Three essays on individual behavior and new technologies

Husiatyński, Maciej

DOI:
10.26116/center-lis-2121

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

CentER

Three essays on individual behavior and new technologies

MACIEJ HUSIATYŃSKI

Tilburg University

Three essays on individual behavior and new technologies

Proefschrift ter verkrijging van de graad van doctor aan Tilburg University

op gezag van de rector magnificus, prof. dr. W.B.H.J. van de Donk, in het

openbaar te verdedigen ten overstaan van een door het college voor

promoties aangewezen commissie in de Aula van de Universiteit op vrijdag 22

oktober 2021 om 10.00 uur

door

Maciej Husiatyński,

geboren te Warschau, Polen

Promotores:	prof. dr. T.J. Klein (Tilburg University) prof. dr. M.C. Mikkers RA (Tilburg University)
Leden promotiecommissie:	prof. dr. C. Propper (Imperial College London) prof. dr. G. Tezel (Tilburg University) prof. dr. E. Bartelsman (Vrije Universiteit Amsterdam) prof. dr. J.H. Abbring (Tilburg University) prof. dr. J. Boone (Tilburg University)

Een deel van dit onderzoek is uitgevoerd tijdens een stage en een deeltijdbaan bij de Nederlandse Zorgautoriteit.

Acknowledgments

It has been nine years since I have left Warsaw to start my studies abroad. Since then I have attended a bachelor degree at the University of Glasgow, a Research Master degree at Tilburg University and, most recently, a PhD program in Economics at Tilburg University. It is needless to say that getting to this stage of my life would not be possible without the ongoing support of my family and friends. While I am unable to mention everyone who played a role in my journey here, I would like to take this opportunity and thank those who have been there during the last few years.

I would firstly like to thank my supervisors, Tobias Klein and Misja Mikkers, who guided me throughout the doctorate program. Tobias always pushed me to explore new directions in my research and was ready to answer all my questions even if I had just dropped by his office. Misja continuously supported me throughout my work and always shared his expertise, especially during my time at the Dutch Healthcare Authority. They were providing me with valuable advice, but at the same time they were also giving me the space to do things my own way. Perhaps most importantly, they gave me a necessary push forward in the middle of my PhD track, for which I am extremely grateful. Thank you, I am here now partially because of your ongoing support.

I am further grateful to the committee that has provided me with many helpful comments. They have highlighted many issues in the essays and pointed to directions that could improve the thesis substantially. I am thankful for all the feedback I have received for my papers, especially during the Structural Econometrics Group meetings. All these comments allowed me to improve the thesis and make it into what it is now.

I have spent the last few years in a very supportive environment of friends that I have met throughout my studies. They were helpful in research, providing valuable comments and long, engaging discussions, but perhaps even more importantly, a great company to spend time together outside of work. Here I would like to especially thank Rafael and Tung who started with me from the first day of the Research Master program and with whom I have spent most of my time, as well as my other classmates: Ming, Paan, Benz and Nora. I am grateful to my office roommates Daniel, Oliver and Lkhagvaa with whom I have met for countless lunch breaks, my band members Joris, Ricardo and Marie, but also Ana, Martin, Clemens, Merve, Hazal, close friends from Glasgow: Yannick, Janne and Chris, and everybody else who were there along the way. It was always good to know that whenever I visit my hometown, I can count on Robert, Kamila, Mateusz, Stanisław, Mateusz, Patryk and Marcin. You all made this journey a lot easier, and a lot more entertaining too. Thank you.

It goes without saying that I would not be here without the support of my mother who was continuously encouraging me to work hard throughout my studies. Together with my siblings, they were always there to consult me on the toughest decisions I had to make. Piotr helped me improve my programming skills and it was great to know that I can spend time with Tomasz, Magda and Maciej whenever I am in Warsaw. I am really lucky to be a part of such a great family.

Towards the end of my PhD, I was also extremely lucky to meet my girlfriend Ewa. She listens to my monologues about economics and took me and our dog on countless beautiful hikes on the weekends. I enjoyed every bit of it, even the one that involved running down the mountain. Thanks to her, the last year of my PhD program was actually the best one so far. There are no words that can express how grateful I am for that.

Contents

1 Introduction **4**

2 Increasing price transparency in the Dutch health care market does not affect provider choice **6**
2.1 Introduction 6
2.2 Institutional setting 8
2.2.1 The Dutch health care system 8
2.2.2 Information 9
2.2.3 Selection of procedures 12
2.3 Data 15
2.4 Empirical approach 19
2.4.1 Effects on spending 19
2.4.2 Effects on choices 22
2.5 Results 23
2.6 Discussion 26
2.7 Appendix 30

3 Indirect reciprocity stimulates viewership gifts in the video game streaming industry **35**
3.1 Introduction 35
3.2 Data 39
3.2.1 Data collection and hosting 39
3.2.2 Sample selection and summary 42
3.3 Empirical results 48
3.3.1 Models 48
3.3.2 Results 52
3.3.3 Heterogeneity 54
3.3.4 Robustness 56
3.4 Discussion 58
3.5 Appendix 61

4 Friendships accelerate exit from video game streaming platform once shutdown is certain **69**
4.1 Introduction 69
4.2 Data 73
4.2.1 Platform shutdown and sample selection 73
4.2.2 Costreaming in relation to the activity post-announcement 76
4.3 Model 80
4.4 Results 82
4.4.1 Baseline results 82
4.4.2 Robustness results 84
4.5 Discussion 88
4.6 Appendix 91

5 Bibliography **96**

1 Introduction

Appearance of new technologies allows us to analyze consumer behavior under original and often understudied conditions. Furthermore, availability of new and detailed data on an emerging market can greatly facilitate investigating known economic phenomena. This thesis consists of three essays in which I examine innovative tools and markets that became available thanks to the recent technological advancements. In particular, I study how they affect the individual behavior of the market participants.

In the first essay, we looked at the introduction of web-based tools that allow consumers to quickly compare prices of health care products offered by different hospitals and clinics in the Netherlands. Price transparency is often viewed as an effective way to encourage price shopping, leading to a lower health care expenditure for the entire system. Not surprisingly, publication of the health care prices in the Netherlands was recently supported by the minister of health.

While we observed that the visits to the website containing price comparison tool surged after the publication of prices, spending, the likelihood to visit a new provider, distance traveled, and type of provider visited remained unaffected. This is even more surprising given that we observed large savings opportunities that were available to individuals in the sample, ranging from 23% to 25% of the price paid among those who could have saved.

We argue that the potential reasons behind these results are a complex system of health care procedures and an unpredictable price dispersion coupled with a very low annual deductible of just 385 euros that is chosen by a vast majority of the consumers. The results seem to confirm the general notion in the literature that price transparency can work well but it still needs to be properly designed. Our results suggest that the current price transparency initiative in the Netherlands did not affect consumer behavior in the short run.

In the second essay, I investigated hosting: mirroring other video game streamer's content coupled with viewership transfers, to study indirect reciprocity. Indirect reciprocity can be defined as being kind to somebody after receiving an act of generosity from a third party. In contrast to direct reciprocity that involves returning the favor directly to the sender, indirect reciprocity results in returning a favor to an individual or individuals not involved in the initial encounter.

Using a phenomenon of hosting that is unique to the new and emerging video game streaming industry, I show that streamers are more likely to host others if they were hosted in the recent past. Furthermore, being recently hosted motivates them to choose less popular channels as the recipients, and for whom such gifts are arguably of much higher value. This particular type of upstream indirect reciprocity affects less popular and less experienced channels most.

The results of this study provide robust evidence for upstream indirect reciprocity in the growing market of video game streaming. Even when monetary gains are at stake, people tend to be reciprocal. Furthermore, the norm of hosting that is quite popular on the platform can be, at least to some degree, a result of indirect reciprocity.

Finally, in the third essay I study consumer churn and platform collapse. A major video game streaming platform has announced a shutdown to occur within a period of one month. However, the exit from the platform was gradual and started long before the actual announcement. Using a unique feature called costreaming that allows up to 4 streamers to join their video broadcast, I investigate how the size of the costreaming networks affects the decision to stay on the platform.

Users with more costreaming links invested less hours and logged in fewer times after the announcement. However, I show that even though networked streamers were less likely to leave prior to the announcement, they were also strongly influenced by their peers who did.

Introducing means of cooperation such as costreaming, an optional feature that allows co-producing the broadcast for the viewers, can strengthen the attachment of users to the platform. However, once knowledge about the shutdown becomes common, costreaming networks may speed up the aggregate exit from the platform.

2 Increasing price transparency in the Dutch health care market does not affect provider choice

This chapter is based on an identically named paper coauthored with Tobias Klein and Misja Mikkers.

2.1 Introduction

Providing consumers with information about prices is often considered to be an effective way to offset increasing health care expenditure, as such price transparency initiative combined with cost sharing should drive consumers towards a more cost-conscious choice. Thereby, it also promotes competition between providers (Mehrotra et al., 2017; Hibbard et al., 2012; Volpp, 2016).

This is one of the reasons why many US states have adopted some degree of price transparency legislation (Volpp, 2016) and calls for increased transparency in the Netherlands were supported by the ministry of health (Kleijne, 2016).

Although studies report that only a small fraction of individuals engage in comparing prices (Desai et al., 2016; Chernew et al., 2018), several papers highlight the potential for supply effects (Brown, 2019; Wu et al., 2014). Additional information can also affect the bargaining process (Tu and Lauer, 2009) or cause some providers to adjust prices due to reputational concerns (Christensen et al., 2018).

An increase in transparency can take several forms: from publishing charge prices (Christensen et al., 2018) or median estimated costs (Tu and Lauer, 2009) to equipping employees with privately owned transparency tools (Lieber, 2017; Whaley et al., 2014). In this paper, we study the effect of price transparency on provider choice in a new setting.

For a long time prices negotiated between providers and insurers in the Netherlands were considered private information (Douven et al., 2018). However, in 2016 one of the major insurers in the market, CZ, unexpectedly published a set of prices for procedures below the 885 euro maximum deductible threshold, with main competitors releasing similar information (Kuijper, 2016; De Jong, 2016). In this paper, we use a difference-in-difference approach to estimate the short-run effects that this partial nationwide introduction of a price comparison tool had on health care spending and provider choice. Using a subset of relatively elective and non-emergency dermatological procedures and unique claims data on Dutch health care spending, this paper finds tightly estimated zero short-run effects on both consumer spending and provider choice. We document that there is a clear potential for savings, ranging from 23% to 25% of the price paid among potential savers. As insurance contracts feature a deductible, many patients would financially benefit from this themselves. Nonetheless, our results suggest that consumers do not exploit these financial opportunities.

These results are by no means surprising given the institutional settings in the Netherlands. With no clear price index across providers and over 4000 health care products to choose from, the costs of price comparison can turn out to be quite substantial. Furthermore, in many cases patients face uncertainty about which product will be coded after the visit and similar products within a single hospital can exhibit massive price variation. Recent survey revealed that 23% of patients are not aware they can choose the hospital they go to (Patiëntenfederatie Nederland, 2019). Handel and Kolstad (2015) showed that consumers often may not understand their health care plans or compare alternatives. Arguably, comparing prices among a wide range of health care products can be as complex. Contrasting that with the fact that a vast majority of the individuals in the Netherlands face the lowest deductible of just 385 euro annually which greatly limits the potential for savings, for many the benefits of such price comparison are outweighted by the costs. This may also be reflected in low search rates: website visits were equal to just 2.43% of the average daily provider visits among CZ consumers. Yet, despite authors prior expectation of limited policy effects on consumer choice, we find it important to accurately assess its effects given public pressure that preceded the publication of prices.[1]

This study contributes to a growing literature on the effects of different price transparency policy changes in health care markets, providing reduced form estimates for the short-run demand effects. Most of the existing literature exploits local initiatives such as employer-specific transparency tool introduction that allows comparing prices for health care products (eg.: Desai et al., 2016; or Lieber, 2017) and finds varying results from no change to 10-17% decrease in spending conditional on search or as much as 18.7% decrease in spending and changes to the entire market structure when actively approaching consumers (Wu et al., 2014). The so called New Hampshire experiment is a single contrasting example that involved publishing bundled statewide median esti-

[1] For instance, Open State Foundation filled a lawsuit against Dutch Healthcare Authority for not publishing the prices (Open State Foundation, 2014).

mated prices for approximately 30 common medical procedures. While initial studies of this policy indicated no effects (Tu and Lauer, 2009), a 5 year follow-up by Brown (2019) estimated a 5% decrease in the costs for the patients. The following study takes an advantage of a similar large-scale event at a national level, but in contrast to the existing literature it is also able to analyze the effects of publishing exact contracted and ultimately paid prices. Consequently, the following study uses a transparency initiative where a large sample of individuals across country obtained access to exact information on the prices, a policy difference that in theory should greatly facilitate price shopping among consumers as compared to other such initiatives studied in the literature.

Furthermore, this paper exploits surges in visits on the transparency tool website and the nature of the annual price adjustment in the Netherlands to quantify a short-run demand response that can be confidently attributed to consumer behavior. With much of the literature highlighting low usage rates of the transparency tools (eg. Desai et al., 2016; Chernew et al., 2018), this study also investigates the effect of sending a reminder about the transparency tool few months after the initial publication, an event that resulted in over 105 thousand website visits in the first week and a permanent 60% increase in daily visits. Although the treatment group had access to some price information prior to posting of the reminder and hence results should be interpreted with care, we find no evidence for meaningful decrease in spending that would reflect the increase in website visits.

This paper is structured as follows: Section 2 shortly describes the health care system and the events related to the publication of prices. Section 3 introduces the data. Section 4 describes the models we use for quantifying the effects of the events. Section 5 presents results and finally Section 6 provides our preferred explanation for these findings and a proposal how the system could be changed so that price transparency has an effect.

2.2 Institutional setting

2.2.1 The Dutch health care system

In 2006 the Dutch health care system underwent a major reform that moved it towards more demand-driven service provision (Enthoven and van de Ven, 2007; Rosenau and Lako, 2008).Residents in the Netherlands are required to buy a mandatory health insurance from private insurers. The government determines the coverage of the standardized health insurance package. Dutch enrollees have since 2016 an obligatory annual deductible of 385 Euro.

Insurance is provided by private insurers, which are obliged to accept any enrollee (regardless of health risk or pre-existing medical conditions) without price discrimination, with a few minor exceptions. In particular, insurers are allowed to give a rebate of up to 5% on the premium for collective contracts and are allowed to give a rebate for enrollees who choose an extra voluntary deductible of up to 500 Euro. In addition, insurers are obliged to contract sufficient health care supply to meet the demand of their enrollees. The idea behind the system is that insurers can make profits by contracting health care providers and incentivizing them to provide care efficiently. To mitigate risk selection in the insurance market, the Dutch government runs an elaborate system of risk-adjustment in which insurers are compensated for differences between their populations. Consumers can switch insurers on the annual basis (Kroneman et al., 2016).

Health care providers compete for contracts with insurers. To a large extent (around 70% of hospital revenue), prices of hospitals were liberalized in 2012. In 2005 a case mix system called Diagnosis Treatment Combinations (DTC's) was introduced for the reimbursement of hospital care. According to article 35 of the Healthcare Market Regulation Act hospitals are required to state their invoices in terms of these DTC's, which means that DTC's are comparable between hospitals.

Patients are free to choose any provider of health care conditional on obtaining a referral from their general practitioner who acts as gatekeepers for non-emergency care. Although full reimbursement is available only for hospitals within the network (Kroneman et al., 2016), in 2015 only around 7.5% of enrollees chose restricted plans (Bes et al., 2017; NZa, 2017; this fraction increased to 13.1% in 2017). At the same time, a survey from 2019 revealed that 23% of consumers were not aware they can choose the hospital they visit for services, with 21% choosing the nearest provider and 18% following their GPs advice (Patiëntenfederatie Nederland, 2019).

2.2.2 Information

Motivated by the potential to increase efficiency and contain costs while maintaining quality and accessibility of care (Rosenau and Lako, 2008), the success of a demand-driven health care system such as the one in the Netherlands depends on, among other things, access to information about quality and prices (Enthoven and van de Ven, 2007). In order to ensure competition among market participants as well as promote quality enhancements, consumers should be able to evaluate all the alternatives and make efficient choices (Rosenau and Lako, 2008).

Existing literature suggests that consumers in the Netherlands take quality into consideration when choosing a health care provider.[2] Beukers et al. (2013) found that quality indicators are a significant predictor when choosing a hospital for hip replacement; Varkevisser et al. (2012) estimated that patients are willing to travel 9% further relative to mean travel time for a 1% decrease in a readmission rate following angioplasty. Using the example of cataract surgery Ruwaard and Douven (2014) showed that although 80% of chosen providers are within 20km distance, individuals are willing to travel further, especially for the top performing hospitals.

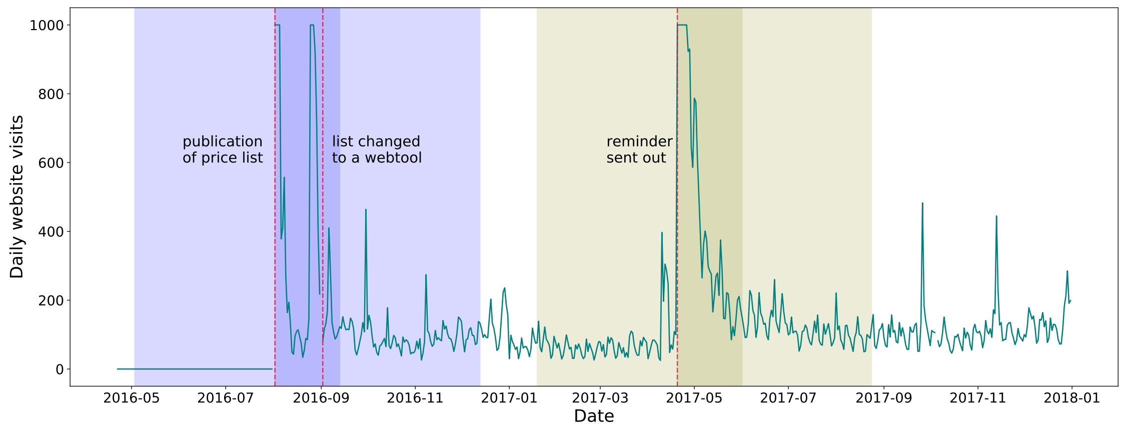
While some degree of information on health care quality is available to consumers in the Netherlands, the same could not have been said about hospital prices which up until recently were kept confidential by both insurers and providers (Douven et al., 2018). In light of the fact that these prices are relevant to consumers, because deductible payments directly depend on them, it is surprising that consumers could not easily compare prices between providers.This situation changed on the 2nd of August 2016 when one of the major insurers in the market, CZ, decided to publish all the contracted Diagnosis Treatment Combinations (DTC) prices below 885 euro (Kuijper, 2016). Soon after, VGZ and Menzis, two other large players in the Dutch market, followed by publishing a similar set of prices (van Bokhorst, 2016; Woldring, 2016). Moreover, another insurance provider, Zilveren Kruis, published more limited data containing only some specific groups of generally described procedures (Skipr, 2016).

Published prices, which are the exact prices contracted between the insurer and provider, became available through freely accessible search engines located at insurer websites as well as external sources such as Consumer Association (Consumentenbond) website that aggregated prices both across insurers and several transparent providers (De Jong, 2016; search tool is now available as an external module; see Open State Foundation, 2019). With these 4 insurers serving 88.3% of the market in 2017 (NZa, 2017), such publication accounts for a major increase in price transparency.

Figure 1 presents daily traffic on the website hosting the search tool published by

[2]Such quality measures can consist of report cards, newspaper rankings or readmission rates, often with specific per-specialism distinctions (Varkevisser et al., 2012; Ruwaard and Douven, 2014).

Figure 1: Daily visits on the CZ website containing the price transparency tool



Note: The figure above plots the amount of daily visits on the CZ price comparison page for the period until 1st of January 2018. The data on prices was initially published in a list format (first dashed vertical line) and replaced a month later by an online tool that allowed for procedure-specific comparison of prices (second dashed vertical line). An email reminder was sent out on 20 April 2017 (third dashed vertical line). In the figure, the maximum number of visits is trimmed at 1000 per day for the clarity of the exposition. Figure 5 in Appendix A zooms in and documents the surge in activity in the first days after the publication of prices and the reminder, respectively. The two shaded areas depict the time periods used in in our main analysis. The darker shades mark the first 6 weeks after the respective event. For some specifications, these data will not be used to account for the fact that it takes time to book an appointment.

CZ, with maximum daily visits trimmed at 1000 for clarity of the exposition. There was a large traffic increase after the publication of prices, with over 75 thousand visits in first week after the event. Interestingly, there is another spike of activity on the 20th of April 2017, which is when a reminder email about the tool was sent to CZ consumers; this resulted in over 105 thousand visits in the first week after the event.[3] Furthermore, both events resulted in a permanent usage level increase. This is trivially the case for the publication of prices, as usage was zero before the introduction of the tool, but also in case of the email reminder where median daily traffic increased by approximately 65%, from 65 in the period of 12 weeks prior to the event up to 107.5 median daily visits in period between 6th and 18th week after the reminder was sent.

Based on these data, back of the envelope calculations indicate that 98 median daily visits over the entire observed period are equal to approximately 2.43% of daily provider visits made by CZ consumers, suggesting that search rates among consumers are relatively low (similar to Gourevitch et al., 2017; or Mehrotra et al., 2017) despite temporal activity rushes.[4] In addition, it has to be noted that the activity presented

[3]See Appendix A for the copy of the reminder sent as well as figures on first weeks post-events.

[4]With approximately 6.9mln procedures in 2017 being recorded (DIS Open Data, 2019) and 21.1% market share of CZ in 2017 (NZa, 2017), this gives an average of 4025 daily visits. Note that while search rates reported in the literature can vary substantially, estimated search rates in the current

resembles visiting the tool's website address and not the actual search activity which may or may not have followed. Furthermore, some consumers may have made several searches during a day. Nonetheless, even such low activity could lead to measurable effects of publishing prices.

The data on prices reveal that a considerable amount of price information is released. For the 176 unique dermatology DTCs there were over 8800 procedure-provider price pairs available online for CZ by the end of 2016, out of a little above 17 thousand provider-specific prices for that year in total. However, at the same time it is important to keep in mind that due to the two-stage nature of the negotiation process where insurers and hospitals first agree on budget and then negotiate the DTC-specific prices throughout the year (Douven et al., 2018) it could be that some prices were not yet available at the date of publication, despite the publication taking place in the second half of the year.

2.2.3 Selection of procedures

The aim of our paper is to study the effect of price transparency on patient behavior when patients had enough time to make choices and could benefit from price transparency. We would like to study this for situations in which treatments are fairly simple and standardized, so that quality differences across providers are less important than for other types of care.

With this in mind, we first looked for a type of care that is high volume, not urgent, widely available, and for which the price is relatively low so that cost-sharing matters. For this reason, we focus on dermatology. There are 176 available dermatological procedures. We selected 6 procedures out of those: 4 outpatient visits for consultations and 2 visits for surgeries. Table 1 contains descriptive statistics for these procedures. Selected products are relatively simple and high volume: they account for approximately 66% of all patient visits within dermatology, with 4 out of 6 selected procedures being top 4 most popular dermatological procedures in both 2016 and 2017. They are also relatively low priced which means that even for a consumer with the annual de-

study can be considered as relatively low. Gourevitch et al., (2017) reported 12% of the sample using transparency tool at least once during 12 months (1% using more than 3 times) while Mehrotra et al. (2017) reported that only 3% of the individuals in the sample compared providers in terms of cost. Desai et al. (2016) reported 10% of the treatment group that was offered a transparency tool searched at least once within first 12 months (this fraction growing to 18% after 24 months).

Table 1: Summary statistics for the selected dermatological procedures

Short Description of the DTC	Year	Provider prices published	Mean Price	Standard Deviation	Minimum Price	Maximum Price	Volume	Percentage of the total
1-2 surgeries	2016	118	452.21	80.73	228.87	737.86	104863	11%
skin cancer or	2017	114	445.68	77.82	231.63	687.27	108622	11%
signs of it	2018	118	438.43	70.58	231.63	696.43	98363	11%
1-2 outpatient visits	2016	116	112.99	21.43	61.42	170.00	249703	27%
skin cancer or	2017	117	112.93	19.05	59.30	185.00	260911	27%
signs of it	2018	120	109.96	16.41	60.19	185.00	240687	27%
1-2 surgeries	2016	120	403.71	64.77	234.84	620.27	39304	4%
benign tumor	2017	120	402.24	66.23	274.50	746.34	37754	4%
of the skin	2018	123	405.82	59.97	275.00	600.00	32972	3%
1-2 outpatient visits	2016	122	113.89	21.41	70.72	225.00	94884	10%
benign tumor	2017	120	112.14	19.41	77.27	185.00	95009	10%
of the skin	2018	123	111.23	17.57	76.92	185.00	85765	9%
1-2 outpatient visits	2016	113	117.31	24.18	70.68	186.67	98842	10%
skin inflammation	2017	111	118.01	20.93	73.29	185.00	100264	10%
or eczema	2018	114	116.71	18.66	76.55	185.00	92213	10%
1-2 outpatient visits	2016	110	122.88	30.70	72.90	222.00	40900	4%
skin conditions	2017	110	120.98	25.29	70.38	211.99	43173	4%
bumps and flakes	2018	115	121.45	22.49	71.44	185.00	39987	4%

Note: This table presents descriptive statistics for the 6 selected procedures. The table was complied using publicly available price data obtained from the Consumer Association search engine (Open State Foundation, 2019). Volumes were compiled using DIS Open Data (2019). Last column denotes the percentage of the total volume within dermatology DTC subgroup.

ductible of 385 euros, these prices can make a difference in the out-of-pocket spending. Most importantly, selected procedures are among DTCs with most prices published. Figure 9 in the Appendix presents distribution of provider specific prices published by CZ for each DTC. A vast majority of DTCs have between 60 and 90 provider specific prices published, whereas the selected procedures are published for 110-122 providers, arguably providing the largest amount of information.

Selected procedures exhibit substantial price dispersion, with the price range often exceeding twice its mean, a disparity that remains large in size across years. Differences in prices may in principle be completely driven by quality differences across hospitals or differences in the level of competition across geographic areas. This is unlikely for the procedures we chose. To provide empirical evidence for this, Figure 2 shows that negotiated prices often do not seem to follow a systematic pattern. In particular, one would expect that when a hospital has high market power or offers high quality services, then it would generally negotiate high prices relatively to the other hospitals. But then, one should see that many prices for that hospital would be above the average. However, as seen in the upper left and lower right quadrants of the scatter plot in Figure 2, a large fraction of hospitals negotiates a price higher than the average for one procedure and price lower than the average for another, very similar procedure.[5] For instance, in 2016 the Antoni van Leeuwenhoek hospital in Amsterdam charged 143.65 euro for 1-2 outpatient skin cancer checkup visits, while the price for a similar procedure, 1-2 outpatient benign tumor of skin checkup visits, was 80.94 euro. The average price for both procedures was about 113 euro. The bar plots on the right show average prices for each procedure-hospital-insurer combination, for 4 out of the 6 procedures. When analysing it from the side of the insurer, one can see that there is no systematic pattern in prices either.

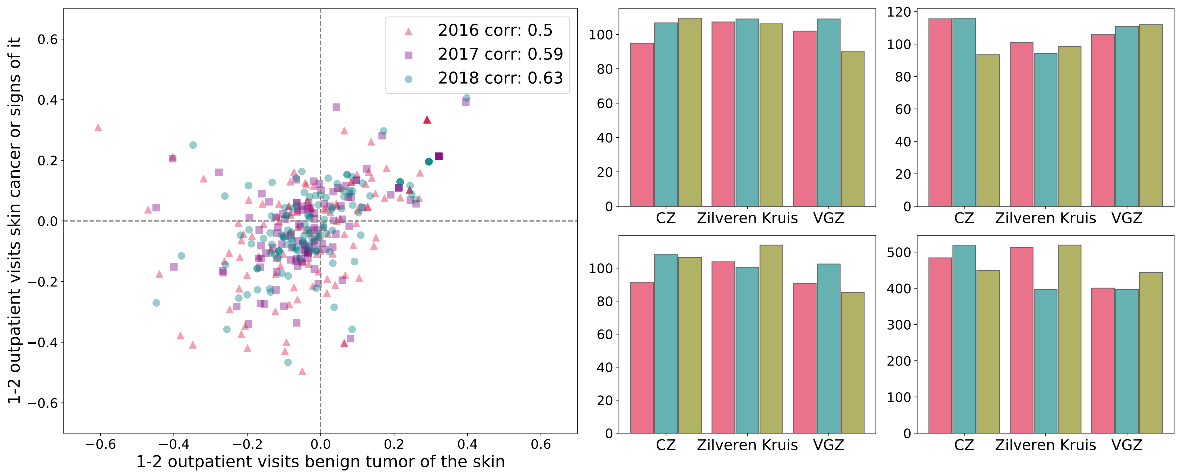
At this point, one may wonder why there is so much unsystematic price variation. The main reason for this is that the contracts between insurers and hospitals go much beyond specifying prices. For instance, they also specify budgets, information exchange and quality requirements among other things.

From the consumer perspective, this gives rise to an additional challenge. It is possible that a consumer who chooses a provider solely on the basis of low price of some anticipated procedure pays more than average because the procedure coded ex-post is actually more expensive than the alternatives in the area. Consider outpatient consultation in case of skin cancer. Mean published prices for CZ across providers is 112.99 euros. Mean price for 20 lowest priced providers is 84.13 euros, which indicates 28.86 euros potential savings or 25.54% savings as compared to the mean price. However, if instead the consumer is coded with outpatient consultations in case of the benign tumor of the skin, these savings from choosing an average of 20 lowest priced providers reduce to just 9.87 euros or 8.74% of the mean price paid. Similar calculations for choosing 20 lowest priced providers for benign tumor of the skin and being coded with skin cancer consultations results in just 7.78% savings instead of 22.45%. While consumers still save by choosing these providers, savings are approximately 3 times lower.[6] Although

[5] See Appendix B for similar scatterplots between different pairs of products; see Douven et al. (2018) for more detailed analysis of correlation between related product groups.

[6] There is a higher correlation in provider prices between similar type of procedures as compared to a similar diagnosis. For instance, average correlation of provider prices across years for checkup and surgery is only 0.49 for skin cancer and 0.43 for benign tumor. In contrast, correlation between skin cancer checkup and benign tumor checkup is 0.58 while similar correlation for surgeries is 0.54. This indicates that even if consumers missclasify the DTC, they may still obtain a relatively cheaper procedure overall, though the correlations between products are still generally low.

Figure 2: Relationship between prices



Note: The scatterplot on the left plots the percentage deviation of the price for one specific DTC (1-2 outpatient visits skin cancer or signs of it; second DTC in Table 1) in one specific hospital and year from the average across hospitals against the percentage deviation of another price (1-2 outpatient visits benign tumor of the skin; fourth DTC in Table 1) from its average. Dots in the upper-left and lower-right quadrants depict hospitals which contracted a price higher than the average for one DTC, and less than the average for another, similar DTC. The bar plots on the right show average contracted prices for 4 out of 6 selected dermatological procedures: 1-2 outpatient visits skin inflammation or eczema (top left), 1-2 outpatient visits benign tumor of the skin (top right), 1-2 outpatient visits skin cancer or signs of it (bottom left), and 1-2 surgeries skin cancer or signs of it (bottom right). Each subplot presents insurer-specific prices for a given DTC contracted with one of the 3 hospitals: Rode Kruis (crimson), Albert Schweitzer (teal) and MC Groep (Zuiderzee Lelystad, Emmeloord, Dronten; olive). The plots show that the price ranking between hospitals is not preserved across insurers and procedures and hence prices negotiated by one insurer are not perfectly informative about prices negotiated by another insurer. See Appendix B for additional scatter plots for other pairs of procedures and the full set of 6 bar plots.

saving opportunities are clearly present in the market due to large price disparities, it remains unclear whether they can be efficiently exploited by the consumers.

2.3 Data

Our goal is to estimate the response to the introduction of price comparison tools in the Dutch health care market on patient behavior. For this, we use individual claims data. Data for the entire population were provided by the Dutch Healthcare Authority (NZa). The data cover spending for the years 2015 to 2017 and contain the date of starting the procedure and the exact procedure code (DTC), provider identity, insurer identity and the price paid for the treatments received.[7]

We restrict our sample to the selected group of 6 transparent and widely available procedures from the dermatology specializations described in Section 2. Starting from this, we keep data for two subgroups of patients: CZ customers and customers of another large insurer who, in contrast to several major players in the market, did not publish price data on the selected dermatological DTCs. The patients of the other insurer serve as the control group in our analysis.[8] Since some insurers are part of larger insurer groups, we limit the treatment sample to entities that have the "CZ" name included in the website address that contains the comparison tool.

In order to separate potential demand effects from annual supply adjustments, we perform the analysis locally in time, by using separate within-year subsets of the data.[9] In particular, we select a subset of 12 weeks before and 18 weeks after each relevant event: publication of prices and reminder sentout. Such sample selection is not only dictated by a desire to study short run demand response to the specified events, but also to avoid noisy data from the beginning and the end of a calendar year when deductible resets and potential health plan switching is occurring. This selection results in two datasets: May 10 until December 6 in 2016 (with the relevant event occurring on the 2nd of August), and January 26 until August 24 in 2017 (with the relevant event occurring on the 20th of April). Figure 1 marks the relevant subsample periods with blue and green shades, respectively.

We also add several consumer characteristics: 8 age bins, gender, dummy variables for deductible level chosen, any additional health care package, dental plan or collective insurance and the registered location of the patient: postcode, municipality and province. In case of within year change to any of these characteristics, we use the characteristic level that was the most frequent in the year. There are some missing values

[7] This price is paid by the insurance company unless the budgetary agreement specifies otherwise. If the treatment falls under the deductible, then the insurance company collects the payment later from the patient. This means that unlike in other settings where patients hand in bills and get reimbursed, we have no missing data on treatments received.

[8] Due to confidentiality issues, neither the identity of this other insurer nor descriptive statistics can be revealed. However, we provide additional evidence that validates the approach taken in the modeling part of the paper.

[9] Although prices are set annually, which should limit the degree of strategic price setting with regards to price transparency, price adjustments within a year are possible. However, data suggest that this possibility mainly concerns lower-volume treatments. Across different procedure groups and years there are within-year price changes for 16-17% of all unique prices, but only 1.3-1.9% of actual observations in the dataset. Hence even if they were an effect of price adjustment and not health plans, the fraction of observations affected is negligible. This observation further suggests that while strategic price adjustment remains a possibility, its effects for consumers should be negligible as well.

on these characteristics, but these result in dropping only approximately 0.2-0.3% of the data.

We restrict the samples to only individuals above 18 years old, which covers little over 94% of the sample.[10] Some visits cannot be matched with distance due to missing provider postcode but this does not affect the main results. Some observations are either duplicated or indicate that consumer visited two different providers for the same procedure within the same day. Since this is rather improbable and the observations constitute less than 0.1% of the samples, we exclude them as well. Finally, we drop 9 observations with zero recorded price.

We assume that if a procedure was transparent in a given year, price data for most providers was already available to consumers at date of the relevant event. While this is a plausible assumption for the publication of prices in August, it is slightly less likely for the reminder sentout since for many hospitals prices may not yet have been available in late April. Although we expect that a substantial amount of prices is already agreed upon and available for publishing if CZ decided to replace old set of prices with new ones, the estimated effect should be interpreted as the effect of such partial information gain, with an expectation that less information was available at the second event relatively to the first one.

Since prices published by CZ are publicly accessible, it can be argued that individuals from the control group were able to view them as a proxy for their own prices. This is especially a concern given that the initial publication of prices by CZ received a substantial attention in the media and there is a high probability that many clients of insurers other than CZ visited the site, contributing to large traffic at the time. Graphs on the right of Figure 2 present prices published in 2016 by 3 transparent hospitals for procedures contracted with two publishing insurers: CZ and VGZ, and Zilveren Kruis who did not publish these particular prices. One can observe that price rankings between hospitals are not preserved among insurers: a hospital that is least expensive for CZ may turn out to be the most expensive for VGZ or Zilveren Kruis. Yet, when comparing prices between CZ and VGZ for 2016, average Pearson correlation coefficient among 6 selected procedures is equal to 0.69, implying that prices of different insurers can indeed be used as proxies.[11] Given that, we use the reminder sentout event that was

[10]Children until 18 years old do not have to pay a deductible.

[11]Similarly, average per-DTC price correlation between CZ and Menzis and CZ and VGZ in 2017 is 0.663 and 0.706, respectively. In contrast, correlation between prices of CZ across years is 0.692 for years 2016-2017. These correlations are restricted to specific years based on availability of published prices that can be meaningfully compared.

Table 2: Potential savings for the selected procedures

	Publication of prices (August 2, 2016)				Reminder (April 20, 2017)			
	Cheaper and closer	% of the sample	Cheaper and within 10km	% of the sample	Cheaper and closer	% of the sample	Cheaper and within 10km	% of the sample
Treatment	8%	100	9%	100	7%	100	8%	100
	25%	30	23%	35	23%	30	22%	36
Control	8%	100	11%	100	6%	100	8%	100
	23%	33	22%	46	18%	32	17%	45

Note: This table provides information on the potential savings. We distinguish between the treatment and the control group. Numbers are for the pre-event period and for the 6 selected DTCs. For each group, the first row presents results averaged over the entire subsample while the second row only averages over the individuals who can save, with the fraction of population averaged over reported in "% of the sample" column. The column "Cheaper and closer" denotes average percentage savings if individuals would choose a provider that is both cheaper and located closer; "Cheaper and within 10km" denotes average savings if individuals would choose a cheaper provider within the distance of 10km. All percentage values are calculated with reference to the actual amount paid. We use a maximum of 100km distance from the individual postcode to construct the set of alternative providers. Some providers may have several postcodes; since we only observe the choice of the provider and not the location, we assume that the individual visits the closest location among the available ones.

delivered only to the consumers of CZ as a second event that substantially increased price transparency.

It is important for the validity of the results (details below) that both treatment and control groups face similar saving opportunities before the events (see Appendix C for additional evidence for pre-event similarities between both groups). To assess this, we check whether there are cheaper providers of the same procedures within a reasonable traveling distance and whether there are providers that are both closer and cheaper than the one that was selected. A breakdown of the saving opportunities for the consumers is presented in Table 2.

Overall, consumers can save up to 6-11% by choosing a cheaper provider within their choice sets, with savings being substantial even when consumers are to choose a provider that is both cheaper and closer to their location. When aggregating only over individuals who can save from switching, savings can be as high as 25% of the price paid, with a fraction of the population that can save ranging from 30% to even 45% of the subsample. More importantly, potential savings before the policy events are similar for both treatment and control groups: the treatment group faces marginally larger opportunities, which may be a result of larger price dispersion seen in Figure 11, but at the same time a higher fraction of the control sample can save (32-45% compared to 30-36% among the treatment group). This is particularly important: although insurers may have different market shares and negotiate different prices across regions, saving opportunities are substantial (which is not surprising given the large dispersion of prices reported earlier) but also similar and available to both groups. Under the hypothesis that consumers make efficient use of the tools, these opportunities should be exploited by the treatment group once prices become public.

2.4 Empirical approach

2.4.1 Effects on spending

Our aim is to estimate the effects of making a price comparison tool available to CZ customers and reminding them of this tool by email. For this, we use claims data over time. There is a control group, which allows us to use a differences-in-differences estimator. For this, we pool observations across procedures and specify

$$\log(p_{ijklmt}) = \beta \cdot (CZ_i * Post_t) + X_i'\alpha + \delta_j + \eta_k + \theta_l + \gamma_t + \kappa_m + \epsilon_{ijklmt}, \qquad (1)$$

where p_{ijkmt} is the price paid by the consumer i for procedure j, who is insured with k and receives the treatment on day m of week t, CZ_i indicates that i is a CZ insuree, and $Post_t$ indicates the time periods after the introduction of the price comparison tool or the reminder (we perform separate analyses for the two events). We include a set of controls in vector X_i with coefficient vector α: gender, age, type of health care package, a dummy variable for collective contract and level of the voluntary deductible. We also control for fixed effects of procedure (δ_j), insurer (η_k), province (θ_l), week (γ_t) and day of the week (κ_m).

Our main parameter of interest is β. We use a log specification. Therefore, β is the percentage change in the price paid after the introduction of the price comparison tool or after the email reminder was sent out.

Equation (1) is estimated by ordinary least squares, following Brown (2018), Lieber (2017) and Desai et al. (2016). There may be a correlation between individual spending within the same household over time. Since we do not distinguish households and most of the individuals appear only once in the samples, we cannot add fixed effects to account for this in a fashion similar to Lieber (2017). Instead, we cluster the standard errors at the four digit postcode level which should not only account for correlation within households over time, but also subtle location differences such as public connection routes or local GP referral preferences. To minimize the chance that our results are driven by outliers, we follow Lieber (2017) and also estimate a winsorised version of the baseline equation where we trim 5th and 95th quantiles, separately for each procedure in the each sample but jointly for both insurers.

Publication of prices creates a solid baseline for a difference in difference design since one group of consumers gained access to their negotiated prices while both groups had similar information on prices before the publication. While we also attempt to exploit the second surge in website activity resulting from sending a reminder to CZ consumers, it has to be highlighted that in this particular case treatment group had access to some degree of price information before the event. While new prices were published within 2 weeks prior to the date of the email reminder, CZ tends to keep old prices as a reference due to correlation with newly negotiated ones.[12]

Correlation between prices of CZ for the selected dermatological procedures across years 2016 and 2017 was equal to 0.692. This means that reminder sentout that closely followed publication of new prices for year 2017 can be treated as increasing the accuracy of information already available to the CZ consumers. Such increase in information is far smaller than publishing previously unknown prices. In practice, however, reminder sentout resulted in far more website traffic than the initial publication of prices, and since the email was sent to CZ insurees, majority of this traffic can be confidently attributed to the treatment group. Consequently, using this event in addition to publication of prices solves some of the shortcomings mentioned before: control group is unlikely to visit the page at the time, consumers should already be more familiar with the idea of price transparency and are far more likely to still be under the deductible.

One of the potential problems with creating a treatment variable in the specification above is the fact that one should account for the waiting time between signing up for treatment and actually receiving it. Using waiting times data for dermatological procedures in 2016 and 2017, we determine that on average the waiting times are little above 3 weeks, with 90th quantile of waiting times distribution equal to 6 weeks. To avoid a situation where some visits are appointed before the event while others are not, we estimate the baseline equation (1) while only using a subsample that excludes first 6 weeks after the event. This way we can ensure that the transition period does

[12] As informed by the insurer, the new prices were published between 5th and 19th of April, with the reminder being sent on the 20th. Unfortunately, it was impossible to track the exact date of the publication; see Appendix A for further discussion of the price publication event.

not affect the results, providing a more direct estimate of the level difference. Since procedures obtained on the weekend may not be elective and therefore less susceptible to shopping, we further estimate the baseline equation while excluding observations on procedures obtained on either Saturday or Sunday. We also attempt to remedy potential discrepancies in prices negotiated by the insurers by including large municipality fixed effects that account for local variation in negotiated prices.[13]

Formally, our main identifying assumption is that the error term in (1) is not correlated with the right hand side variables. Importantly, for this to hold, the usual "common trends assumption" needs to hold. In words, this assumption is that the evolution of the outcome over time is the same for the treatment and the control group. In order to assess this, we revisit equation (1) and replace the treatment dummy with weekly insurer fixed effects in order to determine whether parallel trends in spending are present among the two groups of consumers. Specifically, we estimate the model

$$\log(p_{ijklmt}) = \beta_{kt} + \alpha X_i + \delta_j + \eta_k + \theta_l + \kappa_m + \epsilon_{ijklmt}, \tag{2}$$

where β_{kt} denotes weekly insurer fixed effects separately for the treatment and control group. Then, we plot the fixed effects for the two groups over time and compare the evolution in the pre-treatment period between the two groups. It is useful to also estimate fixed effects for the post-treatment period. If evidence in favor of parallel trends has been gathered from data for the pre-treatment period, then one can use the plot for the post-treatment period to get a first idea about the size of the treatment effect and whether it is constant over time. If a treatment effect is present and the effect is constant over time, then one should see that the curve for CZ patients is shifted, but otherwise the evolution is the same. And if the treatment effect is small, then such a plot will indicate this as well.

In addition, we estimate the baseline equation over the period of 12 weeks before the event with an artificially created placebo treatment dummy in the middle of that period. The aim of this exercise is to ensure that any estimate found was not present directly before the event and that no effect is found just by virtue of a large sample size.

Finally, we estimate the baseline equation with an addition of a linear trend for the treatment group over the entire sample period. This addresses the concern that trends

[13] Fixed effects are only added for municipalities that have at least 100 observations in the pre-policy period; see Appendix C for further information.

are different. This specification can also be used to formally test whether pre-trends are similar if one makes the additional assumption that the treatment effect is constant (which—anticipating the results—is implied by a zero treatment effect).

2.4.2 Effects on choices

Given that prices are largely fixed within a year, any effect on prices paid by the consumers should also be reflected in the underlying provider choice. Moreover, consumers may not react to price differences but instead learn about new alternatives in the area or browse through insurer website to learn about non-monetary provider characteristics, resulting in altered choices despite no effects on the average prices paid. To investigate that hypothesis further, we estimate parameters of the model

$$newProvider_{ijklmt} = \beta \cdot (CZ_i * Post_t) + X_i'\alpha + \delta_j + \eta_k + \theta_l + \gamma_t + \kappa_m + \epsilon_{ijklmt}, \quad (3)$$

where the dependent dummy variable $newProvider$ denotes whether an individual visited a provider that is not in the pool of providers visited by the same patient within 12 months prior to the current visit date. This is a proxy variable that should provide some indication for novelty in choice. The equation includes similar controls and fixed effects as equation (1) and is estimated with least squares, with standard errors clustered at the postcode level.[14] Since switching behavior is only observed conditional on visiting a provider in the last 12 months, the model is estimated over a subsample of individuals.

In a similar fashion as before, we also estimate equation (3) with additional 6 weeks delay to account for the waiting times. We further estimate another variant of equation (3) where we exclude observations related to treatments that were received on the weekend, since such visit may be relatively less elective. Since a one year time span may be considered too long, we estimate a version of the equation (3) where $newProvider$ takes a value of 1 if the last provider visited for a DTC within last 12 months is different from the current choice. Finally, we run a placebo test on the pre-policy subsample, add a linear trend to the baseline specification and estimate weekly fixed effects, in each case investigating whether both groups follow similar trends in a similar fashion as in the case of price regressions.

[14] Results are qualitatively similar when we use a logit model.

A change in provider choice can also be reflected in the choice of the provider type. Therefore, we also investigate whether consumers are more likely to visit a free-standing facility (ZBC; often specialized clinics that offer outpatient procedures), by estimating the model

$$ZBC_{ijklmt} = \beta \cdot (CZ_i * Post_t) + X_i'\alpha + \delta_j + \eta_k + \theta_l + \kappa_m + \gamma_t + \epsilon_{ijklmt}.$$

These facilities may be generally less frequently visited than hospitals, but access to the transparency tool may increase their salience and result in higher probability of a visit. Finally, the effects of transparency change can have an effect on the distance traveled since consumers may find a cheaper provider located further than the most salient close alternative or become aware of a provider that is located closer to them. More specifically, we specify:

$$distance_{ijklmt} = \beta \cdot (CZ_i * Post_t) + X_i'\alpha + \delta_j + \eta_k + \theta_l + \kappa_m + \gamma_t + \epsilon_{ijklmt}.$$

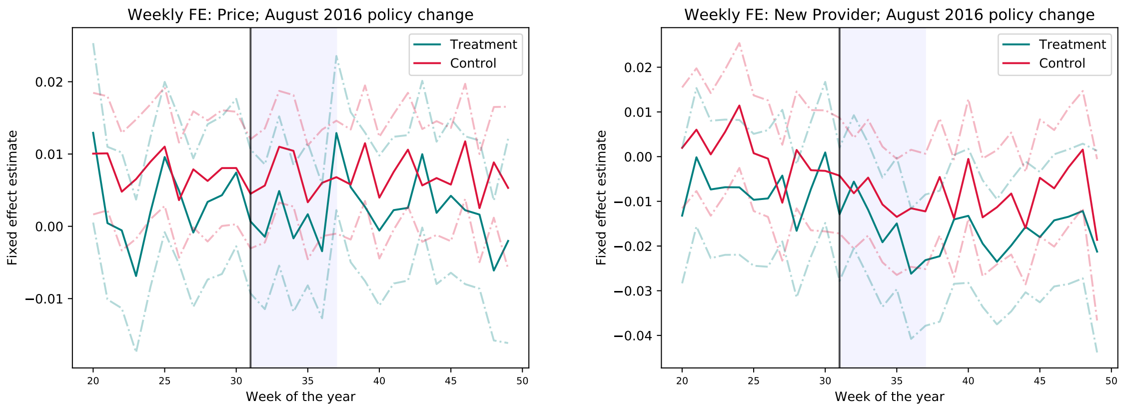
As before, we estimate the parameters with ordinary least squares, but without clustering the standard errors at the postcode level. Since some providers may have several locations and we only observe the choice of the provider and not the exact location, we assume that the consumer goes to the closest among the available locations of the provider. Although plausible, this is a simplifying assumption since not every location may be offering the dermatological procedures and hence the results should be interpreted with caution.

2.5 Results

Figure 3 shows the evolution of main outcomes over time. The graph on the left is for the price consumers paid and shows very similar, almost flat trends for both treatment and control group. There is no indication of a treatment effect (see discussion below equation (2)). The graph on the right is for the likelihood to choose a new provider and shows a small downward trend in the first few weeks, yet also here there is no visible difference in the evolution between the groups both pre- and post-treatment and hence also no indication of a treatment effect.

The estimation results for model (1) are presented in Table 3. The baseline estimate

Figure 3: Evolution of main outcomes over time around the time of the publication of prices



Note: The graphs above presents estimates of insurer-week fixed effects. Based on equation (2) and the corresponding equation with outcome variable $newProvider_{ijklmt}$. Solid lines denote the fixed effects, plotted separately for treatment and control groups. Confidence intervals are constructed using clustered standard errors. Weekly fixed effects are normalized with respect to the first week of the control group (first week omitted in the figure). Vertical lines denote the event date.

of a 0.04% decrease in price paid is insignificant and, given low standard errors, points towards a precisely estimated zero effect of the policy. The effect remains similar in magnitude and insignificant when using a winsorized version of the dependent variable. Although the initial publication in a list format or potential booking delays could have affected the results, excluding the first 6 weeks post event results in an insignificant estimate that further changes in sign. Excluding weekend days from the sample or adding large municipality fixed effects does result in slightly higher estimates in terms of absolute values (-0.07% and -0.13%, respectively) that nevertheless remain insignificant. Importantly, the model passes the placebo test and the estimate of a linear trend for the treatment group is not significant at any conventional level, giving supportive evidence for the difference-in-difference approach taken in this paper.

The reduced form evidence for policy effects on consumer choice are displayed in Table 4. While it is possible that patients pay the same prices while choosing different providers, these results are generally in line with the view that price transparency had no effect on either. There is no significant effect of price publication on choosing a provider different than the pool of providers visited for a DTC within last 12 months. The results are further robust to adding a delay, excluding weekend days or reducing the pool of the providers to just the last visited location. Furthermore, there is no

Table 3: Effects of publication of prices on consumer spending

	Base	Winsorized	Delay	NoWeek	Municipal	Placebo	Base^	Municipal^
$CZ_i * Post_t$	-0.0004	-0.0005	0.0004	-0.0007	-0.0013	0.0026	0.0008	-0.0001
	(0.0014)	(0.0012)	(0.0016)	(0.0016)	(0.0013)	(0.0025)	(0.0030)	(0.0028)
$Linear\ Trend$							-0.0001	-0.0001
							(0.0002)	(0.0002)
Fixed effects	Yes	Yes	Yes	Yes	Partially	Yes	Yes	Partially
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.923	0.945	0.923	0.927	0.933	0.922	0.923	0.933
Observations	190983	190983	151104	159443	190983	72367	190983	190983

Note: The table presents treatment effect estimates for the log price regressions. "Base" displays results for baseline equation. "Winsorised" estimates the baseline equation while using a winsorised dependent variable, with trimming at 5th and 95th quantiles. "Delay" estimates the baseline equation while excluding the first 6 weeks after the event. "NoWeek" estimates the baseline model while excluding the weekend days from the sample. "Municipal" adds large municipality dummy variables to the baseline model. "Placebo" estimates the baseline model over the pre-policy period, with an inclusion of placebo dummy variable for treatment insurer in the middle of that period. "Base^" and "Municipal^" add linear trends for treatment group over the entire sample. Fixed effects include procedure, insurer, province, week and day of the week. Individual controls include age, gender, type of health care package, collective contract and level of the deductible. All regressions are estimated using ordinary least squares, with standard errors clustered at the 4 digit postcode level.

significant effect of the event on choosing a ZBC type provider or on the distance traveled to the chosen provider. The overall robust evidence suggests that publication of prices by CZ had no short-run effects on either prices paid or the underlying choice of the company's consumers as compared to a control group from another insurer. In each case, the results indicate a tightly estimated zero effect of the event, supporting the conclusions that substantial consumer awareness about the existence and availability of price information does not necessarily result in short-run demand effects.

Figure 4 shows the evolution of the main outcomes around the time of the email reminder. There is a visible difference in trends between treatment and control group. The treatment group has a slightly more negative trend in case of the price regression as seen in the graph on the left of Figure 4, and a more positive pre-trend in case of the new provider regression as seen in the graph on the right of Figure 4 (as the trend is close to zero and the trend for the control group is negative). While these differences do not seem substantial in case of the price regression, they may in both cases invalidate the difference-in-difference approach when we do not take this into account. As discussed before, we remedy this by including an additional linear time trend for the treatment group.

Table 5 presents estimation results for the same set of models as Table 3, now estimated using the posting of the email reminder as the relevant event. The last two

Table 4: Effects of publication of prices on consumer choice

	Same as previous provider(s)						ZBC provider type		Distance	
	Base	Delay	NoWeek	Last	Placebo	Base^	ZBC	ZBC^	Dist	Dist^
$CZ_i * Post_t$	-0.002	-0.002	-0.002	-0.001	0.005	-0.002	0.003	-0.003	0.1446	0.2023
	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)	(0.003)	(0.005)	0.1427	(0.2708)
Linear Trend						0.000		0.000		-0.0039
						(0.000)		(0.000)		(0.0158)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.054	0.053	0.046	0.031	0.051	0.054	0.094	0.094	0.085	0.085
Observations	159255	125934	128727	159255	59442	159255	190983	190983	190971	190971

Note: The table presents treatment effect estimates for the choice regressions. "Base" presents baseline results for choosing a new provider different than the pool of providers visited within last 12 months. "Delay" estimates the baseline model with exclusion of first 6 weeks post event. "NoWeek" excludes weekend days from the sample. "Last" modifies the dependent variable to take value of 1 only if the last visited provider 12 months prior to the procedure is different than the current one. "Placebo" estimates the baseline model over the pre-policy period, with an inclusion of placebo dummy variable for treatment insurer. "ZBC" estimates the baseline model with ZBC (specialized clinic) dummy as the dependent variable. "Dist" estimates the baseline model with distance in kilometers as the dependent variable. "Base^", "Dist^" and "ZBC^" add linear trends for the treatment group over the entire sample. Fixed effects include procedure, insurer, province, week and day of the week. Individual controls include age, gender, type of health care package, collective contract and level of the deductible. All regressions are estimated using ordinary least squares and except for the "Dist" and "Dist^" specifications, standard errors are clustered at the postcode level.

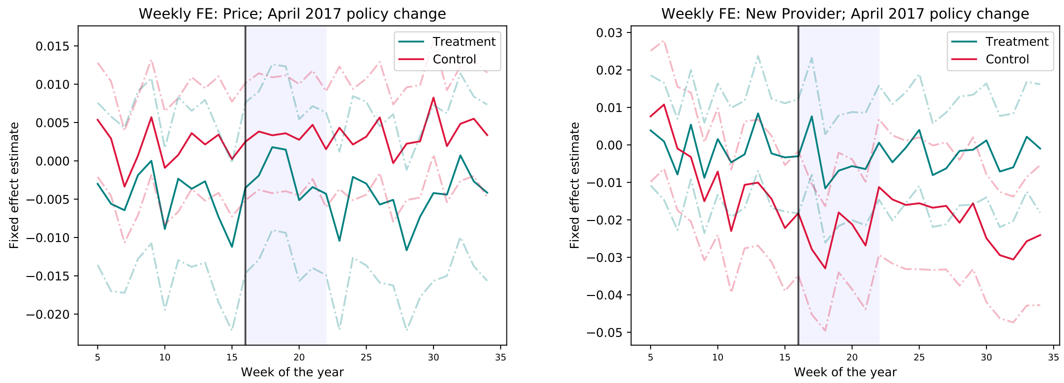
columns confirm that there is a significant difference in the time trend (columns 'Base^' and 'Municipal^'), as already indicated in Figure 4. Once we control for this, we find insignificant effects of the email reminder. Finally, results for the effect of reminder sentout on consumer choice are presented in Table 6. Also here, once we control for linear trends, we find insignificant effects of the email reminder. Furthermore, with no associated trend in the website traffic data - in fact, prior to the reminder sentout seems website traffic seems to be gradually decreasing - it is unlikely that the estimated trend is related to usage of the price comparison tool.

2.6 Discussion

This paper finds a tightly estimated zero effect of the publication of prices on health care spending and provider choice in the Netherlands. This result stands in stark contrast to the majority of the literature. At first this is a surprising result, as consumers did visit the website on which they could look up prices (Figure 1) and there were opportunities to save money (Table 2).

One possible explanation is that large shopping opportunities that are in principle

Figure 4: Evolution of main outcomes over time around the time of the reminder email



Note: The graphs above present estimates of insurer-week fixed effects. Solid lines denote the fixed effects, plotted separately for the treatment and control groups. Confidence intervals are constructed using clustered standard errors. Weekly fixed effects are normalized with respect to the first week of the control group (omitted in the figure).

Table 5: Effects of email reminder on consumer spending

	Base	Winsorized	Delay	NoWeek	Municipal	Placebo	Base^	Municipal^
$CZ_i * Post_t$	-0.0005	-0.0005	-0.0017	-0.0006	-0.0020*	-0.0024	0.0044	0.0032
	(0.0014)	(0.0011)	(0.0014)	(0.0015)	(0.0012)	(0.0023)	(0.0028)	(0.0024)
$Linear\ Trend$							-0.0003**	-0.0004**
							(0.0002)	(0.0001)
Fixed effects	Yes	Yes	Yes	Yes	Partially	Yes	Yes	Partially
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.936	0.960	0.936	0.938	0.949	0.935	0.936	0.949
Observations	186814	186814	151364	157170	186814	74113	186814	186814

Note: See notes to Table 3.

available to the consumers may be difficult to exploit because consumers are unable to know about them, even when the information is in principle available. In the descriptive part of this paper we have shown that prices for very similar procedures offered by the same provider are often very different, in unsystematic ways. Combined with a complex system of DTC relations and uncertainty over which particular procedure will be applied consumers may therefore find it challenging, if not hopeless, to efficiently shop among the available providers. In addition, 87.7% of consumers in 2017 faced only a 385 euro annual deductible (NZa, 2017). Therefore, the incentive for price shopping may be too small to invest into understanding what the different DTCs are so that the prices for the right DTC can be compared.

So, alongside the strong evidence for website usage and consequently a considerable

Table 6: Effects of email reminder on consumer choice

	Same as previous provider(s)						ZBC provider type		Distance	
	Base	Delay	NoWeek	Last	Placebo	Base^	ZBC	ZBC^	Dist	Dist^
$CZ_i * Post_t$	0.014***	0.013***	0.014***	0.007	0.015***	0.005	-0.008***	-0.002	-0.0495	-0.4080
	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.003)	(0.005)	(0.1446)	(0.2870)
Linear Trend						0.001*		-0.000		0.0241
						(0.000)		(0.000)		(0.0161)
Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.054	0.054	0.049	0.034	0.053	0.054	0.099	0.099	0.086	0.086
Observations	155997	126170	127065	155997	61545	155997	186814	186814	186754	186754

Note: See notes to Table 4.

awareness about the availability of the tool and product prices, this may indicate that consumers either checked the website out of interest and did not make use of the information gained or the system is too complex and costly to efficiently shop for the health care products. This conclusion is highly in line with Semigran et al. (2017) who noted that while consumers support the price transparency concept, they face several barriers to efficiently use the tools available.

For data availability reasons, we do not distinguish between consumers who have already crossed the deductible limit (and are therefore not subject to cost-sharing) and those who have not. This, however, does not invalidate the conclusion that access to price information, given estimates with very low standard errors, had no significant overall effect on spending at the population level. It is possible that with two groups of consumers: one under the deductible choosing low priced products to save and the other over the deductible choosing high priced products to proxy for expected higher quality, the overall net effect of price transparency initiative would be zero - exactly what is observed in this study. However, it has to be noted that we also did not find any evidence of direct change in provider choice, type of provider chosen or distance travelled and so we concluded that this scenario is unlikely.

Note also that the results are limited to a selection of arguably the most elective dermatological procedures that are likely to fall under the minimum deductible. Further research could inspect if the results remain qualitatively the same when considering other specializations or types of products.

One may wonder how this could be changed. Instead of using the negotiated DBC-prices as a basis for the deductible payments, insurers could state fixed prices for deductible payments for procedures that patients are likely to understand, such as for outpatient visits, drugs prescriptions, or surgery.[15] This price could then be multiplied

[15]Indeed, a few years ago the Dutch Association of Hospitals has proposed to simplify the deductible

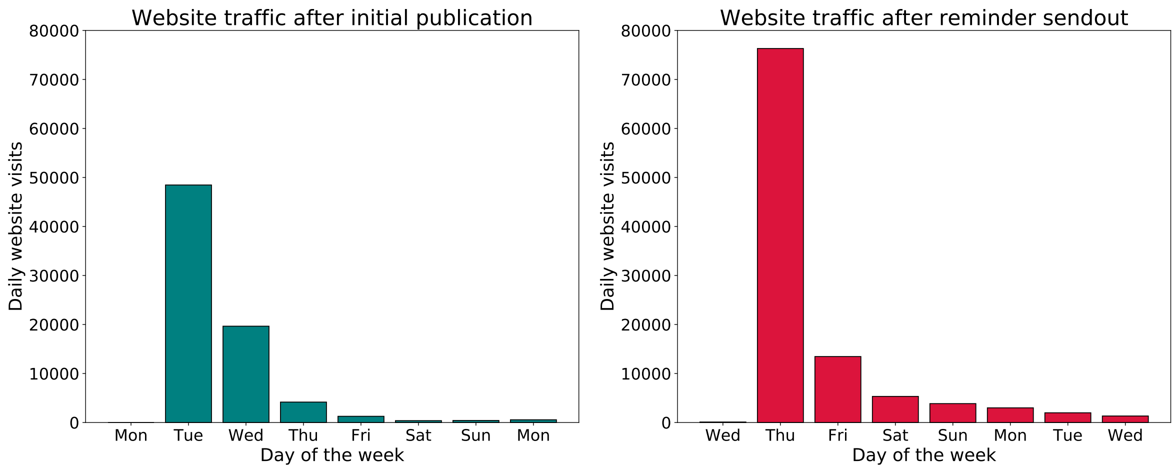
with a provider-specific factor that is related to the average reimbursement the provider receives from the insurance company. If, on average, prices are higher for a provider, then this factor should be bigger than 1, otherwise smaller. Combined with clear information on quality of the provider, consumers should find it easier to efficiently shop for health care services. In addition, one can also consider increasing cost sharing for the patients so that the potential savings overweight the costs of search.

To conclude, our results suggest that price transparency alone does not have an effect on market outcomes in the Dutch settings. Our intuition for this result is that the health care system is too complex from the patient perspective. This suggests that if policy makers would like to enhance competition between providers, then they should simplify the system for the patients. This could be done even without changing the deductible level and would tentatively lead to lower out-of-pocket payments, as it would make it easier for patients to price-shop.

payments (Van Rooy, 2017) and the health insurer Menzis started to implement this proposal in 2017 (Van Aartsen, 2017). But policy makers did not support this and, to date, no big changes have been made. The reason is that any discussion related to cost-sharing is considered politically sensitive in the Netherlands.

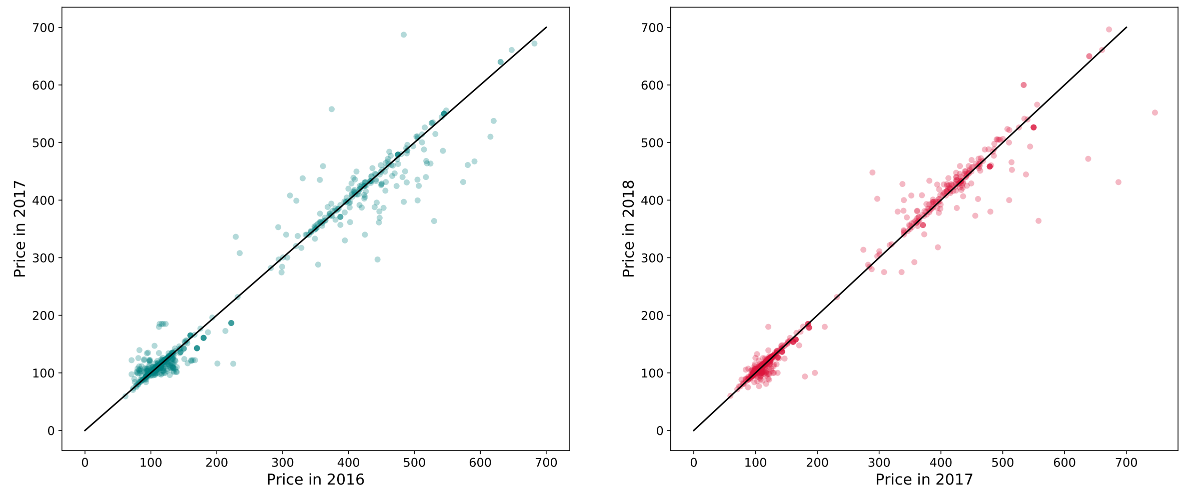
2.7 Appendix

Figure 5: Website traffic in the respective first week after the events



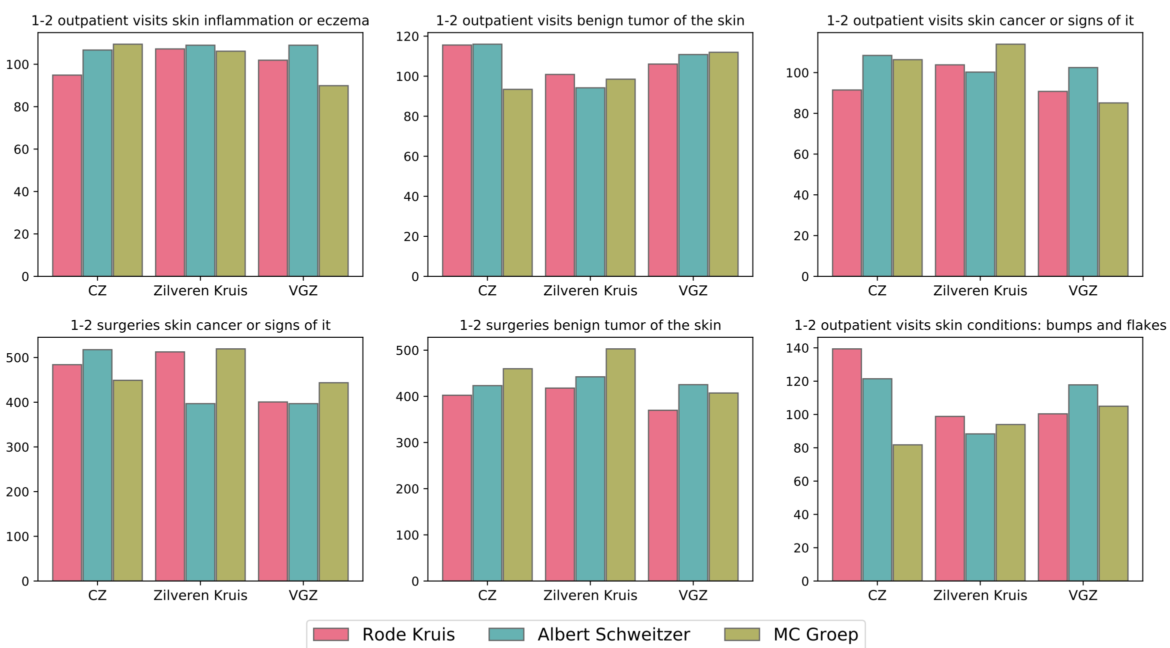
Note: The figure on the left presents the surge in activity in the first week after the publication of the price list. The figure on the right presents the daily amount of visits in the first week after the reminder email was sent out. Note that each figure starts 1 day before the actual event and that traffic on these days was very low. Preceded by a significant social pressure (Open State Foundation, 2014), the publication of price data by CZ received considerable attention. Figure 1 displays the daily traffic on the website that contained the price information published by CZ. Despite the fact that initially prices were published in a list format, there was a substantial amount of traffic on the website that continued well into September and October of the year. Not surprisingly, initial publication resulted in almost 75 thousand visits in the first week after the event, with the majority of that activity happening in the first two days (displayed more clearly on the left of Figure 5). There was also a third (trimmed) spike towards the end of August, but its magnitude was relatively small (around 10 thousand views overall) and there is no explanation for it. It is possible that there were some media reports in this period that related to the publication of prices before it happened, resulting in an increase of the visits on the website. Overall, between 2nd of September 2016 (excluding the list format publication period of August) and 4th of October 2017 there were over 153 thousand visits to the website, with almost 84% of them being unique; users spent on average 2 minutes and 14 seconds on the website. Narrowing this down to the study periods, as indicated in Figure 1, the median number of daily visits was 98, with this amount varying from 51 (10th quantile) to 255 (90th quantile). It remains unclear, however, how many of these visits were followed up by any subsequent search activity as the visits are recorded on the search tool URL. There was an even bigger surge in traffic on the website right after the 20th of April 2017, a result of an email about the comparison tool (displayed in Figure 8) being sent out to the CZ consumers. Since the reminder was sent to CZ consumers only, it guarantees that the vast majority of over 105 thousand visits in the first week after that were made by individuals for whom the prices were indeed relevant. In contrast, the initial publication may have gained attention of the consumers of other insurers as well who, perhaps as a result of the media coverage, visited website out of curiosity.

Figure 6: Relationship between prices negotiated by CZ across years



Note: Figures above present scatterplots of prices in 2017 and 2018 against prices in the respective previous year. Each dot is an available provider-procedure pair. It is often the case that while prices for the current year are not yet available, CZ will keep the old prices on the site for consumer use. A natural question arises whether these are a good indicator of actual prices being paid. While there is some adjustment across years, overall prices seem to be quite similar indicating that old prices could potentially be used as proxies for the new ones. At the same time, they are not perfectly correlated and therefore such information remains imperfect.

Figure 7: Differences in prices across insurers



Note: The bar plots show average contracted prices for selected dermatological procedures. Each subplot presents insurer-specific prices for a given DTC contracted with one of 3 transparent hospitals. The figure shows that the price ranking between hospitals is not preserved across insurers and procedures and hence prices negotiated by one insurer are not perfectly informative about prices negotiated by another insurer. See main text and notes to Figure 2 for additional details and discussion.

Figure 8: Part of the email sent to CZ consumers informing about the transparency tool

Wat kost úw behandeling?

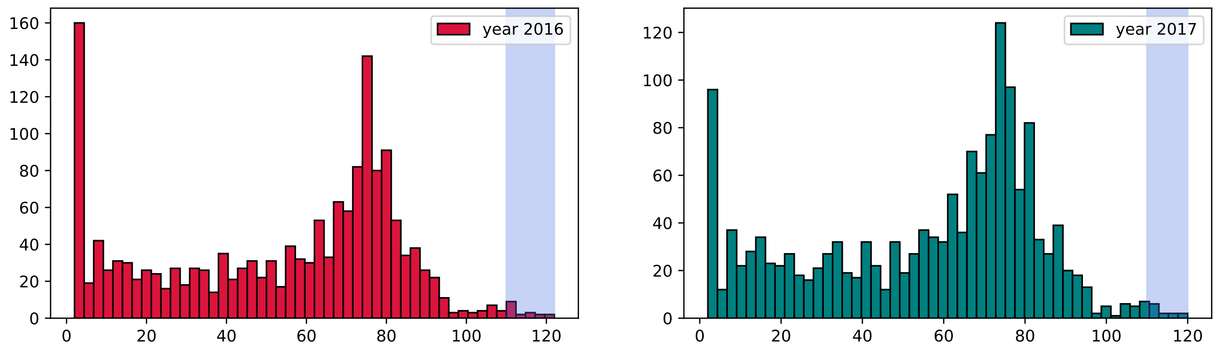
Wat kost een bezoek aan de specialist eigenlijk? En hoeveel betaalt u voor een operatie aan neus- of keelamandelen?

Veel mensen weten niet wat de prijs van een behandeling is. En worden verrast door de rekening. CZ geeft u graag meer inzicht. Bij ons bekijkt en vergelijkt u eenvoudig een groot aantal ziekenhuistarieven tot € 885. Zo weet u waar u aan toe bent.

Dit kost uw behandeling

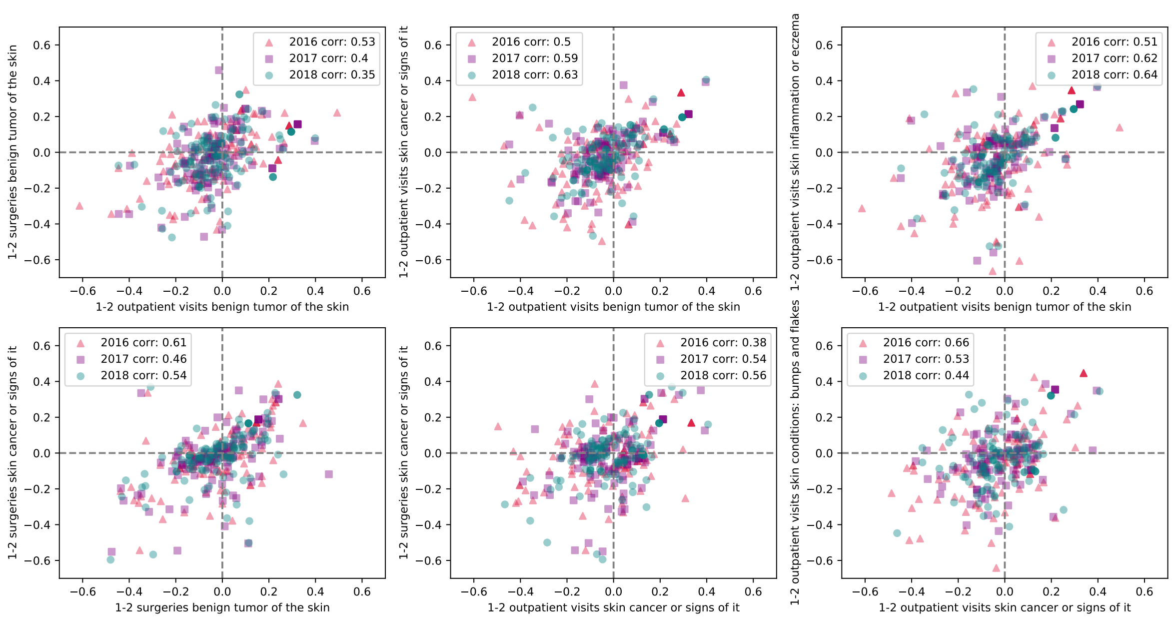
Note: This figure shows the main part of the reminder email. It reads: "What does your treatment cost? What does a visit to the specialist actually cost? And how much do you pay for an operation for nasal or throat tonsils? Many people do not know what the price of a treatment is. And are surprised by the bill. CZ is happy to give you more insight. With us you can easily view and compare a large number of hospital rates up to 885 euros. This way you know where you stand."

Figure 9: Count of provider specific prices published by CZ for each DTC



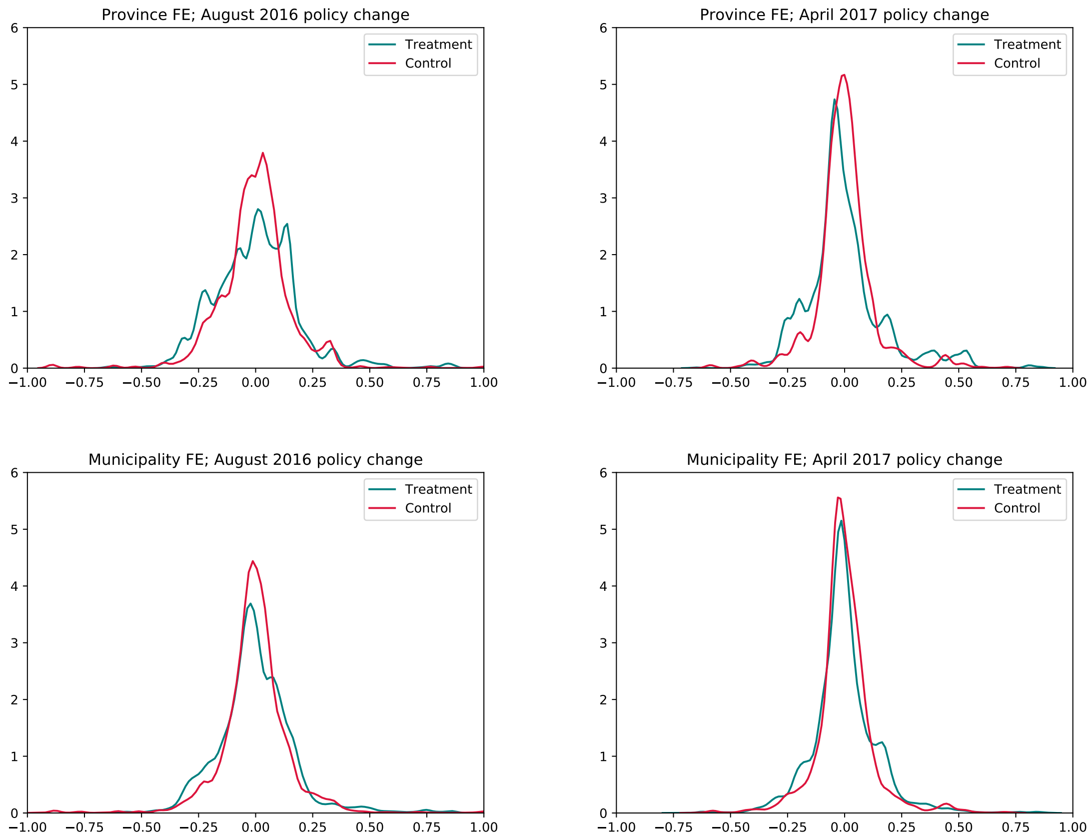
Note: Figures above present provider specific prices published by CZ and grouped by Diagnosis-Treatment combinations. Graph on the left presents the histogram of price counts for year 2016 whereas graph on the right presents a similar histogram for 2017. Shaded areas mark regions where the 6 selected dermatological procedures used in this study are located. One can clearly see that the selected procedures are also the products for which the most prices were published. Note that DTCs that have only one published price are excluded from the histogram for the clarity of the exposition.

Figure 10: Relationship between prices for similar DTCs



Note: The scatterplots above plot pairs of relative prices (computed as a percentage difference to the mean across hospitals) of selected pairs of the 6 DTCs that are used in the analysis against one another, like in Figure 2. See main text and notes to Figure 2 for additional details and discussion.

Figure 11: Log price residuals



Note: The figures above plot residuals from regressing log price on province, procedure, insurer and municipality fixed effects, separately for each event and using a subsample before the actual event. The validity of the empirical approach in this study hinges on the similarity between the treatment and control group. Here, we follow Lieber (2017) and provide evidence on the distribution of the residuals for the pre-periods. The upper row of Figure 11 presents residuals from a regression of the log price paid on province, insurer and procedure fixed effects. The bottom row controls in addition for municipality fixed effects (for those municipalities with at least 100 observations). The distributions are generally similar, although the dispersion is somewhat higher for the treatment group when we only control for province fixed effects. Based on this, we conduct a robustness check in which we also control for these municipality fixed effects. Results are not qualitatively different from the baseline results.

3 Indirect reciprocity stimulates viewership gifts in the video game streaming industry

3.1 Introduction

Prosocial behaviors within the society can often be explained with reciprocity: responding to friendly or harmful actions even if there are no material benefits from doing so (Fehr and Gachter, 2000). Trust and reciprocity can enable many economic interactions where enforcing an executable contract would be prohibitive (Seinen and Schram, 2006). While there exists ample experimental and field evidence of direct reciprocity such as multiple results on the Trust Game introduced by Berg et al. (1995) or field evidence in charity gift exchange by Falk (2007), less research has been carried on the topic of indirect reciprocity.

In contrast with direct reciprocity, which implies repeated interactions, indirect reciprocity can be characterized as a response to an act of kindness that is conveyed by a third party (Seinen and Schram, 2006; Alexander, 1987). For example, Seinen and Schram (2006) showed that the decision to help a stranger is largely motivated by his or her helpfulness score generated from previous interactions with the other participants of the experiment.

Such reciprocity can be defined as 'downstream reciprocity' (Mujcic and Leibbrandt, 2018) where individuals are more likely to be considerate to cooperative people, even if the history of their cooperation was of no consequence to them personally. In contrast, 'upstream reciprocity' describes a situation where people are more likely to be helpful if they received aid from a third party themselves (van Apeldoorn and Schram, 2016).

Several papers explored the topic of indirect reciprocity (see eg. Dufwenberg et al., 2001; or Engelmann and Fischbacher, 2009; for a review see Nowak and Sigmund, 2005), but to my best knowledge only a few tested it outside of the laboratory settings. Notably, van Apeldoorn and Schram (2016) investigated both downstream and upstream indirect reciprocity in a market where people can ask for and give services to each other free of charge. They discovered that individuals who offered such services in the past are far more likely to be accepted for receiving a service in the future. At the same time they found no evidence of upstream reciprocity and highlighted the possibility that reviews of past interactions can simply make an individual more trustworthy.

Upstream reciprocity was directly researched by Mujcic and Leibbrandt (2018) who found that probability to stop and give way doubled for drivers who were themselves given way by the experimenter shortly before. In particular, reputational concerns played no role in this setting indicating that indirect reciprocity is not only caused by other's reputation or an attempt to build it for oneself.

Kizilcec et al. (2018) looked at the effect of receiving a gift on the probability of gifting somebody in the future. Using a large data set of over 1.5mln gift exchanges on Facebook and individual birthdays as break points, they found that people who received birthday gifts were 56% more likely to give such gift in the future. However, a follow up survey revealed that many respondents expected reciprocity for the gifts.

This study uses an example of video game streaming market to study reciprocity in gift giving. Video game streaming is a relatively new and often professional activity where individuals broadcast a live video of their gameplay while interacting with their audience in real time. With 'streamers' usually having a camera feed next to the live video of their gameplay, and viewers being able to chat and cheer for them, this innovative activity can gather several thousands of viewers per streaming channel. Platforms such as Twitch.tv not only allow to broadcast one's gameplay, but also host esport tournaments or charity events run by their users (Deng et al., 2015). This new form of entertainment becomes increasingly popular, too. In the first quarter of 2021, three major streaming platforms accumulated over 9 billion hours watched, with approximately 290 million hours streamed by over 14 million unique channels.[16] Twitch alone has more than doubled in hours watched since the first quarter of 2020 (May, 2021).

It then comes as no surprise that video game streaming can also be a source of significant income for the streamers through donations, subscriptions, advertisement or selling merchandise (Stephenson, 2019). And this quickly adds up: the most popular Twitch streamer used to make at least $500k a month just from the 250 000 subscriptions (Herrman, 2018). Although such a target may be off the limits for an average streamer, some are able to engage in the activity professionally and full time.[17]

[16] Sjöblom and Hamari (2017) reported that individuals who took part in their survey on average followed 26.4 streamers and watched 11 hours of broadcast provided by 5.6 streamers weekly.

[17] The market is quite skewed in terms of popularity: using data on Twitch, Deng et al. (2015) show that majority of viewership is accumulated in a small fraction of both channels and games (top 10% channels gather 93% of the platform's viewership while top 10% of games accumulate 95% of that viewership; Deng et al., 2015).

In this study, I focus on 'hosting': a phenomenon unique to the video game streaming industry where one channel mirrors another channel's content while being offline. If performed at the end of one's streaming session, as it is most often done, hosting allows the audience to stay on the channel and enjoy the content of the selected 'hostee' or, with a single click, transfer directly to the mirrored channel.[18] Furthermore, any new audience arriving at the channel will be showed the content of the 'hostee'. Hence hosting can be viewed, and in practice often is, a viewership gift from a 'hoster' to a 'hostee'. Given that viewership is tightly related to the potential earnings, hosting can be viewed as a monetary gift. Importantly, this phenomenon is a prevalent practice on the platforms and often happens between strangers as a way to show support to the other content creators.

Motivation of streamers when hosting may be multidimensional. They can improve their reputation by supporting other channels and develop tit-for-tat relations with other content creators. Furthermore, they can also signal what channels are worthwhile watching, allowing them to advance their careers as streamers. In that sense, there is a degree of content curation involved (Dale, 2014), though it may also be strategic in nature. At the same time, there are potential costs involved such as exposing the viewership to a potential competitor, though it seems that overlap in streaming time is rather infrequent among streamers and their hostees.

The aim of this paper is to investigate the effects of indirect reciprocity in hosting and hostee choice. I exploit a unique high-frequency data assembled for this project to study whether average number of hosters and average viewership through hosting channels during the last streaming session affects the decision to host or not. Since channels often receive several hosts with varying viewership over the streaming session, the effect of being hosted should mostly be reflected in indirect reciprocity. Having received more such viewership gifts throughout the streaming session, both in absolute terms but also in terms of the viewer count associated with them, should positively influence the decision to host afterwards by generating a sense of gratitude towards the community. In contrast, receiving less gifts than on average may trigger negative reciprocity and motivate an individual to simply go offline without making a gift to anybody.

[18] A similar phenomenon called raiding allows to directly transfer viewers between channels. Hosting is described by Twitch, a major video game streaming platform, as basically embedding a video of another streamer on one's personal streaming page (Twitch, 2014).

This paper contributes to the literature on indirect reciprocity by providing a large sample analysis of a new market where indirect reciprocity motivates viewership gifts. It also differs from previous research on a few important dimensions. Firstly, viewership gifts are exchanged between streamers that potentially compete for market share. Hence the results of this study show that the effects of indirect reciprocity extend to outcomes within a competitive market. While prosocial or altruistic behaviors may be desired by the viewers, the modeling approach taken in this paper and availability of the interaction history between channels allows me to distinguish between reputational concerns and upstream indirect reciprocity.

I find that both these forces play a role: individuals are more likely to host if they were hosted in the recent past, and even more so if they are having many hosters at the moment of making the decision, though the latter may also be reciprocal in nature if by hosting one is 'carrying over' the hosts received. Furthermore, I estimate that the propensity to host depends on gift size, with larger gifts that carry more viewers having a stronger positive effect on the decision outcome. The results remain qualitatively similar even when excluding outcomes that could be a result of direct reciprocity to interactions within the last week.

Secondly, the nature of the hosting phenomenon allows for a free choice of the hostee.In contrast to the existing literature, I am able to investigate whether indirect reciprocity is also expressed through the choice of the receiver. I find that being hosted in the recent past, and in particular with a large viewership through hosting, results in choice of a hostee that is smaller in both current viewership and follower count. Since smaller channels are less likely to reciprocate a gift in its full weight and may be less appealing to one's audience, this may indicate that individuals feel reciprocal towards the community and express their gratitude by a potentially generous action of promoting smaller channels.

Finally, I use detailed data on user characteristics that proxy for popularity and experience on the platform to investigate heterogeneity in upstream indirect reciprocity. Since more experienced channels are likely to have a developed network of friends and cooperating channels, I find that indirect reciprocity is having a stronger effect on less experienced and less popular channels. This heterogeneity may be helpful in promoting cooperative behavior of hosting other channels among less experienced streamers. For instance, receiving gifts may inform them about the existing norm on the platform or allow them to exchange such gifts in the future to form friendships with the other users.

In addition, this paper describes a relatively new market of video game streaming industry and in particular the phenomenon of hosting that is unique to the market. Existing literature studied Twitch platform dynamics (Deng et al., 2015), motivation for the viewers watching video game streams (Sjöblom and Hamari, 2017), or effect of suspense on consumer utility using esports tournaments (Simonov, Ursu and Zheng, 2021), to name just a few. However, to my best knowledge, this is the first paper to study motivations behind hosting and, more generally, reciprocal relations between video game streamers in more detail. The results of this paper suggest that indirect reciprocity can, at least to some degree, explain the prevalence of gift giving among users on the platform.

This paper is structured as follow: Section 2 presents the data collection process, sample selection and descriptive statistics. Section 3 provides a reduced form model used and the results of estimating the effects of indirect reciprocity on both decision to hostee and characteristics of the chosen hostee. Section 4 provides a discussion of the results and contributions of this study.

3.2 Data

3.2.1 Data collection and hosting

The main data set was collected over a period of 48 days using a publicly available part of the API provided by one of the major streaming platforms. A collection instance, which consisted of a data request sent to the API approximately every 5 minutes, gathered information about all currently streaming channels and their characteristics such as number of viewers, game played, whether they are being featured on the main page or static channel popularity measures. With an average of roughly 20 thousand channels online at every collection instance and 288 such collection instances a day, this unbalanced panel of streaming channels constitutes over 250 million observations in total.

However, since this data gathering approach does not allow to track information on hosting activities, a similar request was sent for a selection of 1000 middle-sized and frequently streaming channels regardless of them being online or offline.[19] As a result, the core data set consists of a balanced panel of relatively more popular and

[19] This logic of gathering data is dictated mostly by the API restrictions of the service that allows for limited amount of queries within a given period of time, that limit being much lower for gathering

regularly streaming channels followed consistently over time, augmented by market-level information on all online channels and their activities at a given gathering instance.

Being hosted by other channels on the platform can be regarded as a substantial monetary gift for the hostee. It results in an increased viewership in the short run, exposure to hoster's viewership who may become followers or regular viewers of the stream in the longer run, but also higher position in the viewer-based ranking system that further increases exposure. Graph on the left of Figure 12 plots viewership of both hosting and hosted channels in periods relative to the hosting instance occurring always between 5th and 6th period. It can be seen that hosting channels continue to have an audience despite going offline while at the same time there is a large boost to the viewership of the hosted channel. Note however that the audience that stays on the hosting channel is still watching the content provided by the hostee, though with different chat and layout.[20]

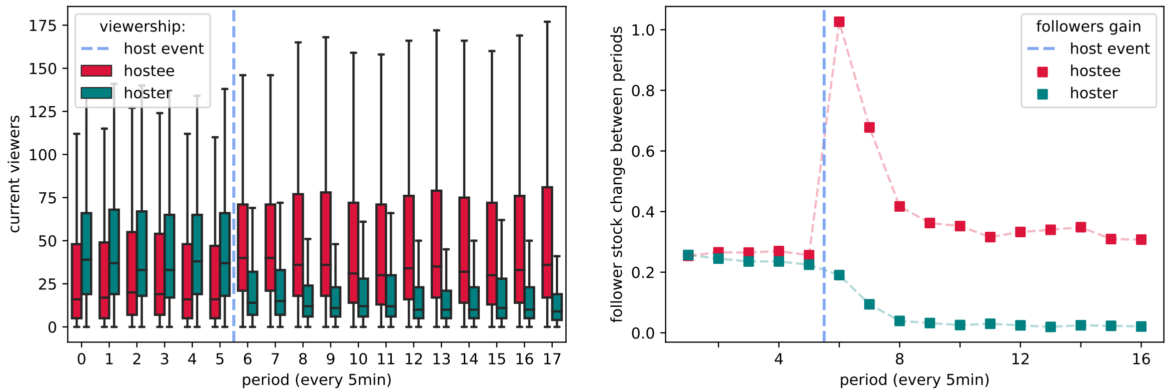
Graph on the left of Figure 15 in the Appendix modifies the viewership as a fraction of hoster viewership prior to hosting, allowing for relative comparisons of viewership flows between channels in the one hour period after. Looking at the median values, hosting channel seems to retain around 40-50% of its audience in the short run and hostee gains approximately 25-40% of that viewership directly moving to the stream, though with both amounts slowly decaying over time.

In the short-run hosting is clearly mutually beneficial, allowing to keep the hosting

information on particular channels as opposed to information on all currently live channels. The subset of channels was selected using an extra sample of 43 days of data gathered beforehand, aggregated to full days based on CET timezone and restricted to observations that were only in the 'games' type category. Sample channels were chosen based on the following criteria: streamed during at least 21 days, had between 10 and 2000 average daily viewers and the average daily hours streamed over days streaming was between 3 and 18 hours. In total, 2186 channels satisfied this requirement. A random draw of 1000 out of them was chosen as the subset of channels used in the study in order to comply with the maximum API request restrictions. The distribution of channels on the platform is highly skewed in terms of popularity and viewership, with a vast majority of channels having few or no viewers at all. In that sense, the selected sample is not representative for the channels existing on the platform, but rather for the channels watched on the platform. This, however, is fully intended in order to use channels that are treating streaming professionally or, at the very least, devoting a substantial amount of time to the activity such that the viewership and popularity on the platform carries a substantial weight for them.

[20] Note also that the hosting channels in this sample are relatively larger and hence their hosts often carry substantial viewership with them.

Figure 12: Short-run effects of hosting on current viewership



Note: graph on the left plots per-period viewership for 30 minutes before the hosting instance (relatively) and 60 minutes after the hosting instance, with periods in approximately 5 minute interval from each other. Graph on the right plots follower stock changes over the same periods of time for both the hostee and the hoster. There is a clear increase in short-run viewership following a host, but also a visible increase in average follower gain among hostees which indicates long-run gains for the channel.

channel busy and hostee's audience together while giving a viewership boost to the hostee. But the benefits are also long-run: graph on the right of Figure 12 shows that the hostee has a large increase in followers directly after the event and an increased follower gain rate afterwards, a soft proxy for long-run popularity. Clearly, hosting does not only result in short-run transfers of the viewership, but also substantial changes in the long-run stock features of the channels that measure their core viewership base. Since streamer earnings are intrinsically related to viewership - directly through the ad revenues, donations and subscriptions, but also indirectly by facilitating receiving the partner status that, among other things, requires at least 2000+ followers - hosting can be regarded as a significant monetary gift to another streamer.

At the same time, the cost-benefit ratio for the hoster is less clear: on one hand, hosting allows the channel to remain 'busy' and may be a sign of a tit-for-tat relation with other streamers. Furthermore, hosts may be announced on the receiving channel, providing exposure to the hoster, and this is especially true for hosts with larger viewership. Still, there are potential drawbacks, too: for instance, it exposes one's audience to potential competition or decreases the likelihood that viewers will watch records of the past streams of the hoster (Videos on Demand, or VODs) while the channel is offline. As seen in the later sections, almost one in five streaming instances ends without a host indicating that the decision to host or not may indeed carry trade-offs.

Regular streaming schedules are important for viewers to be able to return to your channel (Twitch, 2021). For channels in my sample, 40% of the total streaming time happens in the same 3 hours of the day over the sample period. Similar calculations for top 8 and top 12 hours of the day result in approximately 72% and 77% of the total streaming time. This indicates that channels in the sample stream quite regularly and are likely to have fixed schedules.

However, hosting a channel that have a similar streaming schedule poses somewhat greater threat of permanently losing viewers to the hostee than hosting a channel that streams at completely different hours. Plot on the left of Figure 16 in the Appendix shows mode hour of streaming for sampled channels. Plot on the right of figure 16 in the Appendix shows average difference in mode streaming hour between the channels in the sample and their hostees. Mean of this difference is equal to 6.8 hours. With average streaming session length equal to 5.6 hours (over 12 hours prior to the end-of-stream instance), this indicates that the overlap is quite infrequent, though the distribution of differences is moderately skewed to the left. Still, since viewers are likely to devote a fixed amount of leisure time watching, even when the hostee streams at completely different times, viewers may still switch their fixed allocation of time rather than watch more overall.

3.2.2 Sample selection and summary

While extremely rich, the main data set proves prohibitively large for the modeling purposes. In order to leverage the abundance of the data while keeping the estimation process tractable, I focus on hosting decision moments conditional on present characteristics of channel i and past interactions with other channels. Since I observe the platform at approximate 5-minute intervals, I cannot identify the exact moment of hosting or making the decision to host. Instead, I find instances where channel i was online and streaming at period $t-2$, $t-1$ and t, but is either offline at $t+1$, $t+2$ and $t+3$ or hosting the same j at three consecutive periods. As a consequence, I make an assumption that hostee choice occurs at period t while in reality it is more likely occurring in the 5-minute interval between t and $t+1$.[21]

[21] By selecting only the observations t where i was online in 3 consecutive periods before and is offline at least 3 periods afterwards, I make sure that temporal inactivity or potential missing observations due to imperfections in data gathering process are not recorded as end-of-streaming instances.

For each such identified hosting instance, I further gather important features of hosting channel i and hosted channel j at time t. These characteristics contain information on the viewership and number of followers of the channel at time t, but also the popularity of the game played or static characteristics such as the channel language or being partnered with the platform. In addition to that, I collect information on the history of channel i that may have influenced the hosting decision. Prior to the hosting instance, it is likely that i was having a streaming session that lasted longer than 15 minutes, but of varying length depending on the instance. To account for that, I look at last 12 hours before the hosting instance and collect channel characteristics such as the count of periods online, average viewership or followers gained. These variables are aggregated over the periods when channel i was online during last 12 hours to proxy for the last streaming session or events that streamer i considers as 'recent'. Note that while this measure is very crude and can, in some cases, merge two separate streaming sessions, I assume that most channels have regular streaming schedules that consist of a single, continuous session a day and that a streaming session is rarely longer than 12 hours. Descriptive evidence presented later suggests that this is likely the case.

I also make a similar aggregation as above, but over a larger window of last 7 days to account for interactions such as average weekly viewership, how often is i streaming throughout the week or how often i is generally hosted by others. This allows for capturing long run dynamics such as whether the channel has recently developed stronger hosting relations with other streamers, but also limits the sample to 41 days in order to allow for each end-of-stream instance to have the same history length available. Although here the selection of 7 days can be considered arbitrary, note that larger aggregations would limit sample even further whereas one week horizon should be sufficient to capture the dynamics that are relatively longer-run compared to last 12 hours used for last stream aggregations.

Finally, I make a limitation where an individual channel must end a stream at least 10 times within the sample period of 41 days to be included in the data set (this naturally implies that the channel was streaming at least 10 times for more than 15 minutes each time). This way I ensure that channels that became inactive or are not streaming regularly are not included in the data set. This is also done for practical reasons since I am able to consistently and efficiently account for individual fixed effects. As shown later, inclusion of fixed effects in the estimation process is crucial in order to account for habits, settings and hosting networks of channels in the sample that are very likely to be correlated with the variables indicating indirect reciprocity.

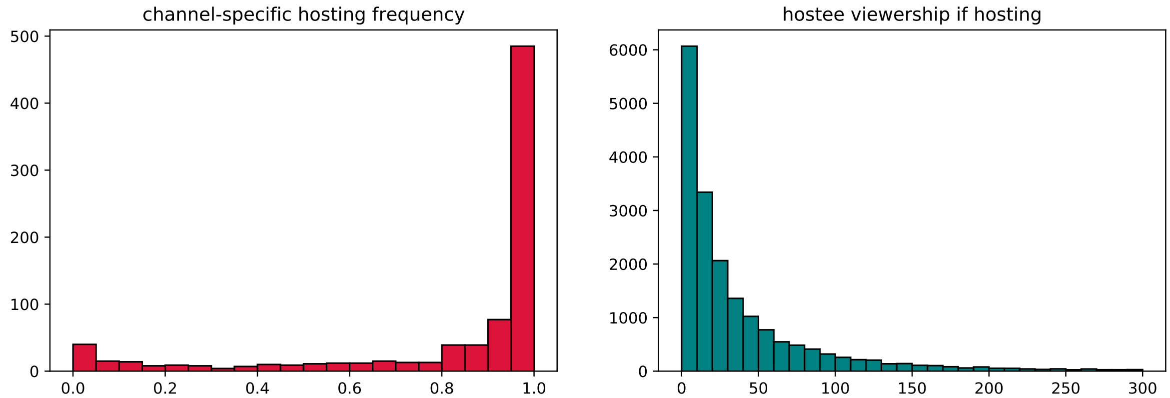
Summary statistics of the selected hosting instances and their features can be seen in Table 10 in the Appendix. I record 23867 such end-of-stream instances made by a sample of 840 channels over the period of 41 days. Strikingly, over 80% of streams end with a host, indicating that this phenomenon is more of a norm than a novelty on the platform. On average individuals streamed between 5-6 hours prior to the end-of-stream instance. Importantly, looking at the average viewership over last 7 days, viewership through hosting channels constitutes almost 30% of all the viewers watching the channel. Note that this is a lower bound on the contribution of hosting towards one's viewership since many viewers choose to switch between channels (see Figure 12 and Figure 15 in the Appendix), which further highlights how important hosting is for the popularity of the channels in the sample.

Finally, note that while the choice of 12 hours to account for last streaming session may have been considered as arbitrary, 90% of the observations in the sample had a stream session duration between 1.92 and 11.58 hours. This indicates that for a vast majority of individuals in the sample these artificially selected cut-off points are not binding and hence should have a negligible influence over the results. At the same time, on average channels in sample stream every 35.5h, and only little below 11% of the instances has started to stream again after less than 12 hours indicating that potential overlaps are quite infrequent.[22]

While extremely prevalent on the platform, decision to host is actually quite heterogeneous across channels. Graph on the left of Figure 13 plots the frequency of ending a stream with a host, averaged separately for each channel in the sample. Approximately 44% of the channels are always hosting - this may be out of a habit or even automatic since it is possible to set up auto-hosting that will select a hostee from a predefined list once a channel is offline - and little more than 4% never hosts. Still, a slight majority of channels seem to make a more ad-hoc choice of whether to host or not: 53.5% of the observations come from the channels with variation in the dependent variable.

[22] Though it has to be noted that in these cases last streaming instance aggregations will partially aggregate over the previous instance. This is an unavoidable trade-off between capturing fully the last streaming session and adding noise from other streaming sessions given heterogeneity among channels in streaming time and a choice of a single cutoff point. Furthermore, I include all the end-of-stream instances even if some of them partially overlap which may lead to situations where previous session among the overlapping pair is shorter. At the same time, I attempt to control for that with records of exact stream length.

Figure 13: Individual hosting frequency and distribution of hostee channel viewership



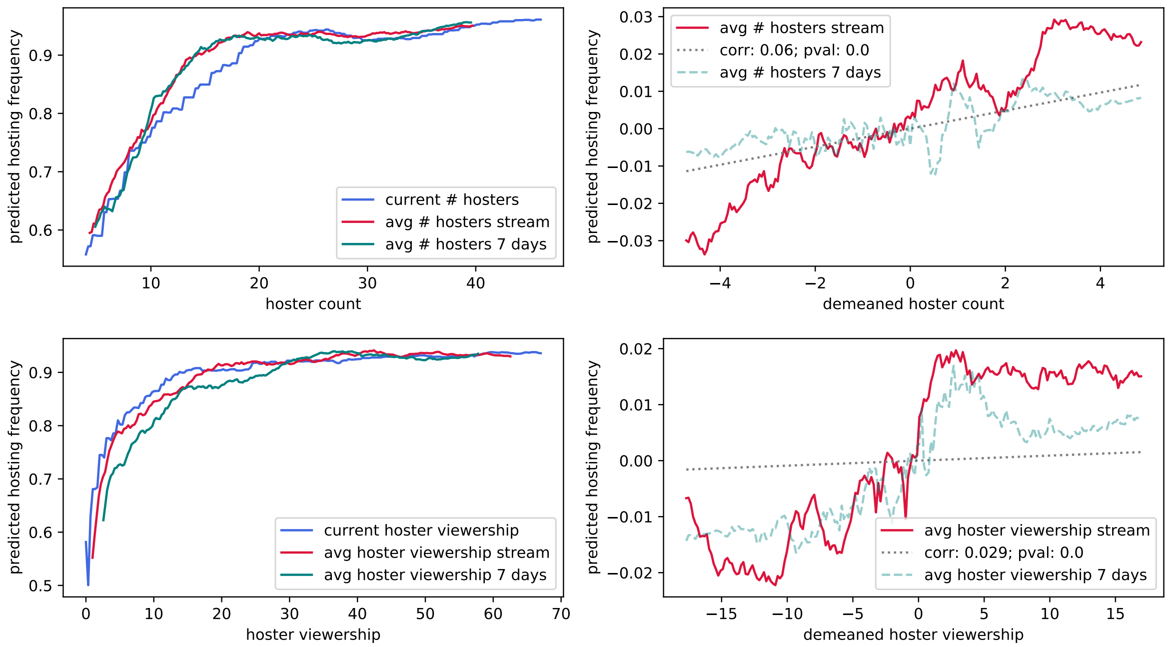
Note: the graph on the left plots the frequency of hosting after ending the stream grouped for each channel seperately. Only channels with at least 10 end-stream instances are used; this results in 23867 instances for 840 unique channels. Note that 380 out of these channels always host and 29 never host. Graph on the right plots histogram of hostee curren viewership at period prior to host conditional on hosting decision. The histogram x axis is trimmed at 300 for clarity of the display.

This indicates that while much of the choice process can be explained by the individual fixed effects, for many channels in the sample the decision is also driven by dynamic incentives and therefore can potentially be influenced by reciprocal considerations.

Graph on the right of Figure 13 plots a histogram of current viewership for the chosen hostees. Clearly, channels with less viewers are hosted much more often than channels with larger viewership base, resembling the general size distribution among channels on the platform. However, current viewership may not be fully reflective of actual popularity of the channel. Some channels may be just starting their streaming session and only in the process of gathering their regular audience. To address that concern, graph on the right of Figure 15 in the Appendix presents a similar histogram for follower stock of hostees at time t. In line with previous observations, a vast majority of hostees have very few followers compared to the median streamer in the sample. Combined with the fact that sampled channels are relatively more popular than the hostee channels (median hoster viewership is almost 63% higher than the median hostee viewership), descriptive evidence suggests that hosting may be a tool to promote less known or new streamers and gifting them with exposure to new audience.[23]

[23] Even if this is an unintentional effect of randomly choosing a hostee. However, with correlation between hostee and hoster followers of 0.098, significant at 1%, there is some indication that channels of similar size generally tend to choose each other as hostees.

Figure 14: Hosting frequency conditional on hoster count and viewership through hosters



Note: graph on the left plots predicted hosting probability conditional on different level of aggregations of hoster count and viewership through hosters using a K nearest neighbors algorithm with k=2000 and uniform kernel weights to smooth out the frequencies. Note that the difference in the variables is only in the level of aggregation: current, last stream average and last 7 days average. Graph on the right present the same KNN smoothed predictions over demeaned variables, ie. KNN model predictions on the right are orthogonalized against individual fixed effects. Grey dotted lines denote the linear predictions of stream aggregation against hosting decision, with label showing Pearson correlation and p-value.

Finally, in Figure 14 I investigate the relation of the variables of interest with the observed probability of hosting at the end of the stream. Note that frequencies plotted are averaged over nearest 2000 observations using KNN algorithm in order to smooth out the plot lines. Clearly, channels that that are hosted regularly are also more likely to host at the end of the stream. This difference in frequencies can increase from 0.5 up to over 0.9, raising sharply in the lower regions which may be indicative of the influence of developed hosting networks or lack of thereof between channels. However, based on the two graphs on the left it is difficult to determine whether more recent aggregations have a stronger effect on hosting decision and it seems that there are some underlying factors that drive most of the relation.

Graphs on the right side of Figure 14 attempt to remedy this problem by demeaning the data for each individual and again plotting the predicted frequencies using a KNN algorithm with a uniform kernel for smoothing. I focus now on aggregations over last streaming session that proxy for indirect reciprocity and only plot weekly aggregations for a comparison. Once demeaned, there is now a clear positive relation between deviations from average hoster count and viewership through hosters size and the frequency of hosting decision. These relations from demeaned data show a strong evidence of upstream indirect reciprocity where channels that are hosted by more streamers, and with larger viewership gifts, are more likely to host at the end of the streaming session. Importantly, the reverse is true for cases when a channel received lower than average amount of gifts recently, showing some signs of negative upstream indirect reciprocity.

Formal analysis shows a significant positive correlation of 0.06 between demeaned hosting frequency and recent hoster count as well as between hosting frequency and viewership through hosters where the coefficient is equal to 0.029 (both significant at 1%). And importantly, same correlations between demeaned hosting frequency and demeaned hostee count is equal to 0.01 and only significant at 10% whereas correlation with viewership through hostees over last week turns negative and equal to -0.01 (significant at 5%). This suggests that recent events have a lot more influence over the decision to host.

Overall, descriptive evidence suggests that hosting can be a substantial viewership gift from one streamer to another. It is also a very prevalent norm on the platform, with over 80% of streaming sessions ending with hosting. While much of the decision to host can be considered static or related to very long-run dynamics, it seems that current considerations also play a role in the decision making of individuals on the platform. In particular, I find evidence that slightly over majority of the channels on the platform have some variation in their decision to host or not and hence are potentially influenced by dynamic considerations. Once these fixed effects are controlled for, one can see a clear positive relationship between being hosted in the last 12 hours and frequency of hosting at the end of the session. In the next section, I introduce reduced form models that allow to capture these interactions in more detail.

3.3 Empirical results

3.3.1 Models

To estimate the effects of indirect reciprocity on decision to host and hostee choice of i, I use the data on how many other streamers were hosting i in the previous and current periods. Individuals who are hosted more should feel more grateful towards the community and may be more likely to express that gratitude by hosting somebody at the end of the streaming session. More importantly, I make an assumption that upstream reciprocity is susceptible to an imperfect recall, ie. having been hosted by more individuals during last stream has an effect on current decision, but same variable aggregated over last few streams should have less influence.

This assumption is implicitly used by field experiments, ie. Mujcic and Leibbrandt (2018) assume that recent experience of being given way has extra influence in current decision while Kizilcec et al. (2018) assume that receiving a birthday gift for recent birthday has an extra effect on gift giving in periods after. Usage of different time lengths was also suggested in van Apeldoorn and Schram (2016) as a mean to distinguish between the forces driving the decision.

In this paper I assume that reputation of a cooperative individual is mostly built at the moment of decision making when she or he is observed by viewers or hosters making the decision to host. Gratitude can be felt as a result of recent past such as being hosted by many people over the streaming session. Finally, hosting network and long-run dynamics between channels are likely to be stable over longer periods of time. This way, by using different aggregation levels over past periods I can distinguish between reputational concerns, indirect reciprocity and long-run dynamics that are not captured by individual fixed effects.

Having that observation in mind, consider a following model:

$$Y_{it} = 1\{\beta_1 X_{it} + \beta_2 \frac{1}{S} \sum_{s=1}^{S} X_{it-s} + \beta_3 \frac{1}{L} \sum_{l=1}^{L} X_{it-l} + \omega M_t + \alpha_i + \gamma_t + \epsilon_{it}\}, \tag{4}$$

where Y_{it} is the binary decision to host at the end of the stream, X_{it} are characteristics of hoster i at time t, α_i is hoster i fixed effect and γ_t are time fixed effects: hour of the day and day of the week dummy variables. To further account for time effects, I consider several market characteristics M_t at time t such as sum of all the viewers, online streamers on the platform or indicators of distribution: count of English live streamers and number of active costreams that may also influence hosting decisions. These variables should jointly account for the cyclical availability of channels to host, but also unexpected time shocks which should be correlated with the aggregate market indicators.

I distinguish three sets of dependent variables: current characteristics at the moment of making the hosting decision X_{it}, average characteristics from online periods over last 12 hours proxying for the last streaming session $\frac{1}{S}\sum_{s=1}^{L} X_{it-s}$, and average characteristics from online periods over last 7 days $\frac{1}{L}\sum_{l=1}^{L} X_{it-l}$. This way I can include present reputational concerns such as how many viewers are watching the decision to host, indirect reciprocity and gratitude over the short window S and average characteristics over a large full week window L that can include longer-run learning. In this fashion, I implicitly make a modeling assumption that different aggregation levels express different forces driving the decision of i.

While I mostly build my analysis on the number of hosters of i at time t, I also include average viewership through hosters. This is driven by the hypothesis that hosts are of unequal value, ie. hosts that contribute many viewers are much more important for i than those that, in an extreme case, carry on no viewers at all. This variable, however, is not a fully reliable measure of how many viewers are contributed by hoster j since viewers can easily switch channels, and in fact often do so. (see Figure 12). Still, it provides a rough measure that should be at least to some degree correlated with large host gifts instances and hence if indirect reciprocity plays a role, having received large viewership gifts should increase the probability of ending the stream with a host.

I include viewership that can be tightly linked to reputational concerns ie. how many viewers see i making the decision, and gain in followers during last stream that can be related to gratitude towards the community. These features can be related to hosting instances since large viewership gifts can result in increased viewership of the channel and increase in follower stock for i, but can also be a result of other events such as costreaming with a particular channel or having a charity streaming session. General viewership can also influence the probability to host since streamer who has many viewers is able to produce a relatively larger gift. I am controlling for that by including different aggregations of the viewership of the channel as well as market indicators - total number of viewers on the platform or how many viewers watch this particular game - that allow for relative comparison of such viewership.

Since the decision to host may depend on availability of potential hostees, market indicators such as number of channels, number of English speaking channels or number of channels playing the same game allow me to control for such availability of live streamers. Still, availability of hostees can be negatively correlated with number of hosts received in the previous session as these gifts are likely to be given by familiar channels who are potential hostees themselves. Although I cannot control for entire history of prior hosting, as a robustness I consider a subsample of choices where i hosts j and j have not hosted i in 7 days prior to that, which, at least to a certain degree, limits the cases when hosting is motivated by availability of individuals who i can reciprocate directly.

By design hosting is highly dependent on channel settings such as auto-hosting, streamer habits and social networks of i, these characteristics likely to be relatively stable over the sample period. These features seems to be the major drivers of the hosting decisions given descriptive evidence in the previous section, and are likely to be strongly correlated with the variables of interest. To account for the unobservables, I include fixed effects α_i and initially model equation 4 using Linear Probability Model with dummy variable fixed effects. This approach allows for an easy inclusion of fixed effects without running into incidental parameter problem and presents clear initial interpretation of the results. I also estimate equation 4 using Conditional Logistic Regression as in Chamberlain (1980) to allow for non-linear relation between variables of interest and the probability of hosting.

Note that inclusion of the fixed effects should also largely account for the long-run weekly aggregations of the variables. In that sense, I include these aggregations to show that they are absorbed by fixed effects of individual channels. Consequently, while the choice of week as a level of aggregation is relatively arbitrary, if they are absorbed into fixed effects and largely insignificant in predicting the probability of hosting, it is even more likely to be the case for any aggregations over longer time horizon. Note also that I am using averages and not sums of the variables of interest over the periods streaming. Since number of periods streaming is a choice made by an individual, cumulative gains from being hosted can be treated by i as partially depending on the choice of stream length and hence elicit less reciprocity. In contrast, averages are more reflective of deviations in gift rate received by i and should be more correlated with the propensity to host at the end of stream if indirect reciprocity plays a role.

Given the benefits of hosting, it is possible that streamers subscribe or donate to each other, or even pay each other outside of the platform in exchange for hosts, and this kind of interactions are not recorded in the available dataset. In fact, such an exchange would be misleadingly recorded in the dataset as a one-sided gift. While this is a potentially important aspect of the interactions on the platform, it should not affect the estimation results since monetary payments should not be correlated with average number of hosts received and their size conditional on the individual fixed effects. In other words, there should be no correlation between gifts received through hosting and potential monetary payments once the popularity of that individual (which likely results in higher number of hosts received as well as higher value of hosting gifts) is controlled for. Hence this issue should not be a concern given modelling approach in this paper that includes several recent popularity measures as well as individual fixed effects.

Since indirect reciprocity can be expressed not only through the decision to host, but also the choice of the hostee, I additionally formulate a simple reduced form model:

$$log(Y_{it}+1) = \delta_1 X_{it} + \delta_2 \frac{1}{S}\sum_{s=1}^{S} X_{it-s} + \delta_3 \frac{1}{L}\sum_{l=1}^{L} X_{it-l} + \psi M_t + \alpha_i + \gamma_t + e_{it}. \quad (5)$$

This model is similar to Model 4, but the binary dependent variable is replaced with current viewership of the hostee in the period prior to host. Since the data is highly skewed towards left, I replace viewership with $log(Y_{it}+1)$ (see Figure 17 in the Appendix for the distribution of the transformed variable). As in the Model 4, I include current, short term and long term aggregated variables as well as market indicators M_t, fixed effects for hosters α_i and time fixed effects γ_t. Note that, similar as in Model 4, individual fixed effects are essential here as they should account for hosting networks and habits of i. Similarly, market indicators and time fixed effects should control for general availability of popular channels or stock of the viewers on the entire platform.

Current viewership may not be reflective of actual popularity of the hostee since chosen channel j may eg. be just starting to stream and hence have low viewership despite being quite popular. As a robustness check I estimate the same model with log follower stock of the hostee. Since these two measures are highly related and both are markers of channel popularity, they should be providing similar coefficient estimates on variables related to indirect reciprocity. At the same time, current viewership is what channel i is more likely to take into consideration when choosing j to host whereas follower stock provides a more stable measure, both in terms of popularity as well as the gathering method.[24]

[24] In some extreme cases, due to the data gathering imperfections, viewership can be recorded several seconds post hosting instance. This is one of the reason for using stable follower stock as a robustness check.

Table 7: Decision to host and viewership of the chosen hostee

	FE LPM		Conditional Logit	
	estimate	std err	estimate	std err
avg # hosters stream	0.092*	(0.0479)	2.7965***	(0.8888)
avg hoster viewership stream	0.0101**	(0.0039)	0.4823**	(0.2269)
Fixed effects		Yes		Yes
Observations		23867		12766
R squared		0.627		-

Note: the table above presents partial results for modelling probability of hosting at the end of the stream. Column 'FE LPM' presents results for fixed effects linear probability model with regular standard errors. Column 'Conditional Logit' presents results for conditional fixed effects logistic regression as in Chamberlain (1980). Note that both dependant variables are expressed in 100 of units for computational reasons (avoiding overflows in Conditional Logit MLE estimation).
Standard errors in LPM model are MacKinnon and White (1985) heteroscedasticity robust standard errors. See Table 11 in the Appendix for full results.

3.3.2 Results

The results of regressing the probability of hosting can be found in the first two columns of Table 7. Each extra hoster averaged over the last streaming session (row 'avg # hosters stream') results in approximately 0.0009 increase in the probability of hosting post-stream in the Linear Probability model or 0.0043 increase in the probability of hosting for an average individual with mean sample probability of hosting in the Conditional Logit model. Furthermore, each extra viewer through hosting channels averaged over the last streaming session (row 'avg hoster viewership stream') increases the probability of hosting by 0.0001 in the linear model or by 0.0007 for an average individual with mean sample probability of hosting in the conditional logit model (both significant at 5%).

Back of the envelope calculations using LPM model reveal that one standard deviation increase in average number of hosters during stream results in approximately 2.26% increase in probability of hosting whereas one standard deviation increase in average number of viewership through hosters results in approximately 0.85% increase in probability of hosting (these are 10.5% and 5.9% for the Conditional Logit when using marginal effects calculated for an individual with the sample average probability of hosting). These results not only support the hypothesis that indirect reciprocity plays an important role in the decision to host since receiving more hosts in recent past positively influences individuals to host others, but also that it is positively related to its intensity expressed by the average viewership through hosters throughout the last streaming session.[25]

Each extra hoster at the moment of decision making increases the probability of hosting by 0.0004 in the linear model and by 0.0027 in the conditional logit (see full results, including control variables, in Table 11 in the Appendix). This could indicate reputational concerns. However, neither the size of the audience through hosters nor channel's direct current viewership had a significant influence over the probability of hosting. This may indicate that perhaps it is not reputational concerns driving the decision, but rather direct reciprocity towards other hosters to carry their hosts, ie. if i is hosted by j and now i hosts k, j will also host k through chain hosting. Positive coefficient on that estimate combined with no influence of direct channel audience could indicate that individuals simply feel obliged to extend further the hosting chain, especially if it carries some viewership with it.

Importantly, weekly aggregations of the very same variables are largely insignificant across the models. This is an important indication that fixed effects seem to account for longer-run dynamics and networks between channels. Combined with significance on recent stream coefficient estimates, this gives some additional support for the hypothesis that the underlying effect driving the results is indeed upstream indirect reciprocity.

The results for regressing the log hostee viewership are presented in Table 8. Note that the interpretation here is of choice rather than affecting hostee viewership since viewership is recorded prior to the hosting instance. Increasing average hoster number over the last stream results in lower chosen hostee current viewership by 0.43%, a result significant at 10%. Furthermore, each extra viewer through hosting channels decreases the chosen hostee current viewership by 0.13%, significant at 1%.

Since a gift of 100 viewers is worth a lot more for a newcomer compared to a popular streamer, these results may indicate that indirect reciprocity is not only expressed by performing kind actions, but also choosing receivers who benefit relatively more from them. Importantly, these channels are less likely to reciprocate the gift in its full size: if they host back, they most likely carry a lot less viewership with such a gift. On the other hand, they may be more likely to directly reciprocate the gift since it was worth much more to them.

[25]Although qualitatively the same, conditional logit predicts several magnitudes larger marginal effects than the linear model. This may be caused by non-linear predictions of the former, especially in the regions close to 0 and 1 and given the large mean value of the dependent variable while calculating the marginal effects using mean sample probability of hosting. However, the exact reason is not clear.

Table 8: Decision to host and viewership of the chosen hostee

	log(viewership)		log(followers)	
	estimate	std err	estimate	std err
avg # hosters stream	-0.0043*	(0.0023)	-0.0059*	(0.0031)
avg hoster viewership stream	-0.0013***	(0.0005)	-0.0013*	(0.0006)
Fixed effects		Yes		Yes
Observations		18623		18623
R squared		0.298		0.308

Note: the table above presents partial results for modelling log viewership and log followers of the chosen hostee conditional on hosting. Column 'log(viewership) presents results for log-linear fixed effects least square regression against current viewership of the chosen hostee. Column 'log(followers)' presents results from similar regression where current viewership is replaced by current follower stock of the chosen hostee.
See Table 12 in the Appendix for full results.

Interpretation is robust to changing dependent variable from log current viewers to log follower stock (see the full Table 13 in the Appendix). Each extra hoster averaged over the last streaming session decreases the follower stock of the chosen alternative by 0.59% while each extra viewer through hosters averaged over last streaming session decreases the follower stock of the chosen alternative by 0.13%. Both estimates are significant at 10%.

3.3.3 Heterogeneity

Given the prevalence of hosting on the platform, I consider heterogeneity in effects of indirect reciprocity related to the popularity of the streamers. I expect that more experienced and established channels are less affected by indirect reciprocity since their decision to host is largely an effect of networks and friendships. Table 9 shows estimation results for interacting hoster count and viewership through hosters averaged over last streaming session with being partnered, follower stock and years since registering on the platform. While all these variables are intrinsically related to experience using the platform, follower stock and being partnered are also measures of popularity and professionalism of the streamer.

As expected, the results suggest that more popular streamers are less likely to be influenced by indirect reciprocity. For instance, each extra hoster on average during last

Table 9: Heterogeneity in indirect reciprocity

	is partnered		# years exists for		current followers	
	estimate	std err	estimate	std err	estimate	std err
avg # hosters stream	0.3305***	(0.0637)	0.4155***	(0.1132)	0.221***	(0.0525)
interaction: # hosters	-0.3099***	(0.0632)	-0.1231***	(0.0406)	-25.1402***	(7.7113)
avg hoster viewership stream	0.0084	(0.0054)	0.0607***	(0.0225)	0.0059	(0.0047)
interaction: hoster viewers	0.0013	(0.0089)	-0.0166**	(0.0074)	1.6257	(2.3857)
individual fixed effects		Yes		Yes		Yes
time fixed effects		Yes		Yes		Yes
nobs		23867		23867		23867
rsquared		0.6273		0.6272		0.6273

Note: the table above presents partial results for Linear Probability Model with additional interaction terms to show heterogeneity in response to being hosted. The models estimated above are similar to the baseline Linear Probability model presented in the modelling part and use a similar set of control variables. Interaction terms are added to highlight that the effect is different for channels depending on their popularity and relative experience with the platform. See Table 15 in the Appendix for full results.

streaming session results in 0.0033 increase in probability of hosting for non-partners, but this effect is negligible for partners. Similar, the effect seems to be fully offset when streamer is registered on platform for more than 3 years or has at least 100 thousand followers.

It is indeed possible that more experienced and popular streamers have a developed network of hosters and their decision to host is driven by long-run commitments. In contrast, less experienced individuals who are still learning the platform or are not so popular may be more driven by recent acts of kindness. Alternatively, it may be that individuals respond to relative rather than absolute changes, ie. a streamer that has on average 20 hosters is going to have a weaker response to an extra host than a streamer that has on average 5 hosters, and more popular streamers are likely to have more hosters generally.

Looking at the wider picture, in both cases indirect reciprocity can be seen as stimulating hosting and cooperation relatively more among newer streamers. This is an interesting result that may indicate that indirect reciprocity promotes norms of cooperation on the platform - in this case, the idea of hosting. With over 80% of streaming instances ending with the host in the sample under consideration, this is not an unlikely hypothesis.

3.3.4 Robustness

One of the concerns regarding the results could be that kindness that I observe among streamers is in fact manifestation of direct reciprocity. This in principle should not be an issue: dependent variables used denote aggregate hoster and viewership counts and channels are often hosted by several individuals (median of 13 hosters on average) over the streaming session. However, closer inspection of the data reveals that 36% of chosen hostees j have been hosting i in last 7 days.

To remedy this, results in Table 13 in the Appendix estimate the key models from Table 7 over a subsample of observations where i is hosting j at t and in the last 7 days prior to t channel j did not host i. Such subsampling allows me to ensure that there is no recent history of j hosting i and so no indication of (recent) direct reciprocity.

Surprisingly, the results on the reduced sample are even more indicative of indirect reciprocity in choice to host. While most coefficient estimates retain their significance level, majority of them increase in magnitude, indicating even stronger effects of indirect reciprocity on the decision to host after the streaming session. This suggests that indirect reciprocity is far more influential for individuals without strong hosting networks. At the same time, results for log viewership and follower stock in most cases lose significance. This may be either due to a drop in sample size or because such choice of smaller channels must be first triggered by some action on the j's side that perhaps makes i aware of j.

Throughout the study, I used averages of the hoster count and associated viewership aggregated over recent time period as the relevant dependent variables. This was dictated by the intuition that total count of hosters and viewership depends on stream length which is chosen by the individual and hence may generate less gratitude. One may still wonder whether using cumulative sums instead of the averages would be a more reasonable as a proxy for gift size as perceived by the streamers. I explore this possibility and interact hoster count and viewership through hosters with the streaming session length.

The results can be found in last column of Table 15 in the Appendix. Only the interaction with hoster count is having a significant, negative effect, indicating that the same average hoster count over a shorter period is having a more positive effect on the probability of hosting post-stream. Importantly, since averages of the key variables are both positive and strongly significant predictors of the decision to host, they indeed seem to be a better choice for variables proxying for indirect reciprocity in this setting.

One of the concerns regarding the results is that the true underlying model may contain many non-linear interactions between controls and the variables of interest. For instance, length of last stream and average viewership are likely to be having a joint effect. Similarly, effect of some variables may be different depending on the market conditions such as number of streamers online or total viewer count. To remedy that, I use double machine learning approach introduced by Chernozhukov et al. (2018) and allow for a fully non-parametric and flexible estimation of the nuisance parameters. Perhaps most importantly, it allows for non-linear control of recent streaming time and its interactions with other variables which may be important given that averages are computed over fixed choice of 12 hours. Although I do not include fixed effects, I expect that the non-linear estimator can distinguish between channels using a combination of followers, unique viewers and years on platform features. Still, since this method does not explicitly control for fixed effects and depends on the tuning of the algorithm, I include the results as a robustness check.[26]

The results for the double machine learning estimation routine can be found in the last column of Table 11 in the Appendix. While the estimate for hoster count loses its significance, viewership through hosters is significant now at 1%. The effect is only a half of what the linear model would predict: each extra viewer through hosting channels averaged over the last streaming session increases the probability of hosting by 0.00005. Such results may suggest that once non-linearities are controlled for, only gifts that carry value are actually triggering indirect reciprocity. In other words, it is not the hoster count which in many cases may not carry any viewers at all, but the viewership associated with it that generates gratitude. However, these results are not exactly comparable since fixed effects are not included in the estimation.

I also estimate the model using double machine learning estimator for log viewership tuned in the same fashion as in last column of Table 7. Again, while the estimates for hoster count lose their significance, viewership through hosters is still significant at 5%.

[26] Double Machine learning method (Chernozhukov et al., 2018) is using Random Forest regressor to orthogonalize againt the nuisance parameters. Random Forest is fitted with 200 trees, using 5 random features at each splitting point. The main estimator is using 30 cross-fitting runs, with 5 folds each. Note that for each variable that enters the model in a linear way, it is orthogonalized against all other variables with nuisance estimator, including the remaining variables of interest. Estimates and standard errors across cross-fitting runs are aggregated using medians, and estimated with DML2 procedure. See Chernozhukov et al. (2018) for additional details of the estimator.

Each extra viewer through hosting results in choice of a hostee with 0.12% lower current viewership. Such results suggest that even when allowing for fully non-linear control variables, having obtained larger viewership gifts results in choosing smaller channels as receivers of the host. Again, since fixed effects are not included, results should be interpreted with caution.

3.4 Discussion

This project investigates the effects of indirect reciprocity on viewership gift giving in the video game streaming industry where hosting is stimulating cooperation among, in many cases, absolute strangers. Average number of hosters and average viewership through hosters during last streaming session - the average amount of viewership gifts and average size of the gifts - are both having positive effect on the probability of hosting and gifting somebody else with viewership after one's stream. One standard deviation increase in either of this variable increases the probability of hosting by 2.26% and 0.85% respectively, indicating that indirect reciprocity plays an important role in viewership gift giving on the platform.

Streamers are indirectly reciprocal when receiving more gifts, but even more so if the gift received actually carries value. Indeed, many hosts may contribute only few viewers and therefore be of little value for the channel, and such 'empty gifts' may trigger little reciprocal action. This may partially explain why when allowing for non-linear controls through double machine learning estimator, only the gift size seems to matter in the decision to host. The estimated effects are also heterogeneous, with more experienced and popular streamers being less affected. Conversely, indirect reciprocity may be much more influential in establishing the norm among newer streamers on the platform.

Perhaps even more interestingly, the results of this study suggest that being hosted and being hosted with more viewers are influencing the choice of a hostee. This is an additional novelty of this paper that stems from free choice of a hostee among streamers and provides additional dimension to the effect. Using reduced form modeling, I estimate a negative link between being hosted and current viewership of the hostee. The parameter estimates indicate that streamers who were hosted more and with more viewers in the last session are both more likely to express their gratitude by hosting somebody else as well as they are more likely to choose and promote smaller channels as hostees.

This result is further robust when replacing current viewership of the hostee with the current follower stock, a much more stable measure of the channel popularity. It indicates that indirect reciprocity can improve discovery of new or niche channels and promote smaller players who otherwise would have problems 'making it big'. Furthermore, these observations seem to go in line with the conclusions of Mujcic and Leibbrandt (2018) who noted that much of the unconditional generosity observed can actually be explained with indirect reciprocity. These results, however, are largely not robust to excluding choices influenced by (recent) direct reciprocity.

It has to be highlighted at this point that the rather small absolute magnitude of the estimated effects should not be surprising given the sample choice. As noted, selected channels consists of mostly professional streamers who broadcast regularly for a large audience. One can then expect that they are mostly driven by strategic considerations that may involve direct reciprocity with other streamers such as exchanging hosts. This hypothesis finds some support in the results since the effect of direct reciprocity is much smaller for more experienced and popular channels. Still, the estimates are highly significant and largely robust across modelling approaches indicating that even in such professional settings indirect reciprocity can play a role in the decision to help others.

Confidently identifying an effect as caused by upstream indirect reciprocity is naturally difficult, with the concept itself being argued to be harder to explain than downstream indirect reciprocity (Nowak and Sigmund, 2005). In case of this study and the streaming market, individuals that exchange hosts may be motivated by how their appear to their viewers or simply be cooperative by nature and set up auto-hosting, these effects potentially being correlated with how many gifts they receive over the last streaming session due to direct reciprocity from former hostees. The modelling approach taken in this paper controls for static individual differences as well as reputation or long-run habits. In robustness analysis I exclude observations that could be considered as recent direct reciprocity, and stable hosting networks should be controlled through fixed effects over the sample period. However, it may well be that receiving many hosts simply makes an individual aware of the norm (or reminds of it) and hosting afterwards is part of simple strategic behavior in a similar fashion as discussed by Mujcic and Leibbrandt (2018). In that sense, although the results of this paper suggest that upstream indirect reciprocity plays a role in decision to host, one cannot fully exclude other potential explanations.

While this project specifically investigates upstream indirect reciprocity by quantifying the effect of gifts received on the propensity to give a gift afterwards, it is possible that streamers obtain gifts because of their cooperative behavior in the past. This would be an example of downstream indirect reciprocity (van Apeldoorn and Schram, 2016) where A is kind to B because B hosted C few days ago. However, it is difficult to determine what is the correct time horizon for downstream reciprocity in case of hosting, especially that it is often mentioned next to 'reputation' or 'social status' (Seinen and Schram, 2006; Alexander, 1987; whereas upstream reciprocity is a response to recent events; Nowak and Sigmund, 2005), both arguably built over a prolonged period of time. While there are some indicators that downstream indirect reciprocity may play a role - correlation between demeaned number of hosts received over the last streaming session and demeaned decision to host from previous period is positive and significant at 1%, same relation between demeaned current gift size and demeaned previous decision to host is negative and insignificant (see Figure 18 in the Appendix for visualization of these relations in a similar fashion as in Figure 14). This may be caused by controlling for individual fixed effects through demeaning that may largely account for reputation of the individual within the sample period while relating being hosted to decision from only the last session.

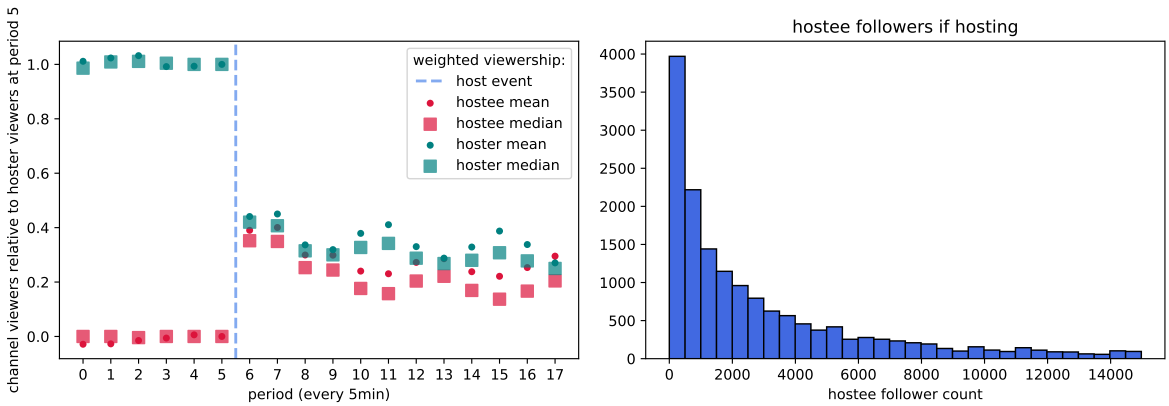
Instead, I argue that even in such case the results still support the existence of indirect reciprocity among individuals in the market. Indeed, the effects of both downstream and upstream indirect reciprocity may be chained together, and while I control for fixed effects and long run dynamics, further research could focus on downstream indirect reciprocity and attempt to quantify its effect, especially in contrast to the results obtained in this paper.

The development of the streaming platform in this study is the best example of how prosocial behaviors can become prevalent norm in the market even when market's nature is competitive. Furthermore, introducing users to a new feature such as hosting can create a more welcoming and supportive community with gains for the platform itself. One can imagine that if inertia plays a role in choice of channels and activities, such bottom up recommendation system should increase the overall viewership on the platform. Future studies could investigate in more detail if this is indeed the case.

In this study, using video game streaming market, I found robust evidence for upstream indirect reciprocity among streamers. At the same time, the following study shows evidence that the widespread occurrence of the phenomenon of hosting can be, at least to some degree, explained by indirect reciprocity.

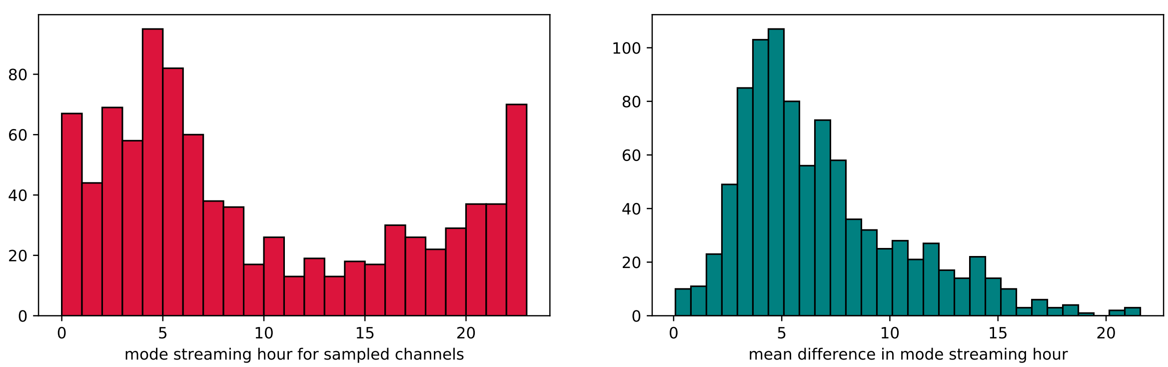
3.5 Appendix

Figure 15: Weighted viewership flows and follower stock of the chosen hostees



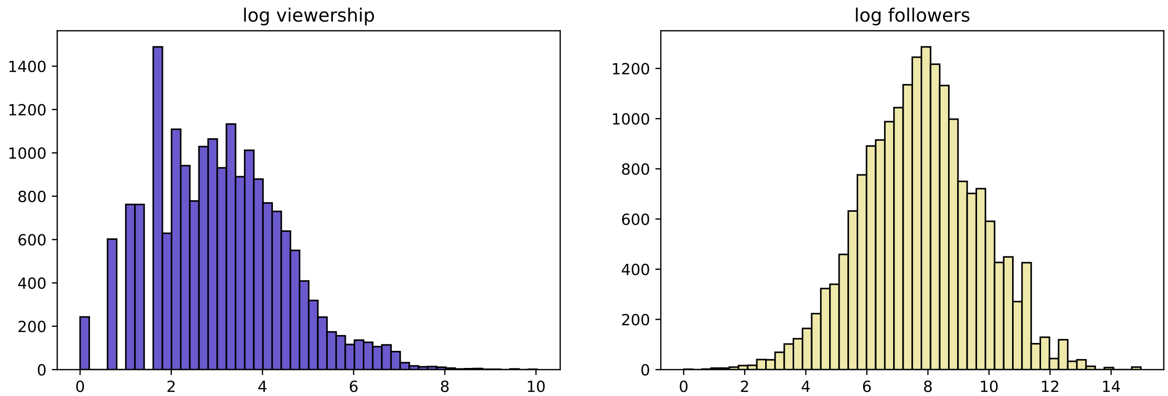
Note: Graph on the left presents weighted flows of viewership where viewership of the hoster is divided by viewership of the hoster at period 5 and viewership of the hostee is demeaned and divided by viewership of the hoster at period 5. In that sense, the Graph on the left plots fraction of pre-host viewership captured by hoster and hostee up to one hours after the hosting instance. Graph on the right plots a histogram of selected hostee follower stock. Clearly, most hostees are actually having few followers compared to median streamer in the sample, to some degree reflecting the popularity distribution on the platform.

Figure 16: Mode streaming hours and differences with hostees



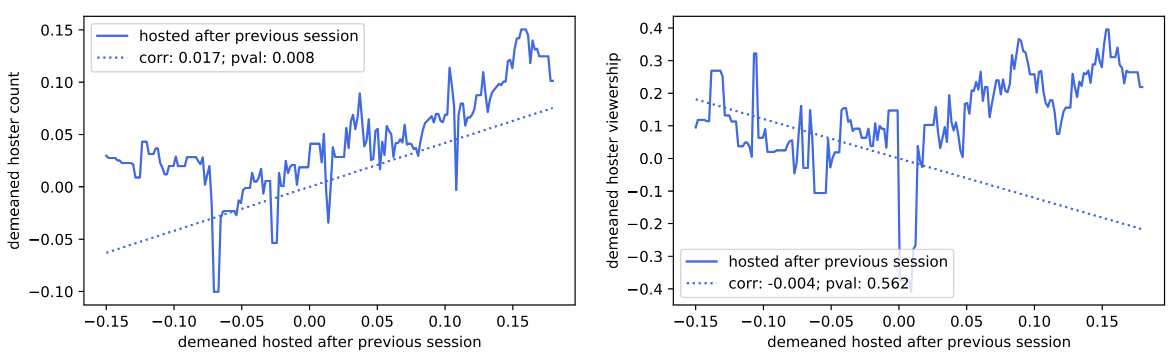
Note: plot on the left presents a histogram of mode streaming hour for the selected channels in the sample. Plot on the right presents a histogram of mean differences between mode streaming time of the channels in the sample and their hostees.

Figure 17: Log transformed dependent variables



Note: graphs above plot log(hostee viewership+1) and log(hostee followers+1) for observations where decision to host was made.

Figure 18: Relation between receiving hosts and hosting in the previous session



Note: plot on the left relates demeaned average count of hosters over the last streaming session given demeaned indicator for whether the channel has hosted at the end of the session prior to that. Plot on the right presents similar relation for the demeaned average viewership through hosters. See additional explanations of how line plots are computed in Figure 14 and related text.

Table 10: Summary statistics sample of the hosting instances

	Mean	Std	Min	Q05	Median	Q95	Max
will host	0.81	0.39	0.00	0.00	1.00	1.00	1.00
hostee viewership	75.74	339.25	0.00	2.00	20.00	271.35	22686.00
hostee followers	15233.00	85696.15	0.00	103.00	2381.50	62111.75	3189017.00
current # hosters	22.62	28.60	0.00	2.00	16.00	64.00	930.00
avg # hosters stream	19.71	24.61	0.00	2.58	13.35	54.91	475.68
avg # hosters 7 days	19.84	23.78	0.00	3.00	13.59	55.25	367.49
current hoster viewership	26.43	51.33	0.00	0.00	10.00	103.00	1919.00
avg hoster viewership stream	26.29	84.38	0.00	0.11	13.22	90.39	11669.14
avg hoster viewership 7 days	27.14	68.09	0.00	0.97	16.22	81.47	3393.86
current viewership	66.86	150.79	0.00	5.00	29.00	222.00	6561.00
avg viewership stream	62.94	167.82	0.00	6.68	28.23	191.78	12283.49
avg viewership 7 days	64.36	142.22	0.00	9.47	29.90	201.69	3815.08
current followers	11790.78	24058.22	185.00	626.00	3882.00	48815.80	240393.00
followers gain stream	18.32	44.00	-8.00	0.00	7.00	69.00	1389.00
followers gain 7 days	128.97	275.20	-37.00	5.00	52.00	510.00	5713.00
# years exists for	2.38	0.90	0.18	0.70	2.47	3.68	5.33
is partnered	0.35	0.48	0.00	0.00	0.00	1.00	1.00
english channel	0.94	0.24	0.00	0.00	1.00	1.00	1.00
autoplay VODs	0.38	0.49	0.00	0.00	0.00	1.00	1.00
is featured	0.01	0.10	0.00	0.00	0.00	0.00	1.00
is costreaming	0.07	0.26	0.00	0.00	0.00	1.00	1.00
streamer level	111.27	9.46	71.00	95.00	112.00	124.00	133.00
streamer experience (in 1000)	1292.90	1051.94	25.58	181.02	983.79	3315.32	7767.94
all-time viewers (in 1000)	292.56	1809.08	1.14	7.87	49.52	797.03	34978.56
stream length hours	5.65	2.77	0.25	1.92	5.17	11.58	12.08
7 days hours streamed	35.06	17.44	0.25	12.50	32.25	67.67	167.83
game viewership	3131.94	3695.09	0.00	29.00	1410.00	9258.70	123966.00
game streamers	1134.35	2082.75	0.00	2.00	318.00	6182.10	79229.00
all current viewers	58331.61	13818.65	13399.00	36112.30	60193.00	79149.00	149719.00
all live channels	18984.96	6027.34	7230.00	8930.00	20860.00	26599.00	50000.00
all current costreamers	4956.60	1770.68	1602.00	2107.00	5371.00	7311.00	19173.00
all live english channels	16679.96	5444.28	6021.00	7492.00	18411.00	23536.00	44119.00

Note: descriptives statistics are summarizes over 23867 hosting instances made by 840 channels over the period of 49 days. Note that the summary statistics hostee viewership and followers are aggregated only over observations when channel decided to host after the stream. 56 observations where sum live channels is lower than 2500 were dropped since these are most likely miscollected. Q05 and Q95 are 5th and 95th quantiles of the distribution, respectively.

Table 11: Full table for the hosting decision results

	FE LPM		Conditional Logit		Double ML	
	estimate	std err	estimate	std err	estimate	std err
current # hosters	0.0379*	(0.0207)	1.7249***	(0.4017)	-	-
avg # hosters stream	0.092*	(0.0479)	2.7965***	(0.8888)	-0.0044	(0.0446)
avg # hosters 7 days	-0.0271	(0.0596)	-0.4737	(1.3253)	-	-
current hoster viewership	0.0034	(0.0034)	0.2296*	(0.1323)	-	-
avg hoster viewership stream	0.0101**	(0.0039)	0.4823**	(0.2269)	0.0054***	(0.0008)
avg hoster viewership 7 days	-0.0129**	(0.0051)	-0.1947	(0.125)	-	-
current viewership	0.0065	(0.005)	-0.0945	(0.1251)	-	-
avg viewership stream	-0.0032	(0.0049)	0.2122	(0.2396)	-	-
avg viewership 7 days	-0.0	(0.0052)	0.0708	(0.114)	-	-
current followers	0.2001**	(0.0843)	-	-	-	-
followers gain stream	-0.001	(0.0029)	0.2393**	(0.1053)	-	-
followers gain 7 days	0.0027	(0.0042)	0.0579	(0.1198)	-	-
# years exists for	-0.2395**	(0.1006)	-	-	-	-
is partnered	-0.0257***	(0.0086)	-	-	-	-
english channel	-0.0037	(0.0046)	-	-	-	-
autoplay VODs	-0.0043	(0.0099)	-	-	-	-
is featured	-0.0473*	(0.0262)	-	-	-	-
is costreaming	0.0064***	(0.0018)	0.137***	(0.0366)	-	-
streamer level	-0.028	(0.0324)	-0.5209	(0.5161)	-	-
streamer experience (in 1000)	0.087	(0.0784)	-1.3634	(1.2201)	-	-
all-time viewers (in 1000)	-0.2082**	(0.1051)	-6.5792	(4.1685)	-	-
stream length hours	0.0376***	(0.0027)	0.5028***	(0.0403)	-	-
7 days hours streamed	-0.0022	(0.0036)	0.0061	(0.0578)	-	-
game viewership	-0.0007	(0.0027)	-0.0194	(0.0509)	-	-
game streamers	0.0079**	(0.0032)	0.1117**	(0.0532)	-	-
all current viewers	0.0027	(0.0038)	0.087	(0.0609)	-	-
all live channels	0.0438	(0.0853)	0.6428	(1.3549)	-	-
all current costreamers	-0.0329	(0.0203)	-0.5514*	(0.2936)	-	-
all live english channels	-0.0165	(0.0778)	-0.0774	(1.2612)	-	-
Fixed effects		Yes		Yes		No
Observations		23867		12766		23867
R squared		0.627		-		-

Note: the table above presents full results for modelling probability of hosting at the end of the stream. Note that all aggregations of hoster number and viewership through hosters (first 6 rows) are expressed in 100 of units, whereas all other control variables are normalized; this is done in order to avoid overflows in MLE estimation. Column 'Double ML' presents additional results for Double Machine Learning routine as in Chernozhukov et al. (2018). Nuisance parameters are estimated through Random Forest (200 trees, 5 features at each split) and 30 cross-fitting runs, with 5 folds each. Estimates and standard errors across cross-fitting runs are aggregated using medians.
See full description in main text and in Table 7.

Table 12: Full table for the hostee viewership results

	log(viewership)		log(followers)		DML: viewership	
	estimate	std err	estimate	std err	estimate	std err
current # hosters	-0.0006	(0.0009)	-0.0007	(0.0013)	-	-
avg # hosters stream	-0.0043*	(0.0023)	-0.0059*	(0.0031)	-0.0016	(0.0023)
avg # hosters 7 days	0.0047	(0.0034)	0.0014	(0.0046)	-	-
current hoster viewership	0.0001	(0.0003)	-0.0002	(0.0003)	-	-
avg hoster viewership stream	-0.0013***	(0.0005)	-0.0013*	(0.0006)	-0.0012**	(0.0005)
avg hoster viewership 7 days	0.0007	(0.0009)	0.0023*	(0.0013)	-	-
current viewership	-0.0002	(0.0002)	-0.0004*	(0.0002)	-	-
avg viewership stream	0.0002	(0.0002)	0.0003	(0.0003)	-	-
avg viewership 7 days	0.0002	(0.0002)	0.0002	(0.0003)	-	-
current followers	0.0	(0.0)	0.0	(0.0)	-	-
followers gain stream	0.0012***	(0.0004)	0.002***	(0.0006)	-	-
followers gain 7 days	-0.0002**	(0.0001)	-0.0002	(0.0001)	-	-
# years exists for	-0.0	(0.0)	-0.0	(0.0)	-	-
is partnered	-0.0933	(0.1145)	-0.025	(0.1545)	-	-
english channel	1.1116	(1.3542)	1.3845	(1.8277)	-	-
autoplay VODs	0.1506	(0.105)	0.1005	(0.1418)	-	-
is featured	-1.5149	(1.3881)	-0.6775	(1.8736)	-	-
is costreaming	0.1025**	(0.0409)	-0.1798***	(0.0553)	-	-
streamer level	0.0275	(0.0207)	0.0594**	(0.0279)	-	-
streamer experience (in 1000)	0.0	(0.0)	-0.0	(0.0)	-	-
all-time viewers (in 1000)	-0.0	(0.0)	0.0	(0.0)	-	-
stream length hours	-0.0001	(0.0004)	-0.0006	(0.0005)	-	-
7 days hours streamed	0.0	(0.0001)	-0.0001	(0.0001)	-	-
game viewership	0.0***	(0.0)	0.0*	(0.0)	-	-
game streamers	-0.0	(0.0)	0.0	(0.0)	-	-
all current viewers	0.0	(0.0)	-0.0	(0.0)	-	-
all live channels	0.0002**	(0.0001)	0.0	(0.0001)	-	-
all current costreamers	0.0	(0.0001)	0.0001	(0.0001)	-	-
all live english channels	-0.0002***	(0.0001)	-0.0001	(0.0001)	-	-
Fixed effects		Yes		Yes		No
Observations		18623		18623		18623
R squared		0.298		0.308		-

Note: the table above presents partial results for modelling log viewership and log followers of the chosen hostee conditional on hosting. Column 'DML: viewership' presents results for regressing log viewership using Double Machine Learning algorithm as in Chernozhukov et al. (2018). Note that nuisance estimators and tuning parameters are set in the same manner as in hosting probability estimation.
See full description in main text and in Table 8.

Table 13: Hosting decision: subsample with no (recent) direct reciprocity

	FE LPM		Conditional Logit		Double ML	
	estimate	std err	estimate	std err	estimate	std err
current # hosters	0.07	(0.0444)	2.2871***	(0.446)	-	-
avg # hosters stream	0.1693*	(0.0904)	3.2654***	(1.0226)	-0.0437	(0.0676)
avg # hosters 7 days	-0.1026	(0.1051)	-0.3218	(1.6242)	-	-
current hoster viewership	0.0086	(0.0068)	0.3009**	(0.1475)	-	-
avg hoster viewership stream	0.0147**	(0.006)	0.513*	(0.2648)	0.0053***	(0.0008)
avg hoster viewership 7 days	-0.0085	(0.0074)	-0.316	(0.4548)	-	-
current viewership	0.0143**	(0.0066)	-0.0071	(0.1365)	-	-
avg viewership stream	-0.0123	(0.0076)	0.0913	(0.2721)	-	-
avg viewership 7 days	-0.0031	(0.0093)	0.0497	(0.1553)	-	-
current followers	0.2112	(0.136)	-	-	-	-
followers gain stream	-0.0005	(0.0047)	0.1972*	(0.1015)	-	-
followers gain 7 days	0.0059	(0.0071)	0.0539	(0.1216)	-	-
# years exists for	-0.0665	(0.143)	-	-	-	-
is partnered	-0.0391**	(0.0161)	-	-	-	-
english channel	-0.0074	(0.0072)	-	-	-	-
autoplay VODs	-0.0076	(0.0149)	-	-	-	-
is featured	-0.0067	(0.0387)	-	-	-	-
is costreaming	0.0087***	(0.0029)	0.141***	(0.0411)	-	-
streamer level	-0.0286	(0.046)	-0.2139	(0.6391)	-	-
streamer experience (in 1000)	-0.0866	(0.1147)	-1.7522	(1.4167)	-	-
all-time viewers (in 1000)	-0.3378**	(0.1553)	-9.9469	(6.1744)	-	-
stream length hours	0.0515***	(0.004)	0.5378***	(0.0466)	-	-
7 days hours streamed	-0.0108**	(0.0052)	-0.101	(0.0689)	-	-
game viewership	0.0007	(0.0041)	-0.0065	(0.0618)	-	-
game streamers	0.0075	(0.0048)	0.0837	(0.0654)	-	-
all current viewers	0.004	(0.0055)	0.0829	(0.0696)	-	-
all live channels	0.063	(0.126)	0.2107	(1.629)	-	-
all current costreamers	-0.0241	(0.0295)	-0.2899	(0.3441)	-	-
all live english channels	-0.0447	(0.1151)	0.0692	(1.5118)	-	-
Fixed effects		Yes		Yes		No
Observations		14048		7865		14048
R squared		0.672		-		-

Note: the table above presents the robustness results for hosting probability models over a subsample of observations where i has not been hosted by chosen hostee j in the last 7 days. Subsampling over main models is done to ensure that there is no (recent) direct reciprocity involved in the decision to host at the end of the stream.
See full description in main text and in Table 7.

Table 14: Hostee viewership: subsample with no (recent) direct reciprocity

	log(viewership)		log(followers)		DML: viewership	
	estimate	std err	estimate	std err	estimate	std err
current # hosters	-0.0008	(0.0013)	-0.001	(0.0018)	-	-
avg # hosters stream	-0.0067*	(0.0036)	-0.0049	(0.005)	0.0014	(0.0035)
avg # hosters 7 days	0.0108**	(0.0051)	0.0002	(0.0072)	-	-
current hoster viewership	0.0006	(0.0004)	0.0005	(0.0005)	-	-
avg hoster viewership stream	-0.0011	(0.0007)	-0.0009	(0.001)	-0.0013	(0.0009)
avg hoster viewership 7 days	-0.0017	(0.0014)	-0.001	(0.002)	-	-
current viewership	0.0	(0.0002)	-0.0002	(0.0003)	-	-
avg viewership stream	-0.0001	(0.0003)	0.0	(0.0004)	-	-
avg viewership 7 days	-0.0002	(0.0004)	-0.0001	(0.0005)	-	-
current followers	-0.0	(0.0)	-0.0001	(0.0001)	-	-
followers gain stream	0.002***	(0.0006)	0.0023***	(0.0009)	-	-
followers gain 7 days	-0.0004**	(0.0001)	-0.0003	(0.0002)	-	-
# years exists for	-0.0	(0.0)	-0.0	(0.0)	-	-
is partnered	-0.2989*	(0.1766)	0.0079	(0.2458)	-	-
english channel	2.8059	(1.8817)	5.1038*	(2.6191)	-	-
autoplay VODs	0.2216	(0.1579)	0.0433	(0.2198)	-	-
is featured	-2.7831	(2.0483)	-3.0194	(2.851)	-	-
is costreaming	-0.0435	(0.0658)	-0.3347***	(0.0916)	-	-
streamer level	0.0104	(0.0312)	0.034	(0.0434)	-	-
streamer experience (in 1000)	0.0	(0.0)	-0.0	(0.0)	-	-
all-time viewers (in 1000)	0.0	(0.0)	0.0**	(0.0)	-	-
stream length hours	-0.0009	(0.0006)	-0.0014	(0.0008)	-	-
7 days hours streamed	0.0	(0.0001)	-0.0	(0.0002)	-	-
game viewership	0.0***	(0.0)	0.0***	(0.0)	-	-
game streamers	-0.0	(0.0)	-0.0	(0.0)	-	-
all current viewers	0.0	(0.0)	0.0	(0.0)	-	-
all live channels	0.0001	(0.0001)	0.0001	(0.0002)	-	-
all current costreamers	0.0001	(0.0001)	0.0001	(0.0001)	-	-
all live english channels	-0.0002	(0.0001)	-0.0001	(0.0002)	-	-
Fixed effects		Yes		Yes		No
Observations		8989		8989		8989
R squared		0.361		0.36		-

Note: the table above presents the robustness results for chosen hostee viewership models over a subsample of observations where i has not been hosted by chosen hostee j in the last 7 days. Subsampling over main models is done to ensure that there is no (recent) direct reciprocity involved in the decision to host at the end of the stream.
See full description in main text and in Table 8.

Table 15: Robustness results: linear probability models with interactions

	is partnered		# years exists for		current followers		stream length hours	
	estimate	std err	estimate	std err	estimate	std err	estimate	std err
avg # hosters stream	0.3305***	(0.0637)	0.4155***	(0.1132)	0.221***	(0.0525)	0.342***	(0.0605)
interaction: # hosters	-0.3099***	(0.0632)	-0.1231***	(0.0406)	-25.1402***	(7.7113)	-25.8316***	(3.8375)
avg hoster viewership stream	0.0084	(0.0054)	0.0607***	(0.0225)	0.0059	(0.0047)	0.0259*	(0.0138)
interaction: hoster viewers	0.0013	(0.0089)	-0.0166**	(0.0074)	1.6257	(2.3857)	-1.6035	(1.1525)
current # hosters	0.0378*	(0.02)	0.0375*	(0.0203)	0.0345*	(0.02)	0.0391**	(0.0199)
avg # hosters 7 days	-0.0194	(0.0596)	-0.0116	(0.0598)	0.0073	(0.0603)	-0.0012	(0.0612)
current hoster viewership	0.0032	(0.0034)	0.0017	(0.0034)	0.0026	(0.0033)	0.0019	(0.0033)
avg hoster viewership 7 days	-0.0128**	(0.0051)	-0.013**	(0.0051)	-0.0121**	(0.0051)	-0.0129**	(0.0051)
current viewership	0.0063	(0.005)	0.0066	(0.005)	0.0054	(0.0049)	0.0037	(0.005)
avg viewership stream	-0.0021	(0.0052)	-0.0029	(0.005)	-0.0008	(0.0054)	-0.0007	(0.0052)
avg viewership 7 days	-0.0001	(0.0051)	-0.0	(0.0051)	-0.001	(0.0051)	0.0004	(0.0051)
current followers	99.6054***	(35.4794)	92.7664***	(35.251)	97.1051***	(35.6227)	71.6026**	(35.1195)
followers gain stream	-0.0001	(0.003)	-0.0011	(0.003)	-0.0001	(0.003)	0.0037	(0.0031)
followers gain 7 days	0.0025	(0.0042)	0.0022	(0.0042)	0.0017	(0.0041)	0.0009	(0.0042)
# years exists for	-0.2808**	(0.1123)	-0.2475**	(0.1124)	-0.299***	(0.1126)	-0.2858**	(0.1122)
is partnered	0.0174	(0.026)	-0.055***	(0.0181)	-0.0546***	(0.0181)	-0.0532***	(0.018)
english channel	-0.0029	(0.0042)	-0.0032	(0.0044)	-0.0049	(0.0049)	-0.005	(0.0055)
autoplay VODs	-0.0038	(0.0099)	-0.0042	(0.0099)	-0.0046	(0.0099)	-0.0052	(0.0099)
is featured	-0.0463*	(0.0263)	-0.0617**	(0.0265)	-0.0462*	(0.0262)	-0.0502*	(0.0262)
is costreaming	0.0068***	(0.0018)	0.0066***	(0.0018)	0.0067***	(0.0018)	0.0068***	(0.0018)
streamer level	-0.0375	(0.0325)	-0.0411	(0.0327)	-0.0309	(0.0324)	-0.0272	(0.0324)
streamer experience (in 1000)	0.072	(0.0785)	0.0843	(0.0783)	0.0767	(0.0784)	0.0791	(0.0783)
all-time viewers (in 1000)	-0.1536	(0.104)	-0.1662	(0.1049)	-0.0816	(0.1039)	-0.1587	(0.1078)
stream length hours	12.9704***	(0.9743)	13.3953***	(0.9729)	13.0791***	(0.9739)	17.8121***	(1.2112)
7 days hours streamed	-0.0021	(0.0036)	-0.0021	(0.0036)	-0.0023	(0.0036)	-0.002	(0.0036)
game viewership	-0.0006	(0.0027)	-0.0007	(0.0027)	-0.0007	(0.0027)	-0.0005	(0.0027)
game streamers	0.0078**	(0.0032)	0.0079**	(0.0032)	0.0079**	(0.0032)	0.0078**	(0.0032)
all current viewers	0.0026	(0.0038)	0.0025	(0.0038)	0.0027	(0.0038)	0.003	(0.0038)
all live channels	0.0405	(0.0852)	0.0446	(0.0853)	0.0396	(0.0852)	0.0383	(0.0851)
all current costreamers	-0.0337	(0.0203)	-0.0331	(0.0203)	-0.0325	(0.0202)	-0.0324	(0.0202)
all live english channels	-0.0114	(0.0778)	-0.0167	(0.0778)	-0.012	(0.0778)	-0.0094	(0.0777)
individual fixed effects		Yes		Yes		Yes		Yes
time fixed effects		Yes		Yes		Yes		Yes
nobs		23867		23867		23867		23867
rsquared		0.6273		0.6272		0.6273		0.628

Note: the table above presents partial results for Linear Probability Model with additional interaction terms to show heterogeneity in response to being hosted. See full description in main text and in Table 9.

4 Friendships accelerate exit from video game streaming platform once shutdown is certain

4.1 Introduction

Social networks are an inseparable part of the individual interactions. Recently, the effects of network structure on economic outcomes have been studied extensively given an increased availability of detailed data at the individual level (Advani and Malde, 2018). Characteristics such as degree centrality were related to diffusion of microfinance (Banerjee et al., 2013) social network adoption (Katona et al., 2011) or aggregation of information (Alatas et al., 2016), to name just a few examples.[27]

Considerably less attention has been devoted to the relation between network structure and collapsing online services. Saavedra et al. (2008) studied gradual collapse of New York garment industry and found better performance in networks with asymmetric degree edges. Garcia et al. (2013) provided a study of network resilience using k-core decomposition and the case of the collapse of Friendster alongside few other social network platforms. Using an example of Hungarian social media platform iWiW and its collapse, Török and Kertész (2017) highlighted the cascading effect of friends triggering each other to exit. Lőrincz et al. (2019) found decreasing positive returns to staying given size of the network, but negative correlation between the probability of staying and clustering of the individuals. Koltai et al. (2021) found that connecting different circles of friends on iWiW decreased the probability of early exit. Farronato et al. (2020) studied the effects of merging two online dog sitting platforms, with one absorbing the entire user base and closing the acquired platform few months after the merger.

In related literature on telecommunication industry where data on connections between users is often available, Kim et al. (2014) noted that node eigenvector centrality helps predict churn. Using a diffusion model, Dasgupta et al. (2008) showed that churn increases conditional on friends who churned. More generally, Verbeke, Martens and Baesens (2014) concluded that network attributes generally facilitate identifying churners and proposed an ensemble of relational and non-relational methods to incorporate

[27] For an extensive literature review of economics and econometrics of social networks, see for instance De Paula (2017), Jackson (2011) or Jackson (2014).

network structure into the prediction process. Óskarsdóttir et al. (2017) extended research by Verbeke, Martens and Baesens (2014), but also provided an extensive review of papers that related social network measures to consumer churn in telecommunication industry.

In this paper I study the relation between platform exit and the degree of the individual network using data from video game streaming, a relatively new market of services where individuals mostly broadcast a live feed of their gameplay. Video game streaming platforms allow for direct interaction between the content creator - the 'streamer' - and the audience, and that sort of community experience can be argued to be one of the reasons behind its popularity (Sjöblom and Hamari, 2017). The underlying market is continuously growing - largest video game streaming platform Twitch with 72.3% market share has doubled in hours watched over the last year to a staggering 6.3 billion in first quarter of 2021 over equally impressive 265 million hours streamed (May, 2021) - and the belief that it will only continue to grow is widespread across users (Johnson and Woodcock, 2019).

While streaming connects the brodcaster with the audience, it also allows streamers to interact with each other. One such activity, unique to the platform under study, is 'costreaming': a phenomenon where up to 4 channels join their streaming sessions in a split-screen format. By doing so, they all temporarily broadcast their combined video and audio feed. Costreamers can be seen as co-producers of a single content for their united audience, giving the viewer an option to switch between up to 4 viewpoints or watch all of them at the same time. Since costreaming can enhance the experience for the viewers and lead to a better content, having many costreamers can be regarded as bridging or coordination type of social capital - connecting several users through one costreaming session - as well as a favor type of social capital - exchanging the choice of games played together for their joint audience (Jackson, 2020, Koltai et al., 2021).

Such social capital can be particularly important when determining when to exit from a platform. In this paper I exploit the availability of unique and high frequency data on a major video game streaming platform that has officially announced its shutdown, to be scheduled within a month from the announcement. Though most streamers could have been expected to switch to another streaming service, the timing of that decision and remaining commitment to the closing platform is not trivial. In addition to individual strategic considerations, streamers may also consider the decision of their peers: they may coordinate the exit time and choice of the new platform or simply follow the first movers.[28]

I take an advantage of almost two months of data prior to the news about the shutdown to formulate full structure of an up-to-date costreaming network. While the true relations between streamers may be much more dense and complex - not all platform users do costream or express their friendships on the platform through costreaming - the network on costreaming itself is extremely accurate. With 40% of the sample costreaming over fair part of their online activity, it is by no means a niche activity. In contrast with social media friendships which often do not need to involve much online commitment, costreaming network consist of links between users that interact in real time in front of their potential audience.

The results of this study show that network size has a negative effect on activity once information about shutdown becomes common knowledge. Each extra costreaming link observed prior to the announcement results in almost 11% decrease in post-announcement hours online and 4.2% decrease in amount of logins to the platform, compared to the mean value in the sample. Although costreaming network formation prior to the announcement should be exogenous to exit after it, general decision to engage in costreaming may depend on individual characteristics such as a preference for multiplayer games that are also related to exit decisions. Taking that into consideration, I show that the baseline results are qualitatively robust to instrumenting potentially endogenous count of costreaming friends with an indicator for Spanish language channel, second most frequent language on the platform after English, including additional game specific control variables that may facilitate or inhibit costreaming among users, or controlling for the total count of costreaming hours prior to the announcement.

Descriptive evidence shows that gradual exit from the platform started already before the announcement. However, despite committing less resources into the platform afterwards, individuals with more costreaming friends are also less likely to churn prior to the news. I find that each extra costreaming friend results in 0.005 higher probability

[28] Streamers, unless being partners and hence having some restrictions, can in theory stream on several platforms at the same time. And in fact it seems that some did do just that: on the last day many used their time on the platform to simply rerun old streaming sessions and redirect viewers to channels on other platforms (Stephen, 2020). However, it is still an investment to manage an additional streaming session for a prolonged period of time, including the potential interactions with chat, and data indicates that most of the streamers seem to have left long before the actual shutdown has happened. In contrast with eg. Farronato et al. (2020) where one can announce services on many platforms with small additional costs (referred to as 'multi-homing'), here such mirrored usage may be much more expensive. Furthermore, there is no apparent link that such activity would in any way be correlated with having more or less costreaming friends prior to the announcement and so it should not be of a major concern to the current study.

of appearing online after the message about the shutdown was published. This result also remains qualitatively robust when adding game controls or total costreaming time. Additional analysis suggests that the negative net effect on total activity afterwards may, at least partially, come from peer effects, with more clustered groups of costreaming friends leaving earlier. Analyzing a subsample of individuals who left after the news, they were strongly influenced by those among their costreaming friends who last time appeared online before the announcement was published. This remains true even when considering churners who have left at least a week prior to the news, ensuring that close coordination between them and their peers who stayed is less likely.

Finally, the results for both online activity after the announcement as well as churning prior to it are robust even when considering a subsample of 2.43% most popular channels with at least 1000 followers. While the effect of network degree was lower than for the baseline sample, it remains strongly significant. With the baseline sample containing over 200 thousand relatively frequently streaming users, these robustness results further confirm that effects of costreaming network are present even when an established channel is at stake.

This paper contributes to the literature on the effects of network degree on consumer churn by analysing a particular type of links formed through co-produced broadcasts. It also provides an analysis of a unique event where a relatively healthy and popular online video game streaming platform has announced a shutdown. Network degree that significantly decreases the probability of churn prior to the announcement, speeds up the process of exit through peer effects - whether through coordination, group decision or simply following the first movers - once shutdown becomes certain. The results of this study show that introducing means of cooperation such as costreaming, a feature that is quite unique in its format to the platform under study, can benefit video game platforms by decreasing churn, but the very same measure can also speed up the collapse.

The rest of the paper is structured as follows: Section 2 introduces the data and presents the descriptive statistics. Section 3 presents the baseline model and section 4 describes the results. Finally, section 5 presents the discussion of the results and the conclusions.

4.2 Data

4.2.1 Platform shutdown and sample selection

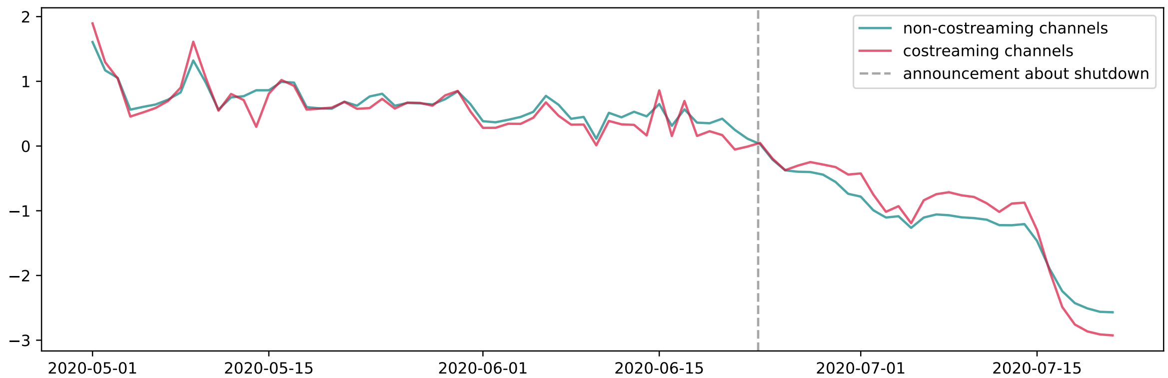
While recent pandemic has resulted in increased popularity of video game streaming platforms overall, growth was not evenly spread. In the last quarter of 2019 and first two quarters of 2020, Twitch, Youtube Gaming and Facebook Gaming all experienced a moderate growth in total hours watched while Mixer noted a steady decline, only to recover from it in the second quarter of 2020 (May, 2020a); this was despite some effort of Mixer to sign exclusive deals that attracted several top streamers from the other platforms. While it captured 31% of all unique channels and 14.4% of total hours streamed, at the same time it only accounted for 1.4% of the sum of hours watched in the second quarter of 2020. Despite that, the announcement about the platform's shutdown did not seem to be expected by the market participants.

Even though the platform under study was partnered with Facebook Gaming and streamers were offered favorable transitions before the shutdown (eg. matching the partnership status on the other platform; Warren, 2020), evidence suggests a moderate growth of the Facebook Gaming in terms of hours streamed. This could suggest that there are some switching costs involved in moving between the two platforms (Farronato et al., 2020), though data also indicates a large growth of the main player of the market, Twitch. With over 14% growth in streaming hours, compared to just 1% growth of Facebook Gaming, it can be argued that it was the biggest market player who gained most from the exit compared and not the partnered platform (May, 2020b).[29]

The following study makes use of a high-frequency unbalanced panel data on all channels active on a major video game streaming platform Mixer. Between the periods of 1st of May 2020 4:20PM CET and 21st of July 2020 9:30AM CET, approximately every 5 minutes a query was sent to the platform API, gathering information about all currently streaming channels and their characteristics such as games played, viewership, being featured or costreaming. In total, a little over 80 days of data was gathered, constituting more than 23200 snapshots captured every 5 minutes, each on average

[29] At the same time, Youtube Gaming lost market share. See May (2020a) and May (2020b) for detailed reports and descriptive statistics on market shares and user statistics for the major video game streaming platforms in the market.

Figure 19: Daily average count of online channels for costreamers and non-costreamers



Note: figure above presents average daily count of channels that have at least 1 follower, with distinction between channels that were part of any costream session before the announcement and channels that were not. Note that both daily counts of channels are normalized against therir means and standard deviations to allows for clear comparison given the substantial difference in group sizes (for that reason, labels for y axis are omitted as they are not informative). Note also that although some channels may be part of a costreaming session that has no other costreamers, they are still included as costreaming channels for the purpose of this plot.

containing approximately 15 thousand live channels.[30]

With the announcement about the platform shutdown occurring on the 22nd of June 2020, data can be split into two parts: post-announcement which is the relevant sample period used to define aggregate online activity, and pre-announcement period that is used for creating control variable such as popularity measures or channel settings. In addition, extensive length of the pre-announcement period allows to characterize an up-to-date state of costreaming networks that spans 50 days of interactions between all the channels on the platform.

Figure 19 presents average daily count of online costreamers and non-costreamers. Note that both counts are seperately normalized to allow for comparison (each time series is divided by its own mean and standard deviation). Interestingly, one can notice a steady decrease of active channels that started already months prior to the announcement, this trend being common to both groups of users. There seems to be a difference between costreamers and non-costreamers, with the first group being relatively more active in first 3 weeks compared to the last week, though the relation is not clear. After

[30] Note that the last approximately 36 hours of data prior to actual shutdown is missing due to termination of the data gathering process given uncertainty about API behavior after the shutdown. Note also that depending on the varying platform traffic, gathering data on all active channels can take 1-3 minutes on average, and hence the distance between snapshots is only approximate. This, however, is of no major consequence to the study since sample data used is largely aggregated.

the announcement, two waves of mass exits common to both groups can be observed: one right after the announcement, and another one a week before the closure.

In the first week of the sample there were 19442 online channels on average during any time of the day, compared to 16919 channels live a week before the announcement, 10816 two weeks after the announcement and just 5039 channels in the 4th week after the announcement right before the closure. Although the process of leaving the platform was gradual and there is no clear unique tipping point, one can conclude that it started long before the actual announcement and a majority of channels left weeks before the actual shutdown has taken place.

With a high fraction of channels appearing on the platform only a couple of times over the sample period, I focus on a subset of regularly streaming channels and their characteristics prior to the announcement being made. I construct a sample of streamers that have logged in at least 7 times and streamed for at least 14 hours in the pre-announcement period. This condition roughly translates to streaming at least once a week for 2 hours, a minimal restriction considering that professional streamers often stream at schedules similar to a full-time job. Furthermore, I select channels that have at least 1 follower at the point when they are first recorded in the data to ensure that they have streamed for some audience in the past.

The sample under consideration consists of 218660 channels in total. Median channel in the sample is 2 years old, has 8 followers and 94 unique viewers ever watching their stream. Only approximately 0.43% of the channels in the sample are partnered. Median channel streams approximately twice as often as the minimum restriction which translates to more or less 2 times a week for two hours each time. Nevertheless, since this can be considered a relatively small commitment to streaming, as a robustness I consider using a subsample of individuals who can be described as professional streamers based on their follower count and for whom the decision how much to stay on the platform and when to switch is of arguably higher importance.

While exit timing is particularly interesting in case of a platform shutdown, in the following study I focus instead on the total time spent streaming post-announcement. Exit from free services where there are no costs to delaying or postponing the decision may not reflect the actual commitment to the platform. This issue can be particularly detrimental in cases when streamers continue to log in at large intervals in order to redirect their viewers to a new channel on another platform or log in solely on the shutdown day despite not being active for few weeks already. In such cases while the actual exit did not occur, one can hardly argue that an individual is active on the platform.[31] In contrast, online activity on the platform bears some cost to the individual and actually measures the involvement of the streamer. If networks do play a role in the decision when to leave the platform, they should also have a strong relation to the sum of online activity after the announcement.Figure 22 in the Appendix shows the distribution of total sum of hours online and count of login instances after the announcement. I define the login instance as appearing online after at least one hour of being consistently offline.[32] Using descriptive statistics in Table 19 in the Appendix one can see that over approximately 50 days prior to the announcement channels streamed on average 49 hours over 22 instances, whereas over approximately 29 days after the news this average drop to just little below 9 hours over 4 instances. Back of the envelope calculations indicate that on average there is a 67.44% decrease in daily hours online and 66.7% decrease in daily login instances. Having in mind that 26.99% of the sample have never appeared online after the announcement was made, this is a substantial drop in online activity that may be attributed to exits from the platform prior to the actual shutdown.

4.2.2 Costreaming in relation to the activity post-announcement

In the pre-announcement period there were 210 thousand distinct costreaming sessions over 0.79 million hours in total (where the sums are calculated by adding individual streaming time of all channels involved in the given costream). A total sum of 324 thousand unique channels were involved in costreaming, creating 446 thousand costreaming pairs among them. With information over the entire platform, this dataset allows to define the entire network of costreaming over the period of almost two months which should be an accurate and up-to-date state of the costreaming relations among channels.

[31] It is possible to determine an artificial cutoff point for consumer churn; eg. Periáñez et al. (2016) define that cutoff point to be 10 days of inactivity while Óskarsdóttir et al. (2017) suggest 30 days based on the existing literature. However, with only 30 days of data before the shutdown, it is hard to determine a reasonable cutoff point that does not drop a substantial part of the data set. Furthermore, since last 36 hours prior to the actual shutdown are missing from the sample, using sum of activity over the sample period avoids censoring of the data or ad hoc determination which observations to censor (which also causes censoring to be dependent on exit).

[32] Given the nature of the main data set, this translates to channel being offline for 12 consequent snapshots that are gathered approximately every 5 minutes.

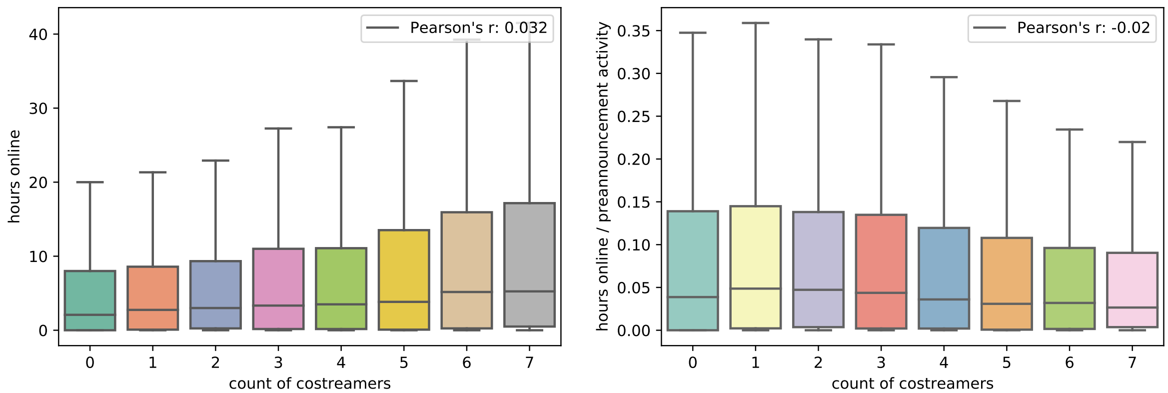
Two channels are connected with an edge if they have costreamed together at any point during the sample period prior to the announcement about the platform shutdown. Based on such constructed network, for each channel I generate a count of costreamers: a count of distinct channels that i shares an edge with. In addition to counting the degree centrality - how many costreamers does i have - I also construct the local clustering coefficient defined as a fraction of possible connections between friends of i that actually exist. As such, it will take value between 0 and 1 depending on how many of $\frac{n_i(n_i-1)}{2}$ potential links are present among n_i costreamers of i.

Table 19 in the Appendix presents the summary statistics on a selection of network characteristics. Approximately 39.62% of the channels in the sample have at least 1 costreamer recorded in the pre-announcement period. This is a substantial fraction, showing that this form of cooperation is not rare among platform users. At the same time the distribution of the costreamer count is highly skewed, with a majority of costreaming channels having 1 (20.54% of the sample) or 2 (9.74% of the sample) related friends. Only approximately 4.82% of channels in the sample have 4 or more costreamers.

Mean time costreaming is quite low at 3.38 hours with an average of 1.96 costreaming sessions (note that the sum of hours is multiplied by the costreamer count in case there is more than 1 costreamer in the session). However, once all non-costreaming channels are excluded, this conditional mean raises to 8.53, with an average of 4.95 sessions among costreamers in the period prior to the announcement. Similarly, when excluding the channels that have less than 2 costreamers (and hence by construction their local clustering coefficient is equal to 0), mean local clustering coefficient is 0.406. This is expected given that costreaming can allow up to 4 players to join the same costreaming session together and hence many of these groups are likely to be highly clustered.

In order to study the relation between network degree and online activity after the announcement, I group sampled channels by costreamer count and plot the distribution of their hours online in graph on the left of Figure 20. Looking at the median values, there is a clear pattern indicating higher post-announcement activity for channels with higher amount of costreamers. However, this conclusion may be misleading given that by construction count of costreamers is strongly related to the total streaming time pre-announcement. Graph on the right of of Figure 20 attempts to remedy this problem by plotting the fraction of streaming post-announcement divided by the fraction of streaming pre-announcement. Once related to the activity post-announcement relative to the same activity pre-announcement, costreamer count has a visible negative effect:

Figure 20: Hours online post-announcement conditional on the count of costreamers



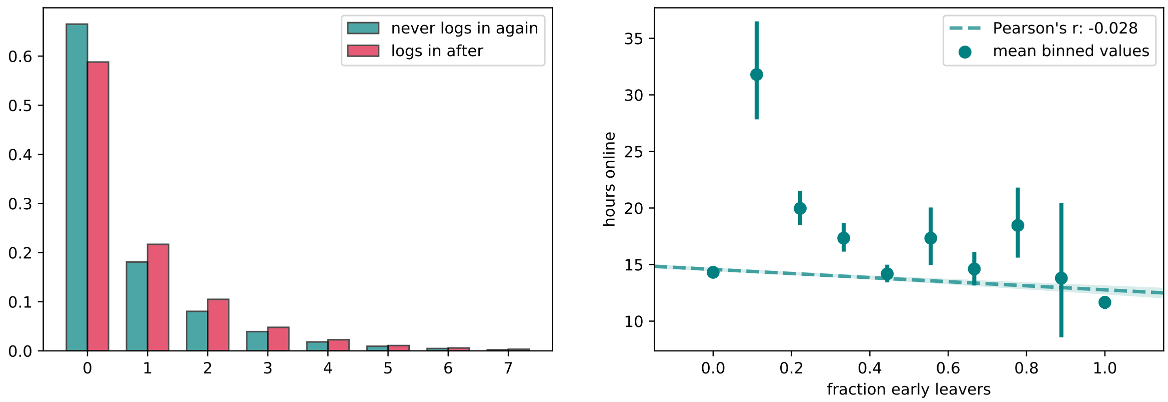
Note: boxplot on the left presents the distribution of total hours streaming after the announcement about the shutdown was made conditional on the count of costreamers, with maximum costreamer count trimmed at 99th quantile for clarity. Boxplot on the right presents a similar plot, but hourly online activity is further divided by the same variable aggregated over the pre-announcement period. Hence plot on the right presents a relative streaming choice conditional on costreamer count. See Figure 23 in the Appendix for similar set of plots for login instances as well as the relation between hours online and login instances with local clustering coefficient of the individual costreaming network.

as it increases, channels tend to spend less time online proportionally. Note that graphs in the top row of Figure 23 in the Appendix show that similar a conclusion can be drawn when replacing hours online with the count of login instances.

It is possible that channels with more costreamears find it easier to adapt and hence are more likely to be the first to exit or even leave prior to the announcement. I define early leavers as having zero hours online within the sample period post-announcement, indicating they have likely left before or immediately after they have heard the news about the shutdown. Plot on the left of Figure 21 shows proportional distribution of the sample split across different costreamer counts, with distinction between channels which were the early leavers and those which logged at least once after the announcement was made. One can clearly see that early leavers are much more likely to have no costreamers at all. Note that while the difference seems to be smaller as the costreamer count increases, it remains proportionally stable at approximately 12% lower fraction of early leavers among individuals with costreamers.

Given the observable gradual decay of usage observed in Figure 19, I further investigate the distribution of exit time among early leavers. Among over 59 thousand individuals who never appear online after the announcement, 52.4% were last seen over a week prior to the announcement, 31.5% over 2 weeks ago and 18.6% in over 3 weeks time. This indicates that a majority of early leavers is likely to be in fact churners that

Figure 21: Early exits and peer effects on exit



Note: graph on the left presents early leavers (individuals who never appear online after the announcement was made, in teal color) against individuals who appear online at least once more after the announcement was made (red color), split by the costreamer count. Both these groups are further divided by their respective group size to obtain comparable fraction and their distribution conditional on costreamer count. Streamers with no costreaming friends are much more likely to be early leavers, while the reverse is true for individuals with at least one costreamer. Graph on the right plots total count of hours online for streamers that did log in at least once after the announcement, against fraction of their costreamers that are early leavers. Note that the fractions are binned into groups (with bootstrapped 95% confidence intervals) for the clarity of the exposition. Regression line indicates that as the fraction of first movers among costreaming peers increases, channels are spending less online time on the platform after the announcement.

left the platform before the announcement and for reasons unrelated to the shutdown.

Although there seems to be a negative relation between costreamer count and proportional sum of activity, descriptive evidence suggests that costreamers are less likely to be the early leavers or churners. Coupled with negative effect of clustering within networks on hours and instances online after the announcement as observed in the bottom row of Figure 23 in the Appendix, overall evidence suggests that costreamers may be influenced by their friends. I investigate that possibility by calculating the proportion of costreamers that are early leavers and plot it against hours online for individuals who stay.

Relevant graph is available on the right of Figure 21. Note that for this graph I subsample individuals that appear online after the announcement was made while excluding the early leavers and treating their decision as exogenous (this is not a strong assumption given that a majority of early leavers seem to have left at least a week before the announcement about the shutdown was made). There is a clear negative relation between fraction of first movers among friends and the total sum of online activity. Clearly, having a large fraction of potential costreaming friends leaving immediately after the announcement (or, in some cases, even before that) has a negative effect on one's total amount of activity.

Overall, descriptive evidence suggests that channels with more costreaming friends commit less hours and log in fewer times on the platform after it becomes known that it will shut down. While they are less likely to be the early leavers, they are further influenced by their friends and, as a consequence, the net effect on the total count of hours and login instances appears to be negative. At the same time, this negative effect appears only when controlling for general, channel specific streaming habits and hence more comprehensive model is necessary to fully capture the effects observed in the data.

4.3 Model

Descriptive analysis shows substantial heterogeneity in the usage of the platform after the announcement about the shutdown was made. In particular, one can observe a clear, negative relation between relative activity before and after the announcement and the degree centrality of the individual costreaming network expressed through the count of costreamers. I attempt to model this relation with a following linear model:

$$Y_i = \alpha + \beta C_i + \delta X_i + \nu_i, \tag{6}$$

where Y_i is the total sum of hours streamed over after the announcement, C_i is the count of costreamers of i as defined in the previous section, and X_i is a set of controls calculated over the period prior to the news about shutdown.

Although estimating the effect on the total count of hours online after the announcement allows me to draw inference about the commitment of the users to the platform, this variable is not fully robust to the timing of the announcement. While it may appear that some users streamed afterwards, they may have just continued their sessions that started before the announcement was made. To remedy that, I estimate an additional specification where Y_i is the count of login instances. Total time streaming and count of login instances measure slightly different concepts since channels differ substantially in how long and how regularly they stream. Nevertheless, I expect that network effects have a qualitatively similar influence over them.

Given that individuals who stream more are also more likely to have costreaming partners, I include count of hours streaming and count of login instances prior to the announcement as control variables. I also add several popularity characteristics such as follower count or partnership status that are related to the commitment to the platform.

These measures are fixed at the value recorded when the channel first appeared in the sample in order to control for general popularity and not recent dynamics that may have been partially caused by costreaming.

Since costreaming may be significantly altered after the announcement, I use data before the event to identify prior costreaming links on the platform. For that reason, the effect of the count of costreaming links is limited to the particular type of recent connections. However, such time restriction also provides some confidence that the recorded networks are established before the decisions about the timing of the exit or how much activity to commit to the platform.

Nevertheless, the number of costreamers one has may still be endogenous. Costreaming in the split-screen format is a feature relatively unique to the platform in this study and individuals who find it unintuitive or unappealing may also be more likely to move first and exit early. They may also have a preference for multiplayer games that facilitate costreaming and can be related to the timing of the exit decision. This potentially creates an omitted variable problem that would bias the estimates on the costreamer count.

To remedy this, I take an instrumental variable approach, using indicators for Spanish speaking channels as instruments for the potentially endogenous count of costreamers. Spanish is the next most spoken language on the platform after English. I assume that non-English streams, constituting only approximately 8% of the total count of channels in the sample, have much smaller communities and consequently may face more difficulties finding a costreamer who broadcast in the same language. At the same time, video game streaming platform under study provides full support in over 21 languages, including Spanish, and so language of the stream should not affect familiarity with the platform or attitude towards the costreaming as a feature of the platform.

Language may still be related to cultural differences between users which in turn may affect the rate of exit and remaining commitment to the platform and this can invalidate the exogeneity of the instrument. Given unavailability of better instruments, I treat instrumental variable approach as a qualitative robustness check rather than the potential main modeling approach.

I also investigate whether individuals with more costreaming partners appear for less hours online simply because they are exiting early, ie. they never appear online after the announcement. This is particularly interesting given the descriptive analysis that revealed substantial exit prior to the announcement, with as much as 30% of the early leavers logging in last time more than 2 week prior and likely being churners who

Table 16: Effects of network clustering on platform usage after the announcement

	OLS reduced		OLS baseline		OLS instances	
	estimate	std error	estimate	std error	estimate	std error
count of costreamers	0.83***	(0.046)	-1.008***	(0.057)	-0.17***	(0.01)
Control variables		No		Yes		Yes
R squared		0.004		0.247		0.208
Observations		218660		218660		218660

Note: table above presents partial results for regressing count of online activity on the platform after the shutdown announcement on the size of individual costreaming networks. Column 'OLS reduced' presents initial results without inclusion of any control variables. Column 'OLS baseline' presents results for regressing total hours online post-announcement in the baseline specification specified in the previous section. Column 'OLS instances' runs a similar regression, but using login instances instead of total hours online as the dependant variable. Standard errors are heteroscedasticity robust standard errors.
Full results of the estimation can be found in Table 20 in the Appendix.

made a decision regardless of the shutdown news. To investigate that, I run a similar regression as the baseline equation 17, but replacing hours online with an indicator variable for having zero hours online after the announcement was made and estimate it with a linear probability model.

4.4 Results

4.4.1 Baseline results

Partial results for baseline model are displayed in Table 16. Each additional costreamer in the presample period increases activity on the platform post-announcement by approximately 50 minutes. Once additional controls are included (column 'OLS baseline'), the effect switches sign to a one hour decrease for every extra costreamer. Given the average count of streaming hours post-announcement in the sample, this is approximately 11% decrease in the total activity. Hence the results are in line with the plots in the descriptive part: since costreamer count is highly correlated with pre-announcement streaming time, the correlation is positive. However, once this is accounted for, the actual underlying effect is negative and strongly significant at 1%.

Running a similar regression on the number of login instances yields very similar conclusions: each extra costreamer in the pre-announcement period results in 0.17 less login instance post announcement, a 4.2% decrease in comparison to the mean value.

Individuals with fewer cooperative ties on the platform log in fewer times after the announcement and spend substantially less hours in total streaming on the platform during the sample period.

The results for a linear probability model that estimates the probability of having no activity online after the announcement are available in the column 'exits before' in Table 17. Each extra costreamer results in 0.005 decrease in probability of early exit, the effect being significant at 1%. Channels that have more costreaming links are actually more likely to login at least once after the announcement, but for much shorter periods, resulting in overall net effect on activity post-announcement being negative. This is also in line with the descriptive evidence in the previous section.

In order to identify potential mechanisms behind the baseline results, I run few additional models similar to the baseline equation 6, results of which are partially displayed in Table 17. Since costreaming allows for cooperation of up to 4 channels, one can expect that the outcome is different for clustered groups. In particular, if there is coordination involved or peer effects present, groups where costreamers of i are also costreaming with each other should exhibit a stronger effect.

The results of including a local clustering coefficient and its interaction with costreamer count are displayed in the second column of Table 17. While the effect of costreamer count remains negative and significant, clustering positively influences the decision on how much longer to stream on the platform. Importantly, there is an additional negative effect for each extra costreamer depending on the degree of clustering of the network. Having a larger network pre-announcement indicates lower activity, the effect being approximately 5 times as strong for fully clustered networks as compared to a network where none of the costreamers of i had costreamed with each other. This supports the idea that clustered networks of costreamers generally stay longer, but this effect is decreasing the larger the network.

Given that the effect is stronger among clustered groups of friends and that individuals with more costreamers are less likely to exit early despite having lower total activity on the platform, I investigate whether peer effects play a role in the decision on how much to stream on the platform after the announcement. To do so, for each channel I count costreamers that are early leavers and run a regression on a subsample of channels that appeared online at least once after the announcement about shutdown was made. In that sense, I am estimating the effect of early leavers on their peers who are active at least to the point of the announcement. Consequently, the estimated effect is exclusively valid for the subsample that does not include early movers and under the

Table 17: Early exits and potential mechanisms behind the baseline results

	exits before		with clustering		no early exits	
	estimate	std error	estimate	std error	estimate	std error
count of costreamers	-0.005***	(0.001)	-0.569***	(0.067)	-0.599***	(0.088)
local clustering coefficient			4.278***	(0.632)		
costreamers * clustering			-2.384***	(0.284)		
count early leavers					-2.094***	(0.162)
R squared		0.074		0.241		0.275
Observations		218164		218164		159278

Note: table above presents partial results investigating the potential mechanisms behind the baseline results. Column 'exits before' runs a Linear Probability model to predict the probability that a channel will never appear online again after the announcement about the shutdown was made. Column 'with clustering' adds the interaction between the count of costreamers and local clustering coefficient, ie. what fraction of possible edges among the costreamers of a given channel do actually exist. Finally, column 'no early exits' reduces the sample to channels that have appeared online at least once after the announcement was made. Furthermore, it counts costreamers of a given channel that are first movers, ie. channels that were dropped from the sample since they have never appeared online after the announcement about the shutdown was made, and includes that variable in the regression.
Full results of the estimation can be found in Table 21 in the Appendix.

assumption that their exit is known and happens before the decision on online activity by other channels in the sample. Again, this is not a very strong assumption given that much of the exits from happened long before the news about the shutdown.

The results are presented in the last column of Table 17. While the baseline effect remains negative, it decreases substantially in magnitude. At the same time, early exits among peers have a very strong and negative effect on the choice of total online activity post-announcement. For an individual with a single friend who never logs in again after the news, this effect sums up to approximately 2.693 hours decrease in online activity post-announcement or almost 31% decrease compared to the mean sample value. Early exits of friends are extremely influential in the decision about online activity for those who decide to log in at least once after the announcement, indicating strong peer effects that are likely to partially drive the negative results of the network size.

4.4.2 Robustness results

One of the concern regarding usage of the network data is the potential endogeneity of the network characteristics. Despite the fact that the announcement about the shutdown was generally unexpected, one can argue that costreaming channels are different than non-costreaming ones. To remedy that, I estimate the baseline regression while

Table 18: Instrumental variable estimation results

	first stage		IV baseline		IV instances	
	estimate	std error	estimate	std error	estimate	std error
count of costreamers			-2.833***	(1.005)	-0.579*	(0.336)
spanish language	-0.223***	(0.02)				
Hausman p-value		-		0.109		0.217
R squared		0.085		-		-
Observations		218660		218660		218660

Note: table above presents partial results for regressing count of online activity on the platform on the costreamer count using instrumental variable approach to account for potential endogeneity in the network measure. Column 'IV baseline' presents results for regressing total hours online post-announcement similar to the baseline model presented in the paper. Instrument used for endogenous network characteritic is whether the channel streamed in Spanish language at any point during the preannouncement period. Column 'IV instances' runs a similar instrumental variable regression, but using login instances instead of total hours online as the dependant variable. Finally, column 'first stage' presents the first stage regression for costreamer count. Note that F statistic for significance of the instrument in the first stage is equal to 122.35.
Full results of the estimation can be found in Table 22 in the Appendix.

instrumenting count of costreamers with Spanish language indicator.

The results for the instrumental variable approach to estimating the baseline regression are available in Table 18. While the second stage estimates have the same sign as the baseline OLS, they are almost three times as large in magnitude for the IV regression: one unit increase in costreamer count results in a decrease in 2.83 hours of streaming and 0.58 login instances, respectively. Both estimation results qualitatively confirm the results of the baseline OLS regression, though with a much larger marginal effect.

First stage regression results confirm the intuition that Spanish language indicator variable is negatively correlated with the costreamer count given smaller community of streamers on the platform, especially when compared to most common English language channels. However, the null hypothesis of the Hausman test cannot be rejected in neither of the cases. Given the uncertainty about exogeneity of the instruments discussed in the previous section as well as the results for the Hausman test, I continue the main estimation with the OLS.[33]

[33]Note that adding additional language indicator variables for Portuguese, German or French does not qualitatively change the results, but may cause large switches in Hausman and Sargan test statistics. This is an additional reason why I restrict instruments to the first most-common non-English language indicator and treat the results as a qualitative robustness to the baseline approach.

It is possible that the choice of games increases costreaming capabilities and is also correlated with post-announcement streaming time, creating an omitted variable bias. Although I am able to include this information, there is no certainty about causality. It is unclear whether the choice of games is a result of a consensus with friends about what to play or perhaps because of a given choice of games channel i begins costreaming with channel j. For that reason, I do not include game characteristics among control variables in the baseline results but instead run this specification as a robustness test.

Estimates for the specification with additional game characteristic controls can be found in the first two columns of Table 24 in the Appendix. I include total count of hours streaming the top 15 costreaming games on the platform as well as several other general features such as average viewership of games played, count of games played or average number of streamers playing the games that i played. Furthermore, I run this robustness check for two dependent variables: hours online as well as the indicator for early leavers. Note that the sum of hours playing the selected games account for almost 65% of the total streaming time in the sample, indicating that these few most popular titles account for the majority of the observed activity.

In both cases, the coefficient estimate on costreamer count decreases in comparison to the baseline results, but remains negative and strongly significant. Each extra costreamer presample results in approximately 40 minutes less streaming time after the announcement about the shutdown was made. This is still a substantial change of almost 8% decrease compared to the mean value in the sample. For the linear probability model with additional game controls, the probability of leaving early decreases by 0.002 units for each extra costreamer.

Costreaming in the described format can be considered quite unique to the platform of interest. That implies that some individuals may stay longer on the platform simply because they enjoy the feature and not because of the amount of costreaming links made with others. Furthermore, individuals with recent viewership boosts or higher exposure may feel more attached to the platform and leave later. I consider that possibility by including additional control variables for total count of hours costreaming, average viewership and count of hours featured prior to the announcement. Since causality is again not clear here - individual may be costreaming more simply because they have many potential costreaming friends or may have higher viewership because of costreaming partners - I consider this regression as an additional robustness check for the main results of this paper.

Results of the estimation are available in first column of Table 23 in the Appendix. As expected, individuals who costream more in absolute terms are also spending more hours online after the announcement. While there is evidence for the intuition that a degree of uniqueness of costreaming keeps users on the platform, count of costreaming links has a qualitatively similar effect as in the baseline results. Furthermore, since users are also influenced by their friends leaving, it seems that it's not only a matter of costreaming, but also connections with specific friends before the announcement.

I also run a similar robustness check for the linear probability model, estimating the effect on leaving early. The estimates can be found in the second column of Table 23 in the Appendix. While the main estimate for count of costreamers substantially decreases in magnitude, it remains significant at 10%. At the same time, count of costreaming hours has negative effect on leaving early. Such results may again suggest that it is not only the count of links made, but also the feature itself that keeps users on the platform.

Since the announcement about the shutdown coincided with the holiday season, many users may have simply gone on their planned vacation. With a group of friends that always plays together, such a situation can lead to temporal exits of entire groups that are unrelated to the closure of the platform but appear to be related to costreaming and coincide with the closure time. Results in the previous section showed that more clustered groups of costreamers generally stay longer, but this effect is quickly offset and turns negative for larger groups.

To investigate whether the effect remains similar for individual who are not a part of a single group of friends, I consider a subsample of users who did not leave the platform prior to the announcement, have at least 3 costreaming friends and their local clustering coefficient is equal to 0 which implies that their costreaming friends did not costream with each other prior to the news about shutdown. I then regress this subsample in a similar fashion as in the last column of Table 17 to check if they are influenced by early leavers in their relevant group of fully disconnected costreamers.

I find a relatively smaller, but still negative and significant effect for the influence of early leavers count, but the effect for costreamer count loses significance. While this may be due to substantially lower sample size that reduced to just 2.2% of the initial sample size for the regression in the last column of Table 17, the results for the effects of peers remain similar even when clustered groups of friends are excluded from the sample, supporting the social capital hypothesis. Early exits of costreaming friends decrease one's investment of time in the closing platform even when these friends are not part of a single group and do not depend on each other in a sense where single exit triggers all others to cease their activities.

Finally, although I imposed a restriction on the sample in terms of number of followers and minimum streaming time prior to the announcement, these channels can still be considered unprofessional streamers. In fact, most of them seem to stream casually, with median channel appearing approximately every third day for little less than 2 hours. One can consider that the effect is much more interesting for individuals who potentially make a living out of streaming. To investigate whether the effect differs for such a group of users, I create a subsample of channels that have at least 1000 followers. In contrast to the median channel in the baseline sample, they stream on average 161 hours over the presample period which indicates at least 3 hours of streaming every day.

The results of regressing hours online and indicator for early leavers using the subsample of channels with at least 1000 followers can be found in last two columns of Table 24 in the Appendix. Note that there is a sharp reduction in the sample size to just 5165 individuals which is 2.3% of the initial sample size. However, in both cases I still find negative and significant effect of costreamer count. For each additional costreaming link, streamers with at least 1000 followers are 0.004 units less likely to leave prior to the news about the shutdown. In case of hours online after the announcement, each additional costreamer decreases the dependent variable by approximately 43 minutes. While this is just an approximate 2.7% decrease for each extra costreamer (channels in the subsample are on average streaming approximately 27 hours after the announcement), the results consistently suggest that the effect is present even among professional and established users of the platform.

4.5 Discussion

Costreaming is a fairly recent phenomenon originating from video game streaming industry where up to four distinct channels can join in a split-screen gaming broadcast. This form of activity is not only an official on-platform feature that allows co-producing content for the viewers, but also an indicator of social relations among the channels involved. Consequently, count of costreamers is likely to be an important factor determining how much time to invest in a platform that is ceasing its activities. Individuals that have a group of costreaming friends may actually be more attached to the platform and the community, finding it harder to leave immediately. At the same time, if they rely on others with regards to streaming, they can also leave earlier once they observe their friends leaving.

This paper studies the effect of individual costreaming network size on the decision about the amount of online activity on the video game platform. Channels that have more distinct costreaming friends recorded prior to the announcement are also more likely to cut their usage of the platform once it becomes known that it will shut down within a period of month. This is a substantial decrease of 11% in hours online and 4.2% decrease in login instances for each extra costreamer as compared to the mean values of these variables post-announcement. The results are further qualitatively robust to instrumenting potentially endogenous costreamer count with Spanish channel language, controlling for game choice or even the total costreaming time prior to the announcement.

The source of the net negative effect of costreamer count on online activity appears to be non-linear. Individuals have 0.005 higher probability of appearing online after the announcement for each distinct costreaming friend they have, indicating that churning prior to the announcement is far less likely among costreamers, but at the same time they spend less total time on the platform. Furthermore, the effect is stronger in clustered groups of costreamers, suggesting peer effects. Modeling the decision on activity online for the subsample individuals who decide to appear at least once after the announcement shows that much of the negative effect seems to come from the fraction of their peers that have appeared last time online prior to the announcement. This suggests that individuals with more developed networks are more attached to the platform, but also far more likely to leave if their friends are leaving.

At the same time, this paper did not attempt to explicitly model peer effects or establish causality in the decision to exit despite having high frequency data on online activities. With lack of information on communication between peers, it is likely that two consequent exits would be treated as one causing the other whereas they can in fact be a coordinated decision to exit, with two friends having different streaming schedules. Subsample results that consider individuals who appear at least once after the announcement and condition their commitment to the platform on early leavers among their costreaming peers can fall into a similar problem if some of their peers simply did not log in because of their streaming schedules but still agreed to exit immediately with individuals who did log in.

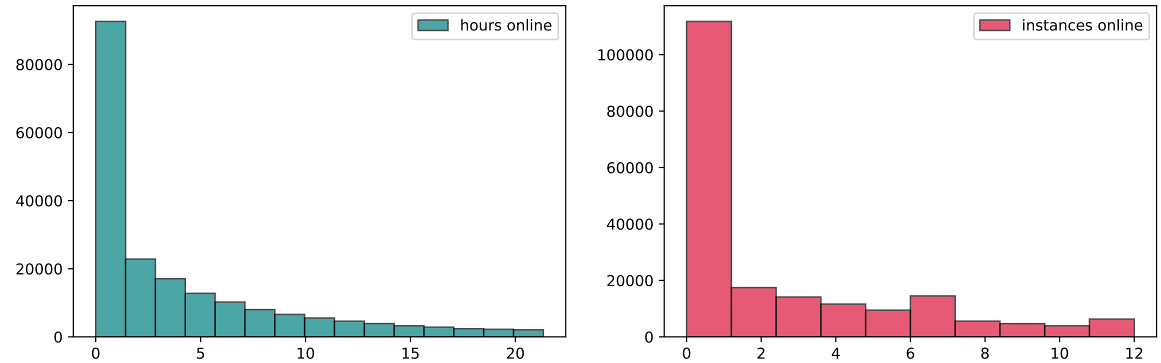
One of the potential solutions to that problem would be to define early leavers as individuals who left at least a week before the announcement and limit the probability of coordination. Third column of Table 23 in the Appendix presents results for early leavers defined in this fashion while still considering a subsample of individuals who logged in at least once after the announcement. As expected, the coefficient on early leavers lowers in magnitude, but remains significant and negative. However, it is impossible to determine a fixed cutoff value that would confidently exclude any coordination from the sample. Furthermore, any such fixed value also removes some part of the sample that did not coordinate despite being close to the announcement date. Future studies could try and decompose the coordination from peer effects and establish clear causation in the exit decisions.

Online platforms can freely design their sites and add unique features. Costreaming, in the format available on the platform under study, allows channels to join their feed and co-produce the content they show to their aggregated viewership. The results of this study show that adding such features seem to generally keep users on the platform: each extra distinct costreaming link results in a lower probability of churning prior to the announcement about the shutdown, and the results is robust even among most popular streamers. Yet, with peer effects playing a role, the very same links can quickly reverse the effect, as seen in the results of this study where net effect on total activity post announcement is in fact negative. Lőrincz et al. (2019) indicated positive effects of network size on staying on the platform while Török and Kertész (2017) argue that the users begin to churn once a sufficient amount of their friends have left the platform. It seems that similar mechanisms can be observed in the current study. This indicates that the net effects on the popularity of the platform of allowing for networking can greatly depend on the dynamics of growth or collapse and hence should be carefully assessed. At the same time it has to be highlighted that this negative net effect appears in a very extreme scenario when closure is known and set and arguably in most cases introducing additional means of networking should have positive effects.

While common intuition would suggest that networks strengthen the resilience of the platform's community, net effect of a larger individual network of costreamers is lower activity that can be translated into an earlier factual exit from the platform. In this particular case it may be explained by both cooperative or even dependent nature of costreaming networks as well as the common knowledge about platform collapse where peer effects actually speed up the exit process rather than preventing it. In consequence, introducing features that allow for networking and cooperation in a video game streaming platform can have substantial effect on the dynamics of user exit.

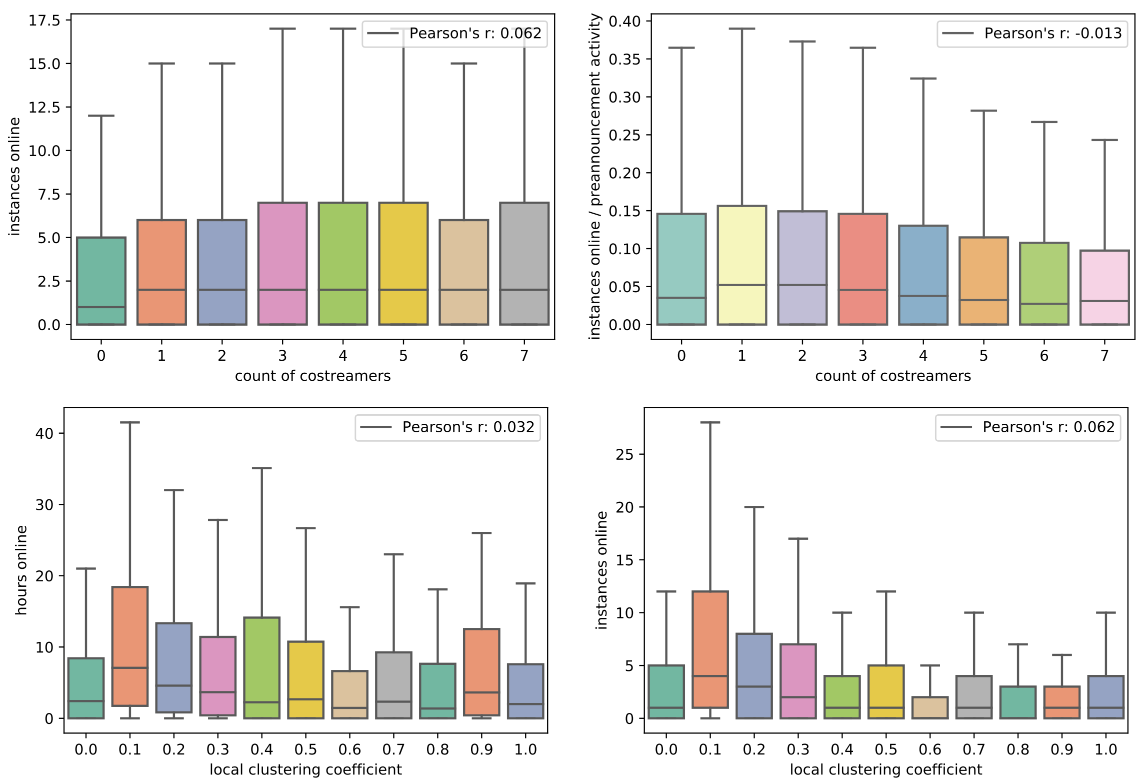
4.6 Appendix

Figure 22: Count of hours online and logins after the announcement



Note: histograms above plot distribution of total hours online and count of login instances after the announcement about the shutdown was made. Note that the maximum count of hours and logins are trimmed at 90th quantile for the clarity of the exposition.

Figure 23: Count of logins, costreamer count and local clustering coefficient



Note: boxplots in the top row present distributions of login counts after the announcement conditional on the count of costreamers. Bottom row presents similar plots for both hours and instances but against the local clustering coefficient. See a detailed description of similar plots in Figure 20 and in text.

Table 19: Summary statistics for the costreaming networks

	average	std dev	min	Q25	Q50	Q75	max
hours online	8.71	22.54	0.00	0.00	2.42	8.67	685.33
instances online	4.01	6.39	0.00	0.00	1.00	5.00	183.00
count of costreamers	0.82	1.62	0.00	0.00	0.00	1.00	70.00
local clustering coefficient	0.08	0.24	0.00	0.00	0.00	0.00	1.00
count early leavers	0.31	0.75	0.00	0.00	0.00	0.00	38.00
count left week prior	0.19	0.53	0.00	0.00	0.00	0.00	33.00
sum of hours costreaming	3.38	15.55	0.00	0.00	0.00	1.33	695.08
count of costreaming sessions	1.96	6.65	0.00	0.00	0.00	2.00	319.00
preannouncement instances online	22.01	14.82	7.00	12.00	18.00	27.00	470.00
preannouncement hours online	48.89	60.00	14.00	19.33	28.50	51.33	1231.58
preannouncement hours costreaming	10.81	21.21	0.00	0.00	1.92	15.25	1048.25
preannouncement hours featured	0.00	0.84	0.00	0.00	0.00	0.00	239.08
preannouncement average viewership	1.64	21.46	0.00	0.08	0.30	0.95	5086.23
exists # years	1.97	0.91	0.02	1.23	2.04	2.74	5.39
VODs enabled	0.31	0.46	0.00	0.00	0.00	1.00	1.00
channel is partnered	0.00	0.07	0.00	0.00	0.00	0.00	1.00
count of followers (in 1000)	0.22	7.76	0.00	0.00	0.01	0.02	3062.64
all-time unique viewers (in 1000)	3.69	196.80	0.00	0.03	0.09	0.34	60512.41
streamer count games (in 1000)	2.39	2.38	0.00	0.39	1.31	4.45	16.06
viewership games (in 1000)	3.71	2.85	0.00	0.92	3.44	6.21	17.56
count games played	5.52	4.45	0.00	3.00	4.00	7.00	110.00
hours playing Fortnite	11.92	24.93	0.00	0.00	1.17	16.08	909.33
hours playing Call of Duty	5.12	20.06	0.00	0.00	0.00	1.25	784.25
hours playing Apex Legends	3.78	18.44	0.00	0.00	0.00	0.00	745.67
hours playing Minecraft	1.56	7.58	0.00	0.00	0.00	0.25	683.75
hours playing Rainbow Six	1.09	6.71	0.00	0.00	0.00	0.00	506.75
hours playing GTA 5	1.42	9.23	0.00	0.00	0.00	0.00	591.17
hours playing PUBG	1.34	13.26	0.00	0.00	0.00	0.00	502.00
hours playing Dead by Daylight	0.67	6.95	0.00	0.00	0.00	0.00	547.50
hours playing Sea of Thieves	0.53	5.33	0.00	0.00	0.00	0.00	489.08
hours playing Roblox	0.89	4.28	0.00	0.00	0.00	0.00	364.42
hours playing Destiny	0.65	6.86	0.00	0.00	0.00	0.00	465.08
hours playing NBAK	1.19	9.22	0.00	0.00	0.00	0.00	836.92
hours playing Minecraft Dungeons	0.48	2.91	0.00	0.00	0.00	0.00	217.17
hours playing Rocket League	0.51	4.73	0.00	0.00	0.00	0.00	433.50
hours playing Forza Horizon	0.41	5.56	0.00	0.00	0.00	0.00	799.08
english language	0.90	0.30	0.00	1.00	1.00	1.00	1.00
spanish language	0.03	0.16	0.00	0.00	0.00	0.00	1.00

Note: table above shows the summary statistics of for 218660 channels in the sample that have at least 14 hours of streaming that include at least 7 login instances in the presample period. I also selected channels that have at least 1 followers at the first time they are observed in the sample. These restriction allows for meaningful computation of the network characteristics across sample while ensuring that channels under consideration are able to costream. It also allows to select channels that do use the platform regularly and have at least few followers. Q25, Q50, and Q75 denote 25th, 50th and 75th quantiles of the observations, respectively.

Table 20: OLS baseline results

	OLS reduced		OLS baseline		OLS instances	
	estimate	std error	estimate	std error	estimate	std error
count of costreamers	0.83***	(0.046)	-1.008***	(0.057)	-0.17***	(0.01)
constant	8.025***	(0.059)	2.681***	(0.204)	0.804***	(0.049)
preannouncement hours online			0.202***	(0.007)	-0.004***	(0.001)
preannouncement instances online			-0.049***	(0.017)	0.207***	(0.002)
exists # years			-0.403***	(0.056)	-0.243***	(0.014)
channel is partnered			13.203***	(3.225)	-0.584***	(0.205)
count of followers (in 1000)			-0.138	(0.111)	-0.015	(0.01)
all-time unique viewers (in 1000)			0.006	(0.005)	0.001*	(0.0)
VODs enabled			-3.818***	(0.186)	-1.66***	(0.029)
R squared		0.004		0.247		0.208
Observations		218660		218660		218660

Note: table above presents full results for regressing aggregate post-announcement online activity on the platform on the individual count of costreamers. Description and relevant explanations can be found in Table 16 and in text.

Table 21: Potential mechanisms behind the baseline results

	exits before		with clustering		no early exits	
	estimate	std error	estimate	std error	estimate	std error
count of costreamers	-0.005***	(0.001)	-0.569***	(0.067)	-0.599***	(0.088)
local clustering coefficient			4.278***	(0.632)		
costreamers * clustering			-2.384***	(0.284)		
count early leavers					-2.094***	(0.162)
constant	0.358***	(0.003)	2.384***	(0.198)	5.516***	(0.26)
preannouncement hours online	0.0***	(0.0)	0.196***	(0.007)	0.226***	(0.008)
preannouncement instances online	-0.007***	(0.0)	-0.04**	(0.017)	-0.111***	(0.019)
exists # years	0.002**	(0.001)	-0.343***	(0.056)	-0.545***	(0.076)
channel is partnered	-0.204***	(0.014)	13.716***	(3.251)	9.439**	(3.675)
count of followers (in 1000)	0.0	(0.0)	-0.138	(0.114)	-0.004	(0.129)
all-time unique viewers (in 1000)	-0.0	(0.0)	0.007	(0.005)	0.005	(0.004)
VODs enabled	0.177***	(0.002)	-3.55***	(0.177)	-1.998***	(0.234)
R squared		0.074		0.241		0.275
Observations		218164		218164		159278

Note: these are full robustness results for potential mechanisms behind the effects estimated in the baseline model. Description and relevant explanations can be found in Table 17 and in text.

Table 22: Instrumental variable estimation approach results

	first stage		IV baseline		IV instances	
	estimate	std error	estimate	std error	estimate	std error
count of costreamers			-2.833***	(1.005)	-0.579*	(0.336)
constant	0.322***	(0.01)	3.255***	(0.373)	0.933***	(0.116)
preannouncement hours online	0.005***	(0.0)	0.211***	(0.009)	-0.002	(0.002)
preannouncement instances online	0.014***	(0.0)	-0.023	(0.022)	0.213***	(0.005)
exists # years	-0.036***	(0.004)	-0.467***	(0.066)	-0.257***	(0.018)
channel is partnered	-0.221***	(0.053)	12.797***	(3.009)	-0.675***	(0.195)
count of followers (in 1000)	0.002**	(0.001)	-0.134**	(0.063)	-0.014***	(0.005)
all-time unique viewers (in 1000)	-0.0***	(0.0)	0.006**	(0.003)	0.001**	(0.0)
VODs enabled	0.052***	(0.008)	-3.729***	(0.189)	-1.64***	(0.033)
spanish language	-0.223***	(0.02)				
Hausman p-value		-		0.109		0.217
R squared		0.085		-		-
Observations		218660		218660		218660

Note: table above presents full results for regressing count of online activity on the platform on the costreamer count using instrumental variable approach to account for the potential endogeneity in the network measure.
Full description of the results can be found in Table 18 and in the related text.

Table 23: Robustness results for further controls, churners and disconnected friends

	hours online: expanded		exits before: expanded		longer delay		unclustered delay	
	estimate	std error	estimate	std error	estimate	std error	estimate	std error
count of costreamers	-1.083***	(0.063)	-0.001*	(0.001)	-1.133***	(0.08)	-0.102	(0.506)
count left week prior					-0.867***	(0.172)		
count early leavers							-1.676***	(0.412)
preannouncement hours costreaming	0.023***	(0.008)	-0.001***	(0.0)				
preannouncement average viewership	-0.011	(0.021)	0.0	(0.0)				
preannouncement hours featured	-0.151***	(0.044)	-0.0	(0.001)				
constant	2.449***	(0.211)	0.361***	(0.003)	5.6***	(0.261)	6.255***	(1.842)
preannouncement hours online	0.194***	(0.007)	0.0***	(0.0)	0.226***	(0.008)	0.139***	(0.019)
preannouncement instances online	-0.038**	(0.017)	-0.007***	(0.0)	-0.109***	(0.019)	0.052	(0.034)
exists # years	-0.362***	(0.056)	0.002*	(0.001)	-0.58***	(0.076)	0.143	(0.444)
channel is partnered	14.821***	(3.457)	-0.215***	(0.016)	9.567***	(3.675)	-30.214	(18.373)
count of followers (in 1000)	-0.127	(0.139)	0.0	(0.001)	-0.003	(0.129)	-5.677*	(3.245)
all-time unique viewers (in 1000)	0.007	(0.006)	-0.0	(0.0)	0.005	(0.004)	0.571**	(0.289)
VODs enabled	-3.498***	(0.196)	0.171***	(0.002)	-2.102***	(0.234)	-4.733***	(1.645)
R squared		0.24		0.075		0.274		0.204
Observations		218164		218164		159278		3558

Note: table above presents additional robustness results. Columns 'hours online: expanded' and 'exits before: expanded' add control variables aggregated over the preannouncement period: costreaming time, average viewership and periods featured, to the baseline models that regress hours online after the announcement and indicator of leaving before the announcement was made. Column 'longer delay' estimates the baseline model for a subsample of individuals who appeared at least once online after the announcement, regressing it on the count of costreaming friends who left at least a week prior to the announcement. Finally, column 'unclustered delay' estimates a similar model as column 'no early exits' in Tables 2 and 6, but for a subsample of individuals who have at least 3 costreamers and local clustering coefficient equal to 0.

Table 24: Robustness results for game choice controls and popularity subsampling

	hours online: games		exits before: games		hours online: popular		exits before: popular	
	estimate	std error	estimate	std error	estimate	std error	estimate	std error
count of costreamers	-0.674***	(0.052)	-0.002***	(0.001)	-0.716***	(0.207)	-0.004***	(0.001)
constant	1.04***	(0.22)	0.425***	(0.004)	9.164*	(5.436)	0.39***	(0.032)
preannouncement hours online	0.242***	(0.011)	0.0***	(0.0)	0.216***	(0.025)	-0.0***	(0.0)
preannouncement instances online	-0.008	(0.014)	-0.006***	(0.0)	-0.423***	(0.116)	-0.004***	(0.0)
exists # years	-0.165***	(0.056)	-0.005***	(0.001)	2.321*	(1.225)	0.031***	(0.007)
channel is partnered	14.94***	(3.349)	-0.184***	(0.014)	14.51***	(3.384)	-0.14***	(0.014)
count of followers (in 1000)	-0.115	(0.114)	0.0	(0.0)	-0.133	(0.117)	0.0	(0.0)
all-time unique viewers (in 1000)	0.005	(0.004)	-0.0	(0.0)	0.007	(0.005)	0.0**	(0.0)
VODs enabled	-2.748***	(0.155)	0.139***	(0.002)	-10.681**	(4.34)	0.031	(0.026)
count games played	-0.14***	(0.023)	-0.006***	(0.0)				
streamer count games (in 1000)	-0.077	(0.073)	-0.045***	(0.001)				
viewership games (in 1000)	0.32***	(0.071)	0.022***	(0.001)				
hours playing Fortnite	-0.068***	(0.012)	0.0***	(0.0)				
hours playing Call of Duty	-0.139***	(0.013)	-0.0***	(0.0)				
hours playing Apex Legends	-0.135***	(0.011)	-0.0***	(0.0)				
hours playing Minecraft	-0.021	(0.024)	-0.001***	(0.0)				
hours playing Rainbow Six	-0.113***	(0.016)	0.0	(0.0)				
hours playing GTA 5	-0.043**	(0.019)	-0.0***	(0.0)				
hours playing PUBG	-0.09***	(0.013)	-0.0***	(0.0)				
hours playing Dead by Daylight	-0.044	(0.029)	-0.001***	(0.0)				
hours playing Sea of Thieves	-0.177***	(0.019)	0.001***	(0.0)				
hours playing Roblox	0.122***	(0.026)	-0.003***	(0.0)				
hours playing Destiny	-0.103***	(0.02)	-0.0	(0.0)				
hours playing NBAK	-0.008	(0.018)	-0.001***	(0.0)				
hours playing Minecraft Dungeons	-0.104	(0.07)	0.002***	(0.0)				
hours playing Rocket League	-0.106***	(0.022)	0.0**	(0.0)				
hours playing Forza Horizon	0.051	(0.045)	-0.0***	(0.0)				
R squared		0.26		0.089		0.164		0.112
Observations		218164		218164		5165		5165

Note: table above presents robustness results for the baseline models that take into account characteristics of the games played as well as the popularity of the channels. Columns 'hours online: games' and 'exits before: games' add count of hours playing top 15 most popular games on the platform to the baseline models that regress hours online after the announcement and indicator of leaving before the announcement was made. Choice of the games played can influence the costreaming probability of the individuals given in-game cooperation posibilities. Columnn 'hours online: popular' and 'exits before: popular' estimates the same two baseline models while reducing the sample to only channels that have at least 1000 followers. Note that sample size in this analysis is reduced to 2.36% of the initial number of observations.

5 Bibliography

Advani, A., and Malde, B. (2018). Credibly identifying social effects: Accounting for network formation and measurement error. *Journal of Economic Surveys, 32*(4), 1016-1044.

Alatas, V., Banerjee, A., Chandrasekhar, A. G., Hanna, R., and Olken, B. A. (2016). Network structure and the aggregation of information: Theory and evidence from Indonesia. *American Economic Review, 106*(7), 1663-1704.

Alexander, R. D. (1987). *The biology of moral systems.* Transaction Publishers.

Banerjee, A., Chandrasekhar, A. G., Duflo, E., and Jackson, M. O. (2013). The diffusion of microfinance. *Science, 341*(6144).

Bes, R. E., Curfs, E. C., Groenewegen, P. P., and De Jong, J. D. (2017). Health plan choice in the Netherlands: restrictive health plans preferred by young and healthy individuals. *Health Economics, Policy and Law, 12*(3), 345-362.

Berg, J., Dickhaut, J., and McCabe, K. (1995). Trust, reciprocity, and social history. *Games and economic behavior, 10*(1), 122-142.

Beukers, P. D., Kemp, R. G., and Varkevisser, M. (2014). Patient hospital choice for hip replacement: empirical evidence from the Netherlands. *The European Journal of Health Economics, 15*(9), 927-936.

Brown, Z. Y. (2019). Equilibrium Effects of Health Care Price Information. *The Review of Economics and Statistics, 101*(4), 699–712. Available at: <http://www-personal.umich.edu/~zachb/zbrown_eqm_effects_price_transparency.pdf> [Accessed 2 December 2019]

Capraro, V., Jagfeld, G., Klein, R., Mul, M., and Van De Pol, I. (2019). Increasing altruistic and cooperative behaviour with simple moral nudges. *Scientific reports, 9*(1), 1-11.

Chamberlain, G. (1980). Analysis of Covariance with Qualitative Data. *The Review of Economic Studies, 47*(1), 225-238.

Chernew, M., Cooper, Z., Larsen-Hallock, E., and Morton, F. S. (2018). Are Health Care Services Shoppable? Evidence from the Consumption of Lower-Limb MRI Scans.

[online] *National Bureau of Economic Research No. w24869.* Available at: <https://www.nber.org/papers/w24869> [Accessed 3 June 2019]

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal, 21*(1), C1-C68.

Christensen, H. B., Floyd, E. and Maffett, M. G. (2018). The Effects of Charge-Price Transparency Regulation on Prices in the Healthcare Industry. [online] *Chicago Booth Research Paper No. 14-33.* Available at SSRN: <https://ssrn.com/abstract=2343367> [Accessed 23 August 2018].

Dale, S. (2014). Content curation: The future of relevance. *Business Information Review, 31*(4), 199-205.

Dasgupta, K., Singh, R., Viswanathan, B., Chakraborty, D., Mukherjea, S., Nanavati, A. A., and Joshi, A. (2008). Social ties and their relevance to churn in mobile telecom networks. *In Proceedings of the 11th international conference on Extending database technology: Advances in database technology,* 668-677.

De Jong, S. (2016). Database met 200.000 ziekenhuistarieven online. *Consumentenbond.* [online] Available at: <https://www.consumentenbond.nl/nieuws/2016/database-met-200.000-ziekenhuistarieven-online> [Accessed 2 December 2019]

De Paula, Á. (2017). Econometrics of Network Models. In: B. Honoré, A. Pakes, M. Piazzesi, and L. Samuelson (Eds.), *Advances in Economics and Econometrics: Eleventh World Congress* (Econometric Society Monographs). Cambridge: Cambridge University Press, 268-323.

Deng, J., Cuadrado, F., Tyson, G., and Uhlig, S. (2015). Behind the game: Exploring the twitch streaming platform. *2015 International Workshop on Network and Systems Support for Games (NetGames)*, IEEE, 1-6.

Desai, S., Hatfield, L. A., Hicks, A. L., Chernew, M. E., and Mehrotra, A. (2016). Association between availability of a price transparency tool and outpatient spending. *Jama, 315*(17), 1874-1881.

DIS Open Data (2019). Open data van de Nederlandse Zorgautoriteit. [online] Available at: <http://www.opendisdata.nl/> [Accessed 1 December 2019]

Douven, R., Burger, M., and Schut, E. (2018). Does managed competition constrain hospitals' contract prices? Evidence from the Netherlands (No. 378). *CPB Netherlands Bureau for Economic Policy Analysis.* [online] Available at: <https://www.cpb.nl/sites/default/files/omnidownload/CPB-Discussion-Paper-378-Does-managed-competition-constrain-hospitals-contract-prices.pdf> [Accessed 25 March 2021]

Dufwenberg, M., Gneezy, U., Güth, W., and Van Damme, E. (2001). Direct vs indirect reciprocity: an experiment. *Homo Oecon, 18,* 19-30.

Elliott, R. (2020). Streamlabs & Newzoo Q4 Year in Review Live Streaming Industry Report. [online] Newzoo. Available at: <https://newzoo.com/insights/articles/streamlabs-newzoo-q4-year-in-review-live-streaming-industry-report/> [Accessed 8th of March 2021].

Enthoven, A. C., and van de Ven, W. P. (2007). Going Dutch—managed-competition health insurance in the Netherlands. *New England Journal of Medicine, 357*(24), 2421-2423.

Engelmann, D., and Fischbacher, U. (2009). Indirect reciprocity and strategic reputation building in an experimental helping game. *Games and Economic Behavior, 67*(2), 399-407.

Falk, A. (2007). Gift exchange in the field. *Econometrica, 75*(5), 1501-1511.

Farronato, C., Fong, J., and Fradkin, A. (2020). Dog eat dog: Measuring network effects using a digital platform merger [online] (No. w28047). *National Bureau of Economic Research.* Available at: <https://www.nber.org/papers/w28047> [Accessed 2nd of October 2021]

Fehr, E., and Gachter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review, 90*(4), 980-994.

Garcia, D., Mavrodiev, P., and Schweitzer, F. (2013). Social resilience in online communities: The autopsy of friendster. *In Proceedings of the first ACM conference on Online social networks*, 39-50.

Gourevitch, R. A., Desai, S., Hicks, A. L., Hatfield, L. A., Chernew, M. E., and Mehrotra, A. (2017). Who Uses a Price Transparency Tool? Implications for Increasing

Consumer Engagement. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing, 54*, 1–5.

Handel, B. R., and Kolstad, J. T. (2015). Health insurance for "humans": Information frictions, plan choice, and consumer welfare. *American Economic Review, 105*(8), 2449-2500.

Herrman, J. (2018). With Twitch, Amazon Tightens Grip on Live Streams of Video Games. [online] New York Times. Available at: <https://www.nytimes.com/2018/06/17/business/media/amazon-twitch-video-games.html> [Accessed 9th of March 2020].

Hibbard, J. H., Greene, J., Sofaer, S., Firminger, K., and Hirsh, J. (2012). An experiment shows that a well-designed report on costs and quality can help consumers choose high-value health care. *Health affairs, 31*(3), 560-568.

Jackson, M. O. (2011). *An overview of social networks and economic applications.* Handbook of social economics, 1, 511-585.

Jackson, M. O. (2014). Networks in the understanding of economic behaviors. *Journal of Economic Perspectives, 28*(4), 3-22.

Jackson, M. O. (2020). A typology of social capital and associated network measures. *Social Choice and Welfare, 54*(2), 311-336.

Johnson, M. R., and Woodcock, J. (2019). 'It's like the gold rush': the lives and careers of professional video game streamers on Twitch. tv. *Information, Communication & Society, 22*(3), 336-351.

Katona, Z., Zubcsek, P. P., and Sarvary, M. (2011). Network effects and personal influences: The diffusion of an online social network. *Journal of marketing research, 48*(3), 425-443.

Kim, K., Jun, C. H., and Lee, J. (2014). Improved churn prediction in telecommunication industry by analyzing a large network. E*xpert Systems with Applications, 41*(15), 6575-6584.

Kizilcec, R. F., Bakshy, E., Eckles, D., and Burke, M. (2018). Social influence and reciprocity in online gift giving. *In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1-11.

Kleijne, I. (2016). Schippers wil openheid NZa over ziekenhuistarieven. *Medisch Contact.* [online] Available at: <https://www.medischcontact.nl/nieuws/laatste-nieuws/artikel/schippers-wil-openheid-nza-over-ziekenhuistarieven.htm> [Accessed 2 December 2019].

Koltai, J., Lőrincz, L., Wachs, J., and Takács, K. (2021). Do Diversity and Context Collapse Kill an Online Social Network? [online] *SocArXiv*. Available at: <https://doi.org/10.31235/osf.io/4vk2m> [Accessed 11th of June 2021]

Kroneman, M., Boerma, W., van den Berg, M., Groenewegen, P., de Jong, J., and van Ginneken, E. (2016). The Netherlands: health system review. *Health Systems in Transition, 18*(2), 1–239.

Kuijper, K. (2016). CZ publiceert behandeltarief per ziekenhuis. *Zorgwijzer.* [online] Available at: <https://www.zorgwijzer.nl/zorgverzekering-2017/cz-publiceert-behandeltarief-per-ziekenhuis> [Accessed 3 December 2019]

Lako, C. J., and Rosenau, P. (2009). Demand-driven care and hospital choice. Dutch health policy toward demand-driven care: results from a survey into hospital choice. *Health Care Analysis, 17*(1), 20-35.

Lieber, E. M. (2017). Does It Pay to Know Prices in Health Care? *American Economic Journal: Economic Policy, 9*(1), 154-179.

Lőrincz, L., Koltai, J., Győr, A. F., and Takács, K. (2019). Collapse of an online social network: Burning social capital to create it?. *Social Networks, 57*, 43-53.

MacKinnon, J. G., and White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of econometrics, 29*(3), 305-325.

May, E. (2020a) Streamlabs & Stream Hatchet Q2 2020 Live Streaming Industry Report. *Streamlabs.* [online] Available at: <https://streamlabs.com/content-hub/post/streamlabs-stream-hatchet-q2-2020-live-streaming-industry-report> [Accessed 9th of August 2021]

May, E. (2020b) Streamlabs & Stream Hatchet Q3 2020 Live Streaming Industry Report. *Streamlabs.* [online] Available at: <https://streamlabs.com/content-hub/post/str

eamlabs-and-stream-hatchet-q3-2020-live-streaming-industry-report> [Accessed 9th of August 2021]

May, E. (2021) Streamlabs & Stream Hatchet Q1 2021 Live Streaming Industry Report. *Streamlabs.* [online] Available at: <https://blog.streamlabs.com/streamlabs-stream-hatchet-q1-2021-live-streaming-industry-report-eaba2143f492> [Accessed 8th of June 2021]

Mehrotra, A., Dean, K. M., Sinaiko, A. D., and Sood, N. (2017). Americans support price shopping for health care, but few actually seek out price information. *Health Affairs, 36*(8), 1392-1400.

Mujcic, R., and Leibbrandt, A. (2018). Indirect reciprocity and prosocial behaviour: Evidence from a natural field experiment. *The Economic Journal, 128*(611), 1683-1699.

Nowak, M. A., and Sigmund, K. (1998a). Evolution of indirect reciprocity by image scoring. *Nature, 393*(6685), 573-577.

Nowak, M. A., and Sigmund, K. (1998b). The dynamics of indirect reciprocity. *Journal of theoretical Biology, 194*(4), 561-574.

Nowak, M. A., and Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature, 437*(7063), 1291-1298.

NZa (2015). Zeventien zorgpolissen nader bekeken. [online] Available at: <https://www.zorg-nv.com/downloads/8917-Rapport_zeventien_zorgpolissen_nader_bekeken.pdf> [Accessed 16 August 2021]

NZa (2017). Marktscan Zorgverzekeringsmarkt 2017. [online] Available at:<https://puc.overheid.nl/nza/doc/PUC_3656_22/1/> [Accessed 23 May 2021]

Open State Foundation (2014). NZA still declines to disclose healthcare costs. [online] Available at: <https://openstate.eu/en/2014/09/nza-still-declines-to-disclose-healthcare-costs/ >
[Accessed 14 August 2021].

Open State Foundation (2017). More than 715,000 hospital prices online. [online] Available at: <https://openstate.eu/en/2017/09/more-than-715000-hospital-prices-online/ > [Accessed 2 December 2019].

Open State Foundation (2019). Prijzen van bijna één miljoen zorgbehandelingen op de kaart te vergelijken. [online] Available at: <https://openstate.eu/nl/2019/02/prijzen-van-bijna-een-miljoen-zorgbehandelingen-op-de-kaart-te-vergelijken/> [Accessed 3 December 2019]

Óskarsdóttir, M., Bravo, C., Verbeke, W., Sarraute, C., Baesens, B., and Vanthienen, J. (2017). Social network analytics for churn prediction in telco: Model building, evaluation and network architecture. *Expert Systems with Applications, 85*, 204-220.

Patiëntenfederatie Nederland (2019). Hoe kiest u een ziekenhuis? [online] Available at: <https://www.patientenfederatie.nl/downloads/monitor/123-patientenmonitor-hoe-kies-je-een-ziekenhuis/file> [Accessed 16 August 2021]

Periáñez, Á., Saas, A., Guitart, A., and Magne, C. (2016). Churn prediction in mobile social games: Towards a complete assessment using survival ensembles. *IEEE International Conference on Data Science and Advanced Analytics (DSAA),* 564-573.

Rosenau, P. V., and Lako, C. J. (2008). An experiment with regulated competition and individual mandates for universal health care: the new Dutch health insurance system. *Journal of Health Politics, Policy and Law, 33*(6), 1031-1055.

Ruwaard, S., and Douven, R. (2014). Quality and hospital choice for cataract treatments: the winner takes most (No. 272. rdf). *CPB Netherlands Bureau for Economic Policy Analysis.* [online] Available at: <https://www.cpb.nl/sites/default/files/publicaties/download/cpb-discussion-paper-272-quality-and-hospital-choice-cataract-treatments-winner-takes-most.pdf> [Accessed 16 August 2021].

Saavedra, S., Reed-Tsochas, F., and Uzzi, B. (2008). Asymmetric disassembly and robustness in declining networks. *Proceedings of the National Academy of Sciences, 105*(43), 16466-16471.

Seinen, I., and Schram, A. (2006). Social status and group norms: Indirect reciprocity in a repeated helping experiment. *European Economic Review, 50*(3), 581-602.

Semigran, H. L., Gourevitch, R., Sinaiko, A. D., Cowling, D., and Mehrotra, A. (2017). Patients' views on price shopping and price transparency. *The American journal of managed care, 23*(6), 186-192.

Simonov, A., Ursu, R., and Zheng, C. (2020). Do Suspense and Surprise Drive Entertainment Demand? Evidence from Twitch.tv. [online] *Columbia Business School Research Paper Forthcoming, NYU Stern School of Business Forthcoming.* Available at: <https://ssrn.com/abstract=3711801> [Accessed 22 July 2021].

Sjöblom, M., and Hamari, J. (2017). Why do people watch others play video games? An empirical study on the motivations of Twitch users. *Computers in human behavior, 75*, 985-996.

Skipr (2016). Ook Zilveren Kruis maakt tarieven deels bekend. [online] Available at: <https://www.skipr.nl/nieuws/ook-zilveren-kruis-maakt-tarieven-deels-bekend/> [Accessed 26 March 2021]

Stephen, B. (2020) Mixer shuts down today. *The Verge.* [online] Available at: <https://www. theverge.com/2020/7/22/21334082/mixer-shut-down-twitch-facebook-gaming-partnership> [Accessed 9th of August 2021]

Stephenson, B. (2019). 7 Ways to Make Money Streaming Video Games on Twitch. *Lifewire.* [online] Available at: <https://www.lifewire.com/make-money-streaming-on-twitch-4144817> [Accessed 8th of March 2020].

Török, J., and Kertész, J. (2017). Cascading collapse of online social networks. *Scientific reports, 7*(1), 1-8.

Tu, H. T., and Lauer, J. R. (2009). Impact of health care price transparency on price variation: the New Hampshire experience. *Issue Brief Center for Studying Health System Change, 128*, 1-4.

Twitch (2014). Share Your Favorite Content with Host Mode. *Twitch.* [online] Available at: <https://blog.twitch.tv/en/2014/07/24/share-your-favorite-content-with-host-mode-b496aa6166d7/> [Accessed 20th of July 2021].

Twitch (2021). The Basics. *Twitch.* [online] Available at: <https://www.twitch.tv/creatorcamp/en/learn-the-basics/the-basics/> [Accessed 21th of July 2021].

Warren, T. (2020). Microsoft is shutting down Mixer and partnering with Facebook Gaming. [online] *The Verge.* Available at: <https://www.theverge.com/2020/6/22/21299032/microsoft-mixer-closing-facebook-gaming-partnership-xcloud-features> [Accessed 9th of August 2021]

Whaley, C., Chafen, J. S., Pinkard, S., Kellerman, G., Bravata, D., Kocher, R., and Sood, N. (2014). Association between availability of health service prices and payments for these services. *Jama, 312*(16), 1670-1676.

Woldring, J. (2016). Menzis en Anderzorg geven inzicht in ziekenhuistarieven. *Poliswijzer.* [online] Available at: <https://www.poliswijzer.nl/nieuws/menzis-en-anderzorg-geven-inzicht-in-ziekenhuistarieven> [Accessed 3 December 2019]

Wu, S. J., Sylwestrzak, G., Shah, C., and DeVries, A. (2014). Price transparency for MRIs increased use of less costly providers and triggered provider competition. *Health Affairs, 33*(8), 1391-1398.

Van Aartsen, C. (2017). Menzis spreekt vaste prijzen af met 17 ziekenhuizen. *Zorgvisie.* [online] Available at: <https://www.zorgvisie.nl/menzis-spreekt-vaste-prijzen-af-met-17-ziekenhuizen/> [Accessed 26 March 2021]

Van Apeldoorn, J., and Schram, A. (2016). Indirect reciprocity; a field experiment. *PloS one, 11*(4), e0152076.

Van Bokhorst, A. (2016). Ook VGZ maakt ziekenhuistarieven openbaar. *Independer.* [online] Available at: <https://weblog.independer.nl/zorgverzekeringen/ook-vgz-maakt-ziekenhuistarieven-openbaar/> [Accessed 3 December 2019].

Van Rooy, Y. (2017). Vereenvoudig het eigen risico. *skipr.* [online] Available at: <https:// www.skipr.nl/blog/vereenvoudig-het-eigen-risico/> [Accessed 22 March 2021]

Varkevisser, M., van der Geest, S. A., and Schut, F. T. (2012). Do patients choose hospitals with high quality ratings? Empirical evidence from the market for angioplasty in the Netherlands. *Journal of health economics, 31*(2), 371-378.

Verbeke, W., Martens, D., and Baesens, B. (2014). Social network analysis for customer churn prediction. *Applied Soft Computing, 14*, 431-446.

Volpp, K. G. (2016). Price transparency: not a panacea for high health care costs. *Jama, 315*(17), 1842-1843

Yoeli, E., Hoffman, M., Rand, D. G., and Nowak, M. A. (2013). Powering up with indirect reciprocity in a large-scale field experiment. *Proceedings of the National Academy of Sciences, 110*(Supplement 2), 10424-10429.

CENTER DISSERTATION SERIES

CentER for Economic Research, Tilburg University, the Netherlands

No.	Author	Title	ISBN	Published
617	Matjaz Maletic	Essays on international finance and empirical asset pricing	978 90 5668 618 5	January 2020
618	Zilong Niu	Essays on Asset Pricing and International Finance	978 90 5668 619 2	January 2020
619	Bjorn Lous	On free markets, income inequality, happiness and trust	978 90 5668 620 8	January 2020
620	Clemens Fiedler	Innovation in the Digital Age: Competition, Cooperation, and Standardization	978 90 5668 621 5	October 2020
621	Andreea Popescu	Essays in Asset Pricing and Auctions	978 90 5668 622 2	June 2020
622	Miranda Stienstra	The Determinants and Performance Implications of Alliance Partner Acquisition	978 90 5668 623 9	June 2020
623	Lei Lei	Essays on Labor and Family Economics in China	978 90 5668 624 6	May 2020
624	Farah Arshad	Performance Management Systems in Modern Organizations	978 90 5668 625 3	June 2020
625	Yi Zhang	Topics in Economics of Labor, Health, and Education	978 90 5668 626 0	June 2020
626	Emiel Jerphanion	Essays in Economic and Financial decisions of Households	978 90 5668 627 7	July 2020
627	Richard Heuver	Applications of liquidity risk discovery using financial market infrastructures transaction archives	978 90 5668 628 4	September 2020
628	Mohammad Nasir Nasiri	Essays on the Impact of Different Forms of Collaborative R&D on Innovation and Technological Change	978 90 5668 629 1	August 2020
629	Dorothee Hillrichs	On inequality and international trade	978 90 5668 630 7	September 2020
630	Roland van de Kerkhof	It’s about time: Managing implementation dynamics of condition-based maintenance	978 90 5668 631 4	October 2020

No.	Author	Title	ISBN	Published
631	Constant Pieters	Process Analysis for Marketing Research	978 90 5668 632 1	December 2020
632	Richard Jaimes	Essays in Macroeconomic Theory and Natural Resources	978 90 5668 633 8	November 2020
633	Olivier David Armand Zerbib	Asset pricing and impact investing with pro-environmental preferences	978 90 5668 634 5	November 2020
634	Laura Capera Romero	Essays on Competition, Regulation and Innovation in the Banking Industry	978 90 5668 635 2	December 2020
635	Elisabeth Beusch	Essays on the Self-Employed in the Netherlands and Europe	978 90 5668 636 9	December 2020
636	Sophie Zhou	Innovation and the Macroeconomy	978 90 5668 637 6	November 2020
637	Vincent Peters	Turning modularity upside down: Patient-centered Down syndrome care from a service modularity perspective	978 90 5668 638 3	December 2020
638	Pranav Desai	Essays in Corporate Finance and Innovation	978 90 5668 639 0	January 2021
639	Kristy Jansen	Essays on Institutional Investors, Asset Allocation Decisions, and Asset Prices	978 90 5668 640 6	January 2021
640	Riley Badenbroek	Interior Point Methods and Simulated Annealing for Nonsymmetric Conic Optimization	978 90 5668 641 3	February 2021
641	Stephanie Koornneef	It's about time: Essays on temporal anchoring devices	978 90 5668 642 0	February 2021
642	Vilma Chila	Knowledge Dynamics in Employee Entrepreneurship: Implications for parents and offspring	978 90 5668 643 7	March 2021
643	Minke Remmerswaal	Essays on Financial Incentives in the Dutch Healthcare System	978 90 5668 644 4	July 2021
644	Tse-Min Wang	Voluntary Contributions to Public Goods: A multi-disciplinary examination of prosocial behavior and its antecedents	978 90 5668 645 1	March 2021

No.	Author	Title	ISBN	Published
645	Manwei Liu	Interdependent individuals: how aggregation, observation, and persuasion affect economic behavior and judgment	978 90 5668 646 8	March 2021
646	Nick Bombaij	Effectiveness of Loyalty Programs	978 90 5668 647 5	April 2021
647	Xiaoyu Wang	Essays in Microeconomics Theory	978 90 5668 648 2	April 2021
648	Thijs Brouwer	Essays on Behavioral Responses to Dishonest and Anti-Social Decision-Making	978 90 5668 649 9	May 2021
649	Yadi Yang	Experiments on hold-up problem and delegation	978 90 5668 650 5	May 2021
650	Tao Han	Imperfect information in firm growth strategy: Three essays on M&A and FDI activities	978 90 5668 651 2	June 2021
651	Johan Bonekamp	Studies on labour supply, spending and saving before and after retirement	978 90 5668 652 9	June 2021
652	Hugo van Buggenum	Banks and Financial Markets in Microfounded Models of Money	978 90 5668 653 6	August 2021
653	Arthur Beddock	Asset Pricing with Heterogeneous Agents and Non-normal Return Distributions	978 90 5668 654 3	September 2021
654	Mirron Adriana Boomsma	On the transition to a sustainable economy: Field experimental evidence on behavioral interventions	978 90 5668 655 0	September 2021
655	Roweno Heijmans	On Environmental Externalities and Global Games	978 90 5668 656 7	August 2021
656	Lenka Fiala	Essays in the economics of education	978 90 5668 657 4	September 2021
657	Yuexin Li	Pricing Art: Returns, Trust, and Crises	978 90 5668 658 1	September 2021
658	Ernst Roos	Robust Approaches for Optimization Problems with Convex Uncertainty	978 90 5668 659 8	September 2021

No.	Author	Title	ISBN	Published
659	Joren Koëter	Essays on asset pricing, investor preferences and derivative markets	978 90 5668 660 4	September 2021
660	Ricardo Barahona	Investor Behavior and Financial Markets	978 90 5668 661 1	October 2021
661	Maciej Husiatyński	Three essays on Individual Behavior and New Technologies	978 90 5668 663 5	October 2021

Maciej Husiatyński (Warsaw, Poland, 1992) received his MA with Honours in Economics at the University of Glasgow. After graduating from a Research Master program in Economics, he began his PhD at the department of Econometrics and Operations Research at Tilburg University under the supervision of prof. Tobias J. Klein and prof. Misja C. Mikkers.

This dissertation consists of three essays on individual behavior and new technologies. The first essay presents evidence that despite wide availability of potential savings, publication of contracted hospital prices does not affect the short-run demand for health care. The second essay studies indirect reciprocity in a large market of video game streaming where individuals are more likely to transfer their viewers to each other if they received similar gifts in the recent past. Finally, the third essay shows that while social networks of co-producing content can decrease individual churn from an online platform, they also speed up the exit process once the shutdown of the platform is announced to the users.

ISBN: 978 90 5668 663 5
DOI: 10.26116/center-lis-2121

TILBURG SCHOOL OF ECONOMICS AND MANAGEMENT

CentER Dissertation Series