1.739 ESSO 4.3km lies before 1.729 ESSO 6.1km. Besides, it's suspected that there are hidden factors in play. For example, 1.739 TotalEnergies (5.3km) is not advantageous when it comes to duration of current price and alphabet of operator brand, yet it lies before ALLGUTH (also 5.3km) and ESSO (2.1km). Similarly, it's surprising that three ALLGUTH stations with lowest prices were displayed at the end of list (part 4).

Top X is determined by descriptive data collected on ranking changes, instead of searching for literature. The last price of every station before April 15 are extracted, ordered and saved as the initial prices table. For each price adjustment starting April 15 (regardless of increase or decrease), the initial prices table first feeds out the previous ranking of the adjusting station, then it is updated with the latest price and time, such that it contains the most current price information. The current prices table is then sorted in ascending order by petrol price, date, and brand name of stations, and returns an updated ranking of the adjusting station. Further, since the ranking is complete only when all stations are open, only price reductions between 7am-9pm are selected. Comparisons of station rankings pre and post price reductions are made. Before price reductions, stations overall are mostly ranked on and below ten (probability is in 5-10% for each rank lower than ten) and are rarely ranked within the top five, while after the changes, stations are ranked rather evenly across the entire ranking list (most ranks have probabilities in 3.5-6.5%).

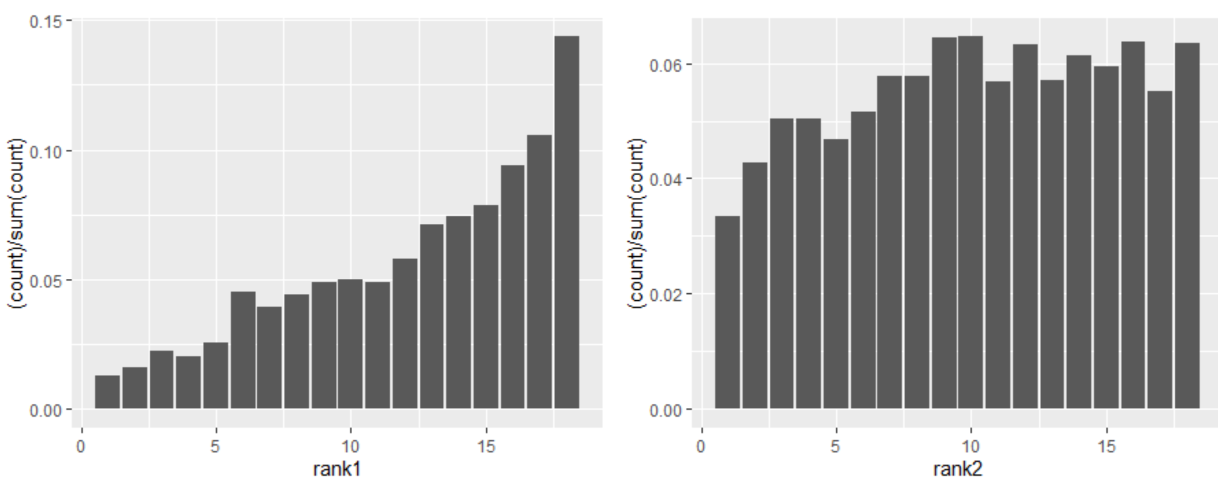

Figure 5. Histogram of station rankings before (left) and after (right) price reductions. Total 4,770 observations.

Figure 6 further illustrates that the chase of ranking differs by brand. Distributions of stations rankings for brands like Agip, ARAL, ESSO, and OMV remain a similar shape before and after price changes, while distributions of brands like ALLGUTH, HEM, JET, Kaufland, Shell become much more concentrated. ALLGUTH, HEM, and Kaufland particularly drive to the in the top three, and JET is rather comfortable being exactly the top five.

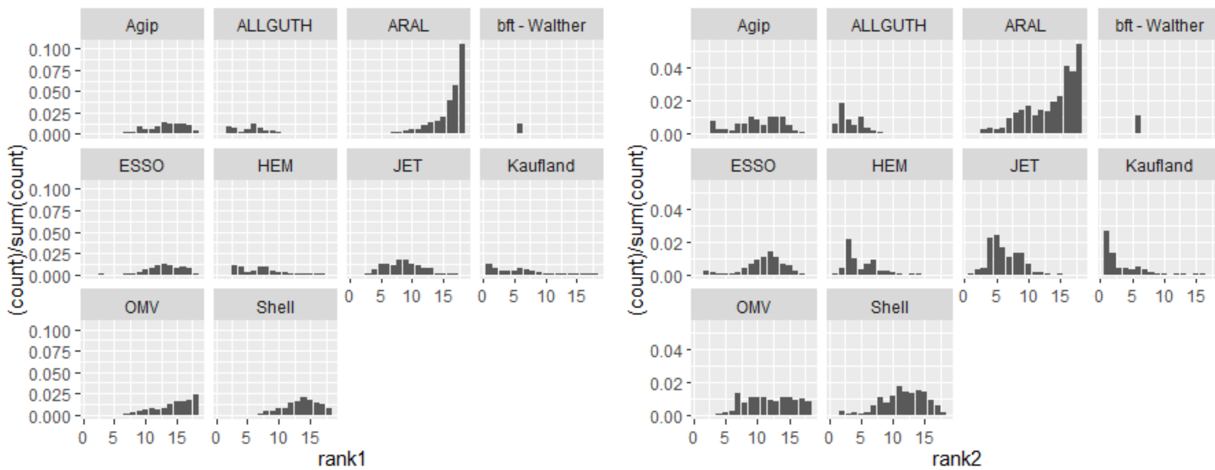

Figure 6. Histogram of station rankings before (left) and after (right) price reductions, per operator brand.

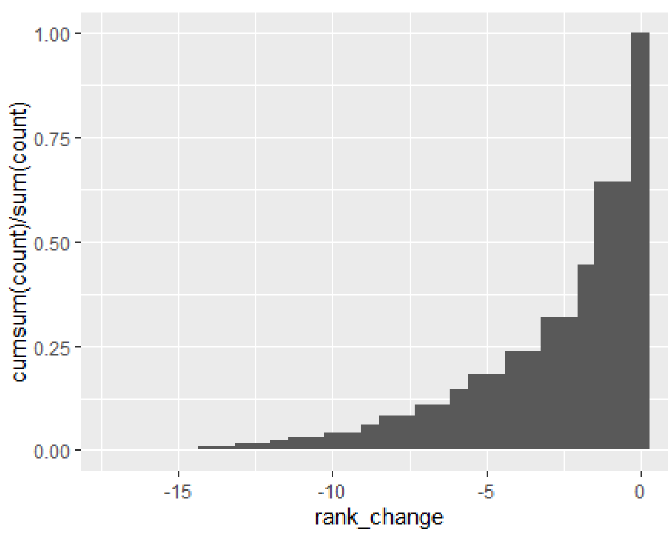

Figure 7. Cumulative histogram of ranking change per price reduction. Negative values correspond to numeric value of ranking being reduced. More than 75% of changes lift the station's ranking up by five levels or less.

Further, Figure 6 shows more than 75% of price reductions lift the station's ranking up by only five levels or less, and this call for conservatism when artificially setting up the X for top X strategy. Given all rankings observations, the X of top X strategy is set as ten.

The selection of station-specific factors is largely based on Wein's paper, with several adjustments mainly due to unavailability of accurate data and time restriction.

Three variables are not included in the category of other products and services offered by gas stations (data extracted from clever-tanken website): consumer satisfaction, number of accepted credit cards, and variables related to opening dates and hours (e.g., whether opens on weekends and public holiday). The exclusion of customer satisfaction is due to limit of free usage of pulling customer reviews data from Google Maps API. In the manual data collection process, I noticed several stations have changed operator brand since Apr-May 2020, which could result in inaccuracy of non-infrastructure related data, such as acceptance of credit cards and opening time. Therefore, a conscious decision was made to exclude these two factors. Acceptance of certain brand's petrol card is mentioned in the Wein paper (2021) but is deliberately not included for the same reason. Infrastructure such as ATMs is unlikely to be removed from the site, so related data are considered as close to true representation of products and services offered in Apr-May 2020. There is no Bavaria public holiday during selected time frame.

Among location category variables, commuter route to neighboring cities is substituted by driving time from Rosenheim station. The variable was included in the Wein paper based on the consideration that traffic volume may affect stations' choice on likelihood, time and magnitude of price changes. Since the selected stations have included 8 stations located in neighboring cities of Rosenheim, I instead used the ordinal scale on driving time from Rosenheim station to represent stations' willingness to compete in the Rosenheim local market. The longer the driving time, the less likely a station is effect by the overall price dynamics of Rosenheim stations, and therefore less likely to initiate a price reduction action. Following figure 2, the variable driving time is transformed ordinally: less than 5 mins of driving time corresponds to value 1, driving time between 5-10 mins corresponds to value 2, and driving time between 10-15 mins corresponds to value 3. The variables "near national road?" and "near motorways?" are adapted to proximity to railway and highway respectively (see Figure 8).

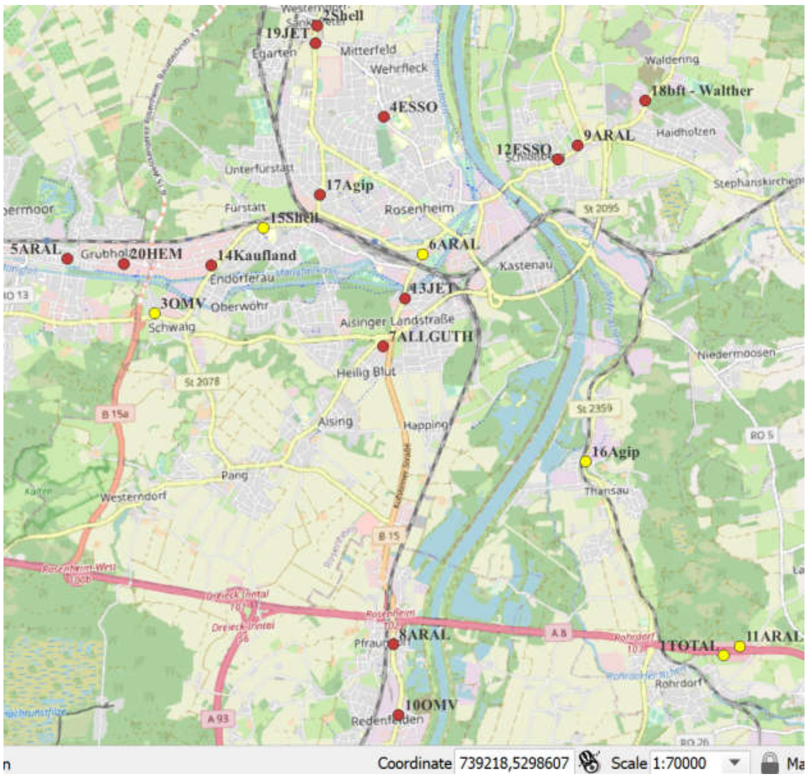

Figure 8. Stations 6, 15, 16 are close to railway; stations 1, 3, 11 are close to highway. Rest 14 stations are only close to secondary highways. Stations are labeled by unique index and operator brand.

4 Method and result

The method is to first apply Lasso fixed effect logistic regression to filter out less relevant variables, then apply fixed effect logistic regression with the remaining variables. Despite many variables in the Wein paper are tested significant, empirically they show little economic implication and non-recognizable patterns, so I assumed sparsity in this model.

9,031 price adjustment observations and 26 variables (later transformed into 38 variables due to changing categorical variables into dummy variables) are used to compute Lasso model (see Table 11 for variable representation). The optimal lambda value that minimizes prediction error is $1.165643*10^{-4}$. The lambda that is within one standard error of optimal lambda and offers the simplest model is $7.66915*10^{-3}$. The later is chosen due to optimal lambda has kept majority of variables. 13 variables are chosen by Lasso with lambda of $7.66915*10^{-3}$ (see Table 10).

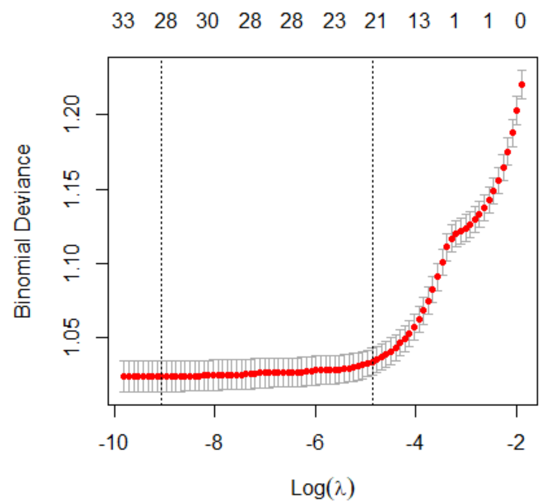

Figure 9. Lasso model deviance with respect to different log(lambda) values.

1	out_top10	6	brandJET	11	high.pressure.cleaner
2	p_weekdayTue	7	brandKaufland	12	work.shop
3	p_weekdayWed	8	index	13	bakery
4	brandALLGUTH	9	non.oli..local		
5	brandbft - Walther	10	vacuum.cleaner		

Table 10. List of variables selected by Lasso.

The chosen variables are then used to perform logistic regression. Variable p_wekdayTue and p_weekdayWed are transformed from categorical variable p_weekday; variable brandALLGUTH, brandbft – Walther, brandJET, and brandKaufland are transformed from categorical variable brand. The original categorical variables are used in the logistic regression, instead of transformed ones. Thus, the logistic regression takes in nine independent variables.

	Variable	Type
1	cut (dependent variable)	dummy
Independent variable		
2	out_top10	dummy
3	p_weekday	categorical
4	is_weekend	dummy
5	brand	categorical
6	index	categorical
7	railway	dummy
8	highway	dummy
9	driving.time	categorical
10	oligopolist	
11	non.oli..national.or.regional	
12	non.oli..local	
13	shop	
14	bistro	
15	WC	
16	vacuum.cleaner	
17	portal.car.wash	
18	high.pressure.cleaner	dummy
19	work.shop	
20	backery	
21	truck.gas.station	
22	ATM	
23	restaurant	
24	diaper.station	
25	shower	
26	self.service	

Table 11. List of original variables used in Lasso.

Variable out_top10 is tested highly significant: when a station's ranking drops from within the top ten to lower than 10, the log odds of this station to cut price increases by 2.66 pp (see Table 12). Among all weekday factors, Tuesday and Wednesday are tested significant and highly significant: when the day of week is Tuesday or Wednesday, the log odds of stations decrease prices increase by 0.26 and 0.32 pp respectively. Among all brand factors, ALLGUTH, bft-walther, JET, and Kaufland are tested significant, showing high probability of engaging in price reduction behaviors. Kaufland has the strongest positive effect: a petrol station switching from non-Kaufland brand to Kaufland will increase the logit odds by 1.78 pp. This is consistent with Kaufland's fixation on being in top three mentioned in the Data selection section. Three variables are thrown out by glm function due to singularity. Variable vacuum_cleaner is tested highly significant: a station with vacuum cleaner increases its log odds to cut prices by 1.04 pp. Dummy variables like as shops and car_wash are tested highly significant in the Wein paper (2021), yet are considered less relevant and thrown out by Lasso, possibly due to fewer outdoor activities during the COVID-19 lockdown period. Varying magnitude of coefficients confirms the original sparsity assumption on Lasso.

Given null deviance 11,026.1 and residual deviance 9,240.1, McFadden's pseudo R squared is calculated as 0.162, which suggests that the logistic model has certain explainable power, although not high.

	term	estimate	std.error	statistic	p.value
1	(Intercept)	-1.29937	0.194665	-6.67489	2.47E-11
2	**out_top10**	**2.658046**	0.073431	36.19767	6.62E-287
3	p_weekdayMon	0.218576	0.099891	2.188158	0.028658
4	p_weekdaySat	0.053617	0.101044	0.530627	0.595677
5	p_weekdaySun	0.182028	0.101143	1.7997	0.071908

6	p_weekdayThu	-0.03001	0.095585	-0.31394	0.75357
7	**p_weekdayTue**	**0.263272**	0.099278	2.651872	0.008005
8	**p_weekdayWed**	**0.319581**	0.094398	3.385474	0.000711
9	**brandALLGUTH**	**0.75973**	0.197712	3.842618	0.000122
10	brandARAL	-0.32777	0.131183	-2.49858	0.012469
11	brandbft - Walther	0.924686	0.27997	3.302807	0.000957
12	**brandESSO**	**0.483072**	0.173795	2.77956	0.005443
13	brandHEM	0.251315	0.243761	1.030989	0.302546
14	**brandJET**	**1.298019**	0.122941	10.55808	4.66E-26
15	**brandKaufland**	**1.776565**	0.13677	12.98947	1.40E-38
16	brandOMV	-0.32671	0.14738	-2.21678	0.026638
17	brandShell	-0.07661	0.131123	-0.58426	0.559045
18	brandTOTAL	-2.88055	1.176194	-2.44905	0.014324
19	index	0.024656	0.009617	2.563835	0.010352
20	non.oli..local	NA	NA	NA	NA
21	**vacuum.cleaner**	**1.043755**	0.171943	6.070339	1.28E-09
22	high.pressure.cleaner	NA	NA	NA	NA
23	work.shop	NA	NA	NA	NA
24	bakery	0.186529	0.114937	1.622873	0.104617

Table 12. Coefficients of logistic regression. Significant variables are in bold.

5 Limitation

Limitation on variable selection and representation: As previously mentioned, the logic of standard sorting method is not published by clever-tanken app, and an exact sorting order could not be replicated, such that the ranks of stations are inaccurate. The Rosenheim station (city center of Rosenheim) is used as the starting point to recreate the rankings, however, customers are more likely to use their own current locations as the starting point to extract prices. This further influences the accuracy of station rankings. There exists omitted variable bias since multiple variables are dropped from the Wein's controlling variable list, due to accuracy concerns.

Limitation on model bias: when computing for the Lasso model, the lambda that gives least prediction error is deliberately not chosen for the purpose of getting a simple model. This may further increase the omitted variable bias in the final logistic regression model.

Reviewing the histogram of rankings before and after price reductions, it's obvious that different brands adopt different Xs for top X strategy. Further research can be done to examine whether cross-region petrol stations with the same operator brand adopt similar Xs.

6 Conclusion

Considering collective threat posed by all local competitors and business implications of rankings on online platforms, a hypothesis is formulated that petrol stations will decrease prices when its ranking is outside of top X. Based on 4,770 price cuts, X is set as ten. Post Lasso logistic regression is performed; eight independent variables are tested as significant and positively influence the log odds of a station engage in price cuts. The model suggests that when a station's ranking is lower than ten, the log odds of it reduce prices increase by 2.66 pp. The model has a pseudo r squared value of 0.162, suggesting certain power of explanation.

Works cited

(2022) Covid-19 government response tracker. In: Research project COVID-19 Government Response Tracker. https://www.bsg.ox.ac.uk/research/covid-19-government-response-tracker. Accessed 1 Jan 2023. Excel file downloaded at https://github.com/OxCGRT/covid-policy-tracker/blob/master/data/timeseries/OxCGRT_timeseries_all.xlsx.

App store Germany, 2022. Clever-tanken.de app. Accessed December 13, 2022.

Clever-tanken.de (December 2022). (n.d.). [Web traffic analytics]. Similarweb LTD. Retrieved January 1, 2023, from https://www.similarweb.com/website/clever-tanken.de/#overview

Eibelshhuser, S., & Wilhelm, S. (2017). Markets Take Breaks: Dynamic Price Competition with Opening Hours. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2879392

Haucap, J., Heimeshoff, U., & Siekmann, M. (2017). Fuel Prices and Station Heterogeneity on Retail Gasoline Markets. The Energy Journal, 38(01). https://doi.org/10.5547/01956574.38.6.jhau

Henderson, C. M., Steinhoff, L., Harmeling, C. M., & Palmatier, R. W. (2021). Customer inertia marketing. Journal of the Academy of Marketing Science, 49(2), 350–373. https://doi.org/10.1007/s11747-020-00744-0

Rosenheim, Bavaria, Germany. (n.d.). DB-City. Retrieved January 1, 2023, from https://en.db-city.com/Germany--Bavaria--Rosenheim

Wein, T. (2021). Why Abandon the Paradise? Stations' Incentives to Reduce Gasoline Prices at First. Journal of Industry, Competition and Trade, 21(4), 465–504. https://doi.org/10.1007/s10842-021-00362-8