

Reaching for The Stars: Discounts and Review Tier Transitions in the Video Games Market

[\[Click for the Latest Version\]](#)

Dmitry Sorokin*

October 2020

Abstract

This paper documents that firms in online marketplaces use price promotions to facilitate transitions to better review tiers (similar to the number of stars on Amazon.com). I find that firms close to going one tier up are 4-9% more likely to discount. Two effects could be at play. First, a selection effect arises because customers who buy during a discount could be different from the regular ones, and could potentially leave more positive reviews. Second, a variance effect reflects the idea that positive reviews could help the firm move up a review tier, while negative reviews would keep the tier unchanged, minimizing the downside risk of giving a discount. In order to understand the relative contributions of these effects, I estimate a tractable structural model of demand and reviewing behavior. I find evidence of the importance of the variance effect: when the selection effect is controlled for, firms close to downgrading their review tier are 6% less likely to give a discount, consistent with preferring less variance. I also find that consumers are significantly more likely to leave reviews during a discount. The average selection effect turns out to be small. Additional findings include developing a new approach to estimating demand from data on product usage, estimates of the causal effect of reviews on sales and equilibrium discount elasticities in an important market not previously studied.

*New York University, dsorokin@nyu.edu. I am deeply grateful to Timothy Christensen, Alessandro Lizzeri, and Luís Cabral for their support and guidance. I benefited enormously from my conversations with Ryan Stevens, Jimena Galindo, Sam Kapon, Dmitry Sedov, Jonathan Elliott, and Timur Abbiasov. I would also like to thank Chris Conlon, Christopher Flinn, Elena Manresa, Maher Said, Mikkel Plagborg-Møller, Milena Almagro, Paula Onuchic, Sharon Traiberman, Stefan Bucher, Vasily Rusanov, and all other great people at NYU for their helpful comments. All mistakes are my own.

1 Introduction

In his book “Economics for the Common Good” Jean Tirole proclaims the economy we live in to be governed by “Economics of Attention” [2017, p.379]. A simple search for wine glasses on Amazon.com results in more than 5,000 options, leading to a choice problem that is virtually impossible to solve satisfyingly without guidance. Such guidance is now routinely provided by different types of recommendation systems. Many marketplaces or third-party websites allow consumers to review products or services they purchase and then aggregate these reviews into a rating, greatly facilitating the comparison between alternatives. These ratings are often given a simple graphical representation; for example, a number of stars between 1 and 5 is a common way to split the alternatives into *review tiers*. Such review tiers are used to review household items (Amazon), movies (IMDb), restaurants (Yelp), doctors (RateMDs), and many other products and services. Both extensive literature in marketing, and anecdotal evidence from these markets, suggest that good reviews are important, if not crucial, for a business’s success. Not surprisingly, firms take an active approach to managing their reviews. Firms solicit reviews from their clients, publicly respond to reviews, and even leave fake reviews for themselves or their competitors. In this paper I study a novel form of review management: I show that firms can use discounting to facilitate *transitions* between review tiers. Imagine a product that is several good reviews short of obtaining an additional star. My question is: would it be profitable for such a product to run a price promotion?

With the help of a stylized model, I identify two effects that influence this decision: the *selection* effect and the *variance* effect. Customers who purchase the good during a discount are likely to be different from those who pay the full price and could leave different reviews. If reviews left during a discount tend to be more favorable, then the induced positive selection effect creates incentives for the product to go on sale, as new customers are likely to leave positive reviews and help the product advance to the next review tier. However, these new customers could also prove to be a worse match with the product and leave less favorable reviews. Thus, the sign of the selection effect is ultimately an empirical question.

Previous work studying the selection effect in digital markets produced mixed findings. Some of the papers that documented a negative selection effect struggle to explain why firms run promotions which are unlikely to be profitable and are unlikely to attract positive reviews (see the Literature Review section). My variance effect provides a new and attractive explanation for this phenomenon.

The variance effect reflects the idea that a product that is close to a transition is willing to gamble on upgrading its review tier. Additional positive reviews can catapult it to the next review bin, which promises higher sales. In contrast, a few negative reviews might not affect its review tier—and, therefore, sales—by much. Thus, even if the expected effect of price promotion on the reviews is negative, the *expected payoff* could still be positive due to the asymmetry that arises from the tier system’s discreteness. Similarly, while extra variance is beneficial for products on the verge of a tier upgrade, it could be harmful to products close to sliding one review tier down. Fearing to lose that precious star, such products would prefer to delay a price promotion, unless the selection effect is positive and strong.

To quantify these effects, I analyze data from Steam, a major marketplace for computer video games. I build and estimate a structural model of demand and reviewing behavior on this platform. My model admits full heterogeneity between products in the engagement of users and the absolute effects of reviews and discounts on sales. Despite this richness, I show how the parameters, including nonparametric distributions of the random effects, could be estimated. I use this model to lay the groundwork for my analysis, as it helps me confirm the importance of better review tiers for sales on Steam. Compared to a mediocre review tier, the best tier comes with a premium of as high as 10% in additional sales, and the worst one entails a 13% reduction in sales. My estimates also highlight how impactful discounts are in this market, with an average discount accompanied by a 38% spike in sales. Thus, firms on Steam have both the ability to use discounts to attract new users, and the incentives to upgrade their review tiers, given their differential contribution to sales.

I further document that these incentives translate into action. I find that products that

are close to upgrading their review tier are 4-9% more likely to sell at a discount, and I find some evidence that proximity to a downward transition decreases the probability of selling at a discount by up to 4%. I then employ my structural model to disentangle the relative contributions of the variance and the selection effects to the observed discounting patterns. To that end, I estimate the effect of discounts on the probabilities to leave positive and negative reviews. I find no evidence that selection is positive on Steam, meaning that, on average, reviews left during a price promotion are not more favorable than the ones left outside of a price promotion. Thus, I attribute firms' behavior to the variance effect. While this effect is harder to test directly, I provide some indirect tests of its impact on the discounting behavior around review tier transitions. I estimate the size of the selection effect at the product level and control for it in my regressions of discounts on proximities to transition. I find that, in the absence of the selection effect, products close to a review downgrade are 6% less likely to discount—an increase in both the size and the significance of the effect compared to the specification that does not control for the selection effect. This finding is consistent with the idea that such products are averse to variance in their review scores. I also provide novel evidence that consumers who buy the product during a price promotion are more likely to leave a review, which speaks to the mechanism through which discounts create “variance”.

My paper makes an important step towards a bigger goal of studying the life cycles of products in online marketplaces. The flip side of my finding that a product is more likely to be sold at a discount when a review tier transition is close is that it is less likely to run a price promotion when such transition is far. This idea would be a natural input into a model of pricing behavior of a product that maximizes its success over the entire life span. Such a model would exhibit a complicated joint dynamics of review, sales, and price processes. A period of aggressive discounting would help a new product attract customers and transition to the best review tier it can, and would then be followed by a period of high prices that exploit the accumulated reputation. This study serves as a laboratory that isolates some of the effects present in such a model and foreshadows some of the findings and issues that

could emerge from it.

I also contribute to the ongoing research effort to improve the design of recommendation systems. Theoretical advancements have been recently made in understanding separate components of the ideal system, such as exploration of new products vs. exploitation of the established options [Che and Hörner, 2018], creation of incentives for firms to invest in quality [Aperjis and Johari, 2010], promotion of competition, and reduction of barriers to entry [Vellodi, 2020]—to name just a few topics and contributions. At the heart of every such inquiry lies a particular strategic interaction between firms, users, and the platform. My paper provides empirical evidence on what some of these interactions result in. I show that firms respond to the tier structure of review systems by changing their pricing behavior around the thresholds and that they use pricing to affect their review process. My estimates can inform future work that will synthesize theoretical and data-driven insights to make concrete design recommendations. For example, Vellodi [2020] shows that entry barriers resulting from the low speed of review accumulation by new firms call for a review system that censors the reviews of the most established products. However, when firms can offer lower prices to build up the crucial review base, the need for the design intervention is diminished. My paper directly informs this design question by quantifying two fundamental forces involved: the frequency with which firms change prices to influence their review status and the response of the reviewing behavior to these changes in prices.

1.1 Literature Review

While some of my findings are new and interesting on their own, others are best evaluated against the literatures in economics and marketing to which I contribute. The oldest of these is the literature investigating the impact of online consumer reviews on sales. An extensive review of the studies focusing on this topic and their meta-analysis is provided, for example, by Floyd et al. [2014]. In a seminal contribution to this literature, Chevalier and Mayzlin [2006] uses a difference-in-differences approach, exploiting the differences in reviews

for the same books across Amazon.com and Barnesandnoble.com, to argue that better reviews causally improve sales. A similar approach applied to the same market was recently used by Reimers and Waldfogel [2020], which reaffirms the importance of reviews for sales of books and quantifies the welfare implications of consumer reviews’ informational content. Zhu and Zhang [2010] exploits the differences between two video game consoles rather than websites and finds that review ratings matter only for less popular games.

For identification, the difference-in-differences approach relies on the existence of several comparable platforms and assumes that the unobserved platform-specific tastes for those platforms are fixed. Another approach to identification exploits the rounding of the review data that platforms use to produce simple visual review labels. For example, Yelp.com, a popular restaurant comparison service, assigns four stars to restaurants that have an average review score between 4 and 4.24 (out of 5), but four and a half stars to restaurants with an average score between 4.25 and 4.74. Anderson and Magruder [2012] and Luca [2016] develop a regression discontinuity approach exploiting this rounding to show that an additional half-star on Yelp.com increases restaurants’ traffic and revenue. In studies relying on a discontinuity for identification, it is crucial that assignment around thresholds is as good as random. Unlike the settings in which the variable that is measured against the cutoff is realized once (a student takes a test once and either earns an “A” or a “B”), review scores of products on review platforms are moving slowly and are tested against the threshold daily. If a product is slightly below the cutoff, why would it not try to do something to cross the cutoff, especially if better review bins are more lucrative? My paper asks precisely that question and shows that such concerns are well-founded. I do not study the implications of such behavior for a regression discontinuity approach used in the papers above, but my findings could inform further inquiry into the subject.

Sorokin and Stevens [2020] raises other concerns about the validity of the regression discontinuity approach in studying the causal effect of reviews on product outcomes. To my knowledge, it was the first paper to use Steam to study the causal effect of online

reviews. In the absence of direct sales data, [Sorokin and Stevens \[2020\]](#) uses a particular regression specification that extracts the information on sales from video games’ usage data. I inherit the data from that paper, but confront the absence of sales data in a new way: by formulating a structural model of the data generating process. With the help of this model, I manage to extract information on product sales from usage data. I compare the structural model estimates with the ones obtained by its simpler regression analog and find important similarities and dissimilarities. The last improvement of the present study over [Sorokin and Stevens \[2020\]](#) is that I employ all the relevant review tiers present on Steam, while in earlier work the regression discontinuity approach dictated the focus on a certain subset of review labels.

With the benefits of having good reviews come the incentives to influence the reviews in order to reap those benefits. There are several ways for firms to affect their online reviews. One is review fraud. [Mayzlin et al. \[2014\]](#) provides evidence that small hotel owners leave negative reviews for their competitors and good reviews for themselves. [Luca and Zervas \[2016\]](#) provides similar evidence for restaurants on Yelp. A more scrupulous way for a firm to influence reviews is by assuming an active approach to review management—for instance, by publicly responding to the reviews it receives. [Gu and Ye \[2014\]](#) provides evidence that consumers who had once left a negative review of their hotel experience were more likely to leave a positive review for the same hotel in the future if their first review was responded to by the management. [Xie et al. \[2017\]](#) shows that lengthy in-depth managerial responses to negative reviews are associated with better financial performance in the future. This literature has also been concerned with how consumers change their reviewing behavior when managers start responding to reviews, essentially moving from studying individual responses to describing equilibrium outcomes. [Proserpio and Zervas \[2017\]](#) argues that hotels responding to bad reviews experience an increase in both the valence and the volume of their reviews, attributing it to consumers’ reluctance to leave short indefensible reviews in a situation when they can be responded to. In somewhat contradictory findings, [Chevalier](#)

et al. [2018] finds the opposite effect in the same market, arguing that managerial responses encourage unfavorable review activity. The idea is that consumers receive a signal that the firm is “listening” to them and are stimulated to leave negative reviews that they deem more influential.

However, direct ways of managing online reputation discussed above are not the only avenues through which businesses can affect their reviews. Another strand of literature is studying whether price promotions have an effect on firms’ online reputation. In a seminal study, Byers et al. [2012] documents that restaurants receive substantially worse reviews after running a price promotion through Groupon, a promotion service. In other words, Byers et al. [2012] documents a negative selection effect on Yelp. It presents evidence that customers attracted by promotion are different from regular customers and that they are more likely to be trying out a new restaurant or cuisine while using the Groupon voucher. Worse match quality between the restaurants and the exploring Groupon customers could explain the dip in the review score following a price promotion. Li [2016] confirms the finding in the same market but finds that restaurants with fewer and worse reviews can benefit from participating in a Groupon promotion. There are a couple of other studies that dispute the negativity of the selection effect. Zhu et al. [2019] shows that customers who received a discount leave more positive, albeit less informative reviews. In an experimental study on eBay, Cabral and Li [2015] also finds that offering a rebate for leaving (any) review significantly decreases the likelihood of negative feedback. The effect is explained by clients reciprocating to the rebate, and not by the differences between customers buying with a rebate and without it. My study contributes to this debate by documenting the absence of a positive selection effect in a new market. Additionally, my paper is the first to focus on discrete changes in review tiers that follow discounts instead of studying the continuous effect of a price promotion on the review score.

Based on its findings, Byers et al. [2012] cautions restaurants to treat Groupon customers well to avoid adverse effects on online reputation. To the extent that managers and scholars

perceive promotions to be an effective way of managing reputation, finding a negative selection effect in the data presents a puzzle¹. A similar puzzle is encountered by Zegners [2017], which argues that less-known authors of eBooks on a crowded online marketplace are more likely to give their books out for free to build a reputation and escape the “zero reviews trap”. While the idea is intuitive, the findings are that products employing free pricing are more likely to solicit *negative* feedback when sold for free, which undermines the strategy. I provide a novel mechanism—the variance effect—that could explain the prevalence of price promotions when the selection effect of reviews is negative. The variance effect would explain the behavior observed in Zegners [2017] as follows: if a product with no reviews is already doomed, it will find it profitable to sell for free even if the probability to solicit favorable reviews is small. After all, sales can not fall below zero, but the upside of accumulating some positive reviews and escaping the trap is substantial. My variance effect conceptually echoes the “Gambling for Resurrection” effect from the political science and finance literatures [Downs and Locke, 1994]. If this type of reasoning is commonplace among firms, then the researcher would observe free pricing to deteriorate review scores, on average, even if the average effect on sales remains positive.

The rest of this paper is organized as follows. Section 2 describes the institutional setting and the data I use. Section 3 develops and estimates the structural model of demand and gaming activity on Steam. The technical nature of the material in that section contributes greatly to its size, but the purpose is to simply lay the groundwork for the main analysis by quantifying the importance of different review tiers on sales and showing that discounts are effective in attracting new customers. The main findings are described in Section 4. It opens with a stylized model of a discounting decision for a product close to a review tier transition and introduces the selection and the variance effects. The section proceeds by studying such decisions in the data. Section 5 supplements the demand model from Section 3 with a model of reviewing activity and estimates its parameters to quantify the relative

¹Direct profitability of Groupon promotions is questionable; see Dholakia [2011].

contributions of the selection and the variance effects to observed patterns of discounting around review transitions. Section 6 concludes.

2 Institutional Setting and Data

2.1 Steam Marketplace

Steam is arguably the largest online marketplace for selling computer games in the world. Founded in 2003, in 2013 it was responsible for about 75% of PC games sold online globally [Cliff, 2013], and in 2017 it earned around \$4.3 billion, with an estimated market share of 18% in the entire market for PC games [PC Games, 2017]. Given an ongoing unprecedented growth in the size of the video games' market [\$43.4 billion in 2018, about the size of the U.S. film industry, Minotti, 2019] and the increasing share of online sales in this market [83% in 2018, compared to 20% in 2009, Entertainment Software Association [2019], Steam is a major player in an increasingly more important industry. Thus, this marketplace is not merely a laboratory for studies hoping to extrapolate the findings to other more well-known platforms, but is an important digital market to study in its own right.

Any user who owns a game on Steam can leave a review for that game. Starting from late 2016 Steam only uses reviews left by customers who have *purchased* the game in its review score calculation, excluding reviews left by customers who had obtained a key to activate the game through other channels (directly from the developer, or by purchasing the code from a different retailer). Steam is, thus, different from some other platforms studied in the literature, notably Yelp and TripAdvisor, where any user can leave a review (and, to a certain extent, Amazon, where non-verified users can also leave reviews). This gives me confidence in the authenticity of most reviews. This confidence is further backed by the importance of Steam to game developers and publishers, and Steam's history of monitoring the platform for fraudulent activity and excluding unscrupulous actors [Batchelor, 2019]. With such high stakes, and a non-negligible risk of being caught, there are good reasons to

Table 1: The mapping between the review score and review bins on Steam

Review Bin	N. of Reviews	Score
No Score	[0, 10)	-
Negative	Any	[0, 40)
Mixed	Any	[40, 70)
Mostly Positive	Any	[70, 80)
Positive	[0, 50)	[80, 100)
Very Positive	More than 50	[80, 100)
Overwhelmingly Positive	More than 500	[95, 100]

expect the developers to behave scrupulously.

A review consists of a binary grade, “thumbs up” or “thumbs down”, and text. A *review score* is defined as the fraction of positive reviews among all reviews a game has, provided it has at least 10 reviews. Based on the score, Steam assigns *review tiers* to each game. These labels² and the mapping rules are summarized in Table 1³. For example, a game with 7 positive and 2 negative reviews would have neither a review score, nor a label, but an extra positive review would promote it to the “Positive” bin.

Steam’s homepage welcomes customers with a selection of featured games, that could include new games, popular games that are currently on a discount, games that have recently released a major update, etc. [Steamworks Documentation, 2020b]. A customer could either click on the games already presented to her, or browse one of the categories that Steam offers. Available categories are based on games’ characteristics such as performance (e.g., “Top Sellers”, “New Releases”), genre (e.g., “Racing”, “Anime”, “Simulation”), or technical characteristics (VR or controller support). Games within a category are organized into a list, with an example session presented in Figure 1. From that list the user can get some basic information about each title, such as its price or if the game is currently sold at a discount.

Importantly, if the user hovers her cursor over a game, she could further see the review label of the game and the number of reviews the game has. I use this feature of Steam’s store to rationalize why later in the analysis I choose to focus on the review labels as my main

²I use the words “tier”, “bin”, and “label” interchangeably throughout the text

³There are labels worse than “Negative”, but they are rare, so I bin all of them into “Negative”

explanatory review variables, rather than, say, the review score or the review texts. Granted, informative review texts could be important for purchase decisions [Chevalier and Mayzlin, 2006], but on Steam, in order to gain access to such information, a customer should be willing to click on the game in the first place. de Langhe et al. [2015] reports that consumers place an unreasonably high weight on the star rating of products⁴. Thus, I conjecture that games with better review labels will, other things equal, attract more customers, and lead to more sales. Crucially, according to Steam’s documentation, all review labels better than “Negative” have a very similar contribution to visibility on the platform, which implies, among other things, that games are not ordered by their review labels when the customer is browsing different categories [2019]. Therefore, any effect of the review labels on sales should come through the perceived quality differences between different bins, rather than some mechanical visibility differences. As an illustration, note that the first game in the list depicted in Figure 1 has “Mixed” reviews.

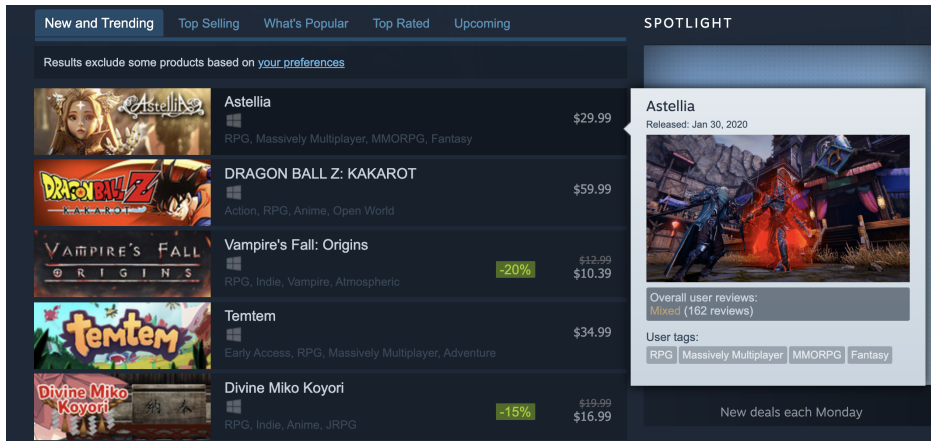


Figure 1: Example Browsing Session on Steam

2.2 Data

Valve Corporation, the owner of Steam, is notoriously secretive about its data and algorithms [PC Games, 2018]. Given the importance of the information on the performance of different

⁴The title of the present study is an homage to the masterfully chosen title of de Langhe et al. [2015].

games for game developers and publishers, the community has responded to this secrecy by establishing projects that monitor Steam in real time and extract information that could be relevant for parties interested in Steam. The majority of data used in this project comes from one such project called “Steam Database”, or “SteamDB”.

Descriptive information on any game, that I will refer to as “static”, such as its title, developer, release date, etc., are readily available from Steam, and this information is also present on SteamDB. More importantly, Steam Database contains time series of prices and player activity for each game on Steam on a daily level. In principle, the pricing information could be obtained by a repeated scraping of the store. However, the player activity information is not trivial to obtain, as Steam does not directly reveal such data, let alone the data on sales of different titles. The creators of SteamDB have declined to explain their collection method, but their reputation in the gaming community most certainly implies that they are doing the best they can. The resulting variable that SteamDB offers, that I will refer to as the *player count*, measures the maximum number of concurrent players for each game on a daily basis. To clarify this definition, consider an example game that, on some day, was played by 10 people at every hour, except noon, when 13 people played the game simultaneously. The recorded player count for this game on that day will be 13. The price and the player count time series are the major “dynamic” variables used in the analysis. The last dynamic variable, namely the review score of each game, is obtained by scraping the reviews directly from Steam, and reverse engineering the evolution of the review score.

Steam is home to more than 25,000 games that differ in their genres, prices, sales, release dates, frequency of updates, and other characteristics. Inevitably, an appropriate sample should be selected for the analysis. [Sorokin and Stevens \[2020\]](#) studies the causal effect of reviews on sales on Steam, and I mirror closely the sample selection procedure used in that paper. First, Steam has made some significant changes to its review system in late 2016, and one of the effects was the removal of a big number of reviews from the review score calculation. While it could be an intervention that is worthy of an independent study, in

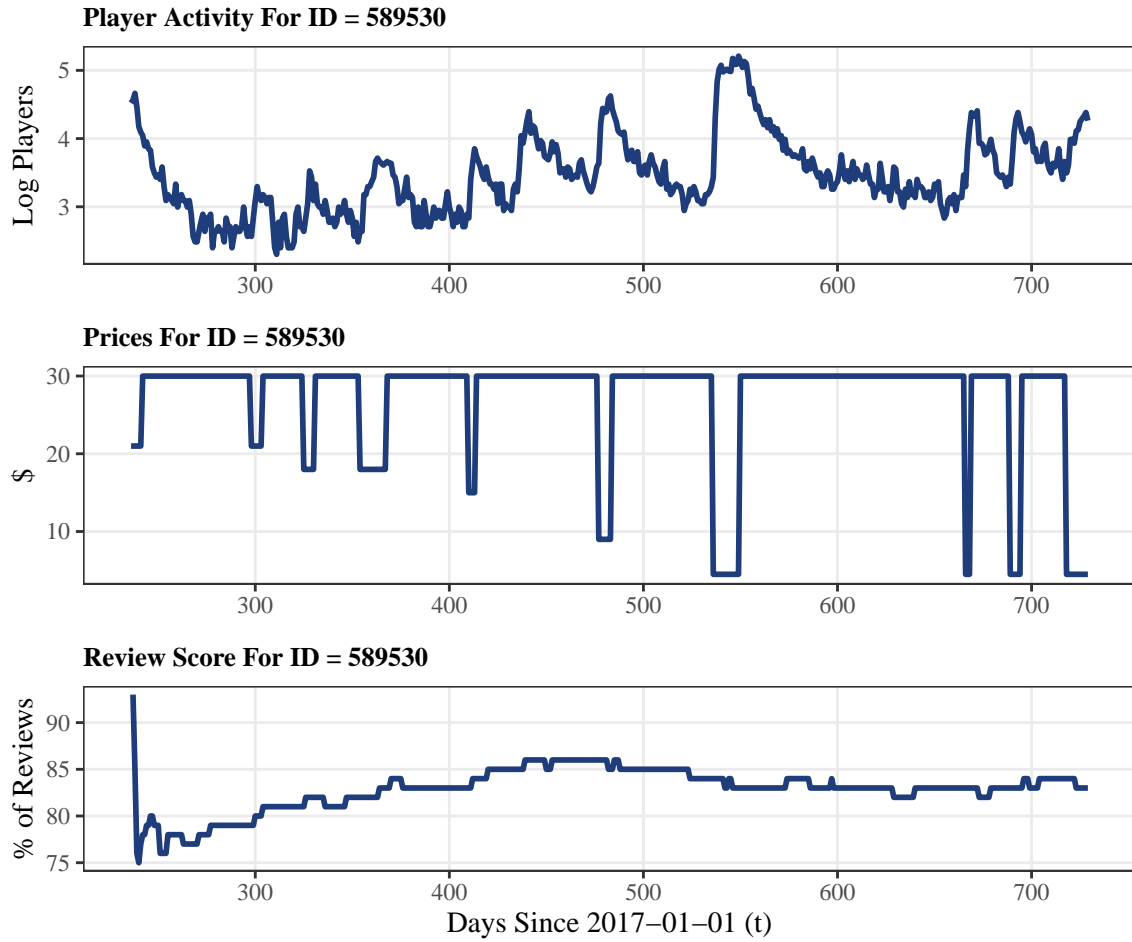


Figure 2: An Example Game in the Sample: Player Count, Price, and Review Score Histories.

practice combining the data from before and after the intervention led to unsound results, so I focus on the two year period from January 2017 to January 2019, keeping the review system as stable as possible. In particular, only games that were released in this time period make it to the sample.

Second, online multiplayer games and games that update frequently are dropped from the sample, as these are products whose quality is changing a lot over time⁵. Unobserved dynamic quality would present a big obstacle to the identification of the effect of reviews on discounting decisions or sales. Consider a game that has just issued a major update. It is a great time for the firm to give a discount in order to rekindle the interest in the game. That update could also help the game transition to a better review tier, thanks to the positive reception of the new content. Thus, game updates could give rise to a correlation between discounting decisions and review transitions that is not causal. Given the fact that the update history is an imperfect way to monitor the update activity, I decided to be conservative and to eliminate as many potential updaters as I could. For that same reason I drop free games and games that are released in a beta-version (the so called “Early Access” program), as these products are likely to be adding new content. Finally, I drop games with a player count of less than four on their median day, as these games are simply very small, and the quality of their player count data is questionable.

2.3 Sample Description

The final sample includes 906 games which I observe for 393 days each, on average. The main variables in the analysis are the aforementioned dynamic variables: player count, price, and review score. An example observation from the sample is given in Figure 2. For the majority of the games in the sample, the player count is the highest around the release date, and then it quickly fades away and oscillates around a smaller level. The player count also

⁵Games on Steam have blogs that allow them to share news about updates with the audience. I used these update announcements to measure the number of updates each game had. Games with more than 5 small updates in the sample were excluded. Publishing news about an update is voluntary, so it is possible that some games that continue to update after the release still made it to the sample.

jumps when a discount is given, reflecting the influx of new players. Games differ a lot in their sizes, as is evident from Figure 7 (in the Appendix).

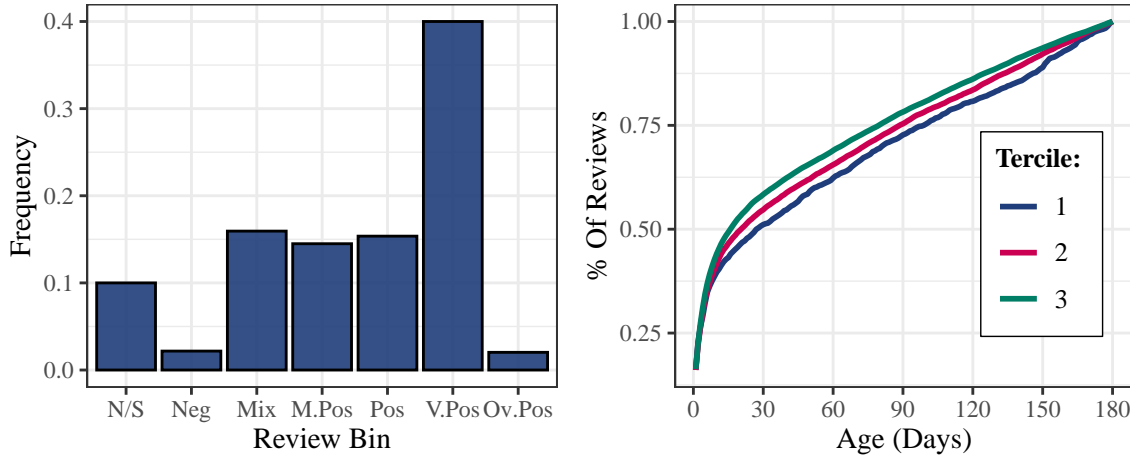


Figure 3: Distribution of Games by Review Bins and the CDF's of Review Arrival Times by Quintiles at 180 Days.

As the goal of the paper is to understand if firms use pricing tools in order to improve their online reputation, a detailed overview of pricing and review variables is in order. A representative game in the sample never changes its price, but instead occasionally goes on discounts, slowly increasing their magnitude as the game ages. There are 43 incidents of price change in the data, compared to 6331 discounts. Steam has a number of rules that regulate price promotions on its platform. A game can have a launch discount, but, otherwise, it has to wait for two months since the release before changing its price or giving a discount. A game can not go on discounts too frequently, and has to wait between four to six weeks after a price promotion to be able to run a new one. The duration of a custom discount is restricted to be at least one full day, and at most two weeks. Besides these custom discounts, that are fully managed by the firms, Steam has a series of curated discounts, when the platform invites selected titles to go on sale. As [Steamworks Documentation \[2020a\]](#) explains it, “while there aren’t strict rules, as a base guideline we tend to focus on the top 10-20% selling games on Steam that are positively reviewed and have otherwise proven to be successful”. Curated promotions are featured prominently on Steam’s main page, and, arguably, lead to more

visibility and sales for the participating titles than custom discounts, which also contribute to visibility, but typically do not get the front-page promotional slots. An important type of curated discounts are the so-called “Seasonal Sales”—big platform-wide events that take place about four times a year around major holidays. Figure 8 (in the Appendix) shows that the biggest sales take place in Winter (Christmas and New Year) and Summer (July 4), but there are also significant discounts in the Fall (Halloween and Thanksgiving). Around 64% of discounts in the sample go live during a Seasonal Sale. Thus, firms have significant agency when it comes to running price promotions, but platform regulations and platform-wide discounts are important determinants of firms’ decision to discount. On the consumer side, it is important to note that Steam users can add games to their “wish lists” in order to be notified when the product goes on sale. This feature is the most likely explanation of why discounts are followed by sharp spikes in player activity: some users who add the game to their wish lists buy it immediately when it is discounted.

The path of the review score of a typical game is quite different from its price history, in that the review score tends to settle quickly. Recall that the review score is just the fraction of positive reviews among all reviews, so the law of large numbers implies that this ratio crystallizes as more reviews flow in. Second plot in Figure 3 shows that out of all the reviews that games accumulate by the age of 180 days, a half arrives in the first month after the release. On average across games, the standard deviation of the review score over time is just 2.2 (out of one hundred), with a split of 1.83 in the first thirty days and 1.83 after that. Given how little the review score changes after the first month, the distribution of games by bins at the age of 180 days, depicted in Figure 3, should be representative of the situation at other ages.

Despite the relative rigidity of the review score, there is enough transitions between review bins to make the analysis of pricing decisions by firms around such transitions possible. Table 2 describes the transitions between review bins in the sample. The number of unique games that have changed review labels in the data is 596, and the number of transitions is

Table 2: Transition Probability And Count Matrices

Probabilities							Counts					
Neg	Mix	M. Pos	Pos	V. Pos	Ov. Pos		Neg	Mix	M. Pos	Pos	V. Pos	Ov. Pos
0	100	0	0	0	0	Negative	0	58	0	0	0	0
24	0	76	0	0	0	Mixed	60	0	189	0	0	0
0	43	0	25	33	0	M. Positive	0	245	0	141	188	0
0	0	40	0	60	0	Positive	0	2	207	0	310	0
0	0	84	0	0	16	V. Positive	0	0	220	0	0	43
0	0	0	0	100	0	Ov. Positive	0	0	0	0	25	0

1688. Sometimes games switch their bins briefly, and go back soon after. The number of transitions that led to the game spending at least 7 days in the new bin is 1225. Table 7 describes the state of the games in the two weeks before such transitions. Discounting does take place before the transitions, games transition at very different ages, but tend to have accumulated only a limited number of reviews by the time their review tier changes.

3 Empirical Model of Demand

Before I start my investigation into the discounting behavior of the products close to a review transition, it is important to address two questions. First, do discounts matter on Steam? That is, do they generate enough additional sales to be relevant? While the answer might be obvious for large sought-after products, the situation is less clear for an average game. The second question is: how big are the differences between various review tiers? If transitioning to a better tier does not increase sales by much, then firms would be unlikely to pursue such transitions when making their discount decisions. In this section I answer these questions and discuss in detail the assumptions behind my analysis, and the interpretation of the parameter estimates.

The main takeaways from the section are: better review labels improve sales, and discounts are important drivers of sales on Steam. Having few reviews or having negative reviews depresses sales by 5-13% compared to having “Mixed” reviews; the “Overwhelmingly Positive” review tier increases sales by about 10% compared to “Mixed”. An average discount

in the sample increases sales by 38% upon introduction, and by another 3-12% every day thereafter. These two facts form the foundation of my analysis of discounting behavior on a verge of a review tier transition that will be carried out in Section 4. I will further extend the model from this section by modeling the reviewing behavior of customers in Section 5.

3.1 Setup

Consider game i that is observed on a daily basis. On day t the game sells one copy to each of the B_{it} short-lived buyers that arrive on that day, a number that is unobserved by the econometrician. Define *active players* of this game, A_{it} , to be the customers who have already purchased the game and are still playing it, either because they have not yet completed it, or because they are not bored with it yet. This number has an empirical counterpart in the player count variable, observable to the econometrician. The game loses E_{it} active players on day t , which is also not observed. I assume that, once a player stops playing the game, she never returns to it again. It is easy to see then that A_{it} follows the following process:

$$A_{it} = A_{it-1} + B_{it} - E_{it} \quad (1)$$

Both the arrival of buyers and the exit of players are not observed, so some assumptions need to be made about these variables in order to be able to disentangle their contributions to the observed player counts.

Assumption 3.1 B_{it} is Poisson with arrival rate $\lambda_{it} = \lambda_i(1 + x'_{it}\beta)$, where x_{it} is a vector of observable characteristics of the game, and λ_i and β are parameters. E_{it} follows a binomial distribution $B(A_{it-1}, 1 - \psi_i)$, where ψ_i is a parameter.

The rationale behind assumption (3.1) is simple. Consumers arrive every day according to the game specific arrival rate λ_i , but that rate can go up or down depending on the values of observable characteristics x_{it} of the game that affect demand: price, reviews, age of the game, or seasonal factors\footnote{I do not impose a restriction that λ_{it} should be

non-negative in estimation. Only 4% of day-product pairs in my data will end up having negative predicted sales, which I take as a sign of a good model fit.}. The mapping between these variables and the number of copies sold is the demand curve for game i . Buyers of the game then become active players, and are subject to a fixed daily risk of $1 - \psi_i$ of abandoning the game. Given that the number of active players “flipping” this coin at the end of day $t - 1$ is A_{it-1} , this process gives rise to the binomial distribution for the number of exiters. The demand model is fully parametrized by the vector $(\{\lambda_i, \psi_i\}_{i=1}^n, \beta)$.

3.2 Identification

There are two key identification challenges I need to address in order to estimate the parameters of the demand model from my data. One challenge comes from the fact that I need to know something about (unobserved) player exit in order to tease out new purchases from the player activity data. The problem is illustrated by equation (1), where the observed part of the data ($A_{it} - A_{it-1}$) only identifies the difference ($B_{it} - E_{it}$). Consider two hypothetical games that have the same observed median daily player count—say, ten people. However, game one offers a lot of replayability and is played by the same ten people over and over again, while game two is played by new ten people every day. Clearly, such two products with identical player activity patterns would have very different sales. I am able to solve this identification problem with the help of Assumption 3.1. The key insight is that observed factors determine the arrival of new players (buyers) in *absolute* terms, while exit is *proportional* to the number of active players. Intuitively, the exit process is identified by the rate of decay of the player count when the count is far from its (slowly-changing) trend or average. Player count is typically far from its average on the days following the game release or the introduction of a discount. On such occasions a lot of new players buy the game simultaneously, which leads to sharp spikes in the player count variable (see example game in Figure 2). For a given decay parameter, the level at which the player count settles identifies the flow rate of new buyers. Given that identifying the exit rate is crucial for teasing out the sales component, it

is a great advantage of my model that it allows full heterogeneity in the exit rate $(1 - \psi_i)$.

The second identification challenge I tackle is obtaining estimates of the causal effect of review tiers on sales. In short, I achieve this by minimizing the omitted variable bias through sample selection, where I eliminate products that are more likely to be affected by unobserved confounding factors. I complement this approach by controlling for game-specific and time-specific shocks in my regressions, in addition to the observable covariates that serve as important control variables on their own. The discussion below elaborates on the details of identifying the aforementioned effects of reviews on sales, as well as the effects of price variables—the main components of the parameter vector β .

The prediction of A_{it} that follows from (1) takes the form of

$$\mathbb{E}[A_{it} | A_{it-1}, x_{it}] = A_{it-1} + \lambda_i(1 + x'_{it}\beta) - (1 - \psi_i)A_{it-1} = \psi_i A_{it-1} + \lambda_i(1 + x'_{it}\beta), \quad (2)$$

so the model implies the following regression equation

$$A_{it} = \psi_i A_{it-1} + \lambda_i(1 + x'_{it}\beta) + u_{it}, \quad (3)$$

with $\mathbb{E}[u_i | A_{it-1}, x_{it}] = 0$. This model is non-linear in parameters, because λ_i is not known and enters the model multiplying β , the parameters common to all games. If not for this commonality in β , estimation of (3) would be straightforwardly achieved by opening the parentheses and estimating

$$A_{it} = \lambda_i + \psi_i A_{it-1} + x'_{it}\beta_i + u_{it} \quad (4)$$

using OLS on a game-by-game basis. This commonality is essential, however, because one game in the sample typically does not exhibit enough variation in review labels to be able to identify the effect of upgrading the review tier on sales.

The fact that $x'_{it}\beta$ multiplies the game fixed effect λ_i allows me to estimate the dependence

of quantity sold on price and reviews in relative terms, i.e., to obtain elasticities. A change of 0.01 in the index $x'_{it}\beta$ means that the quantity demanded of game i goes up by one percent. Normally, a log transformation of the dependent variable is used to estimate elasticities. Indeed, Sorokin and Stevens [2020] is able to estimate these elasticities with a within-estimator, using the following model:

$$\log A_{it} = \tilde{\lambda}_i + \tilde{\psi}_i \log A_{it-1} + x'_{it}\tilde{\beta} + u_{it}. \quad (5)$$

This specification “controls” for the (log) number of continuing players in order to overcome the non-availability of direct sales data; however, it is the numbers of active players and buyers that are additive, not their logs, as (1) and (3) highlight. Thus, the structural model I formulate confronts the estimation problem in a more “heads-up” way, and in that regard offers an improvement over Sorokin and Stevens [2020]. I will contrast the results obtained by the two approaches in the results section, comparing a more detailed structural approach with a less precise, but a more easily implementable, log regression approach.

Another important reason for insisting on estimating the relative effects stems from the limitations of my data. Recall that the player count variable measures only the *maximum concurrent* number of active players every day. This implies that the estimated values of λ_{it} would take into account only those new buyers who contribute to gaming activity during the “rush hour”. An implicit assumption in my analysis is that new buyers of the game choose their gaming time following some fixed game-specific distribution. I use this assumption to say that, if the number of new active players during rush hour goes up by 1%, then sales go up by 1% across all types of players, not only among the “rush hour ones”.

Assumption $\mathbb{E}[u_i | A_{it-1}, x_{it}] = 0$ holds by definition for Model (3), and guarantees identification of all parameters, as long as there is sufficient variation in the data. Of course, this is only true as long as equation (3) is the right model of the data generating process. A threat to identification would come from unobserved demand shifters \tilde{x}_{it} that are correlated

with the observed factors x_{it} (omitted variable bias). For instance, an unobserved advertising intervention that is coupled with a price promotion would increase demand, but the entire effect would be attributed to the observed change in price. My approach is vulnerable to such events, as long as one is interested in getting the causal estimate of the discount elasticity of demand.

However, I argue that (3) is an adequate specification if the goal is to estimate the causal effect of reviews on sales, or to get a predictive model of demand. First and foremost, the reviews are not a choice variable of a firm, and rather serve as a state variable that a firm takes as given every day. Of course, there are various things that a firm can do, that can, in a non-guaranteed fashion, affect this variable. But, as long as the major tools that a firm has access to are controlled for in the regression, the exogenous variation in the review variables is sufficient to identify the causal effect of reviews on sales.

The most direct way in which I control for firm behavior is by including the price in the set of explanatory variables. Any intervention that is correlated with discounts will be attributed to the price variable. Second, recall that the way I constructed the sample rules out the possibility of an omitted variable bias stemming from a number of variables that could collectively be referred to as “changing quality”. The games in the sample are single player games, and thus are not subject to time-varying network externalities or frequent quality updates that could be correlated with reviews, depths of discounts, or the player count. Third, game specific effect λ_i controls for all time-invariant characteristics of game i that determine average sales: initial marketing budget, extraneous popularity of the game’s plot or setting, etc. Similarly, I employ a set of time effects to control for important within-week seasonality in gaming patterns and for the extensive platform-wide sales.

3.2.1 Price Elasticity of Demand

To close off the discussion of identification of the demand parameters in Equation (3), I would like to elaborate on the identification of the price elasticity parameter. As I mentioned before,

at the very least, including the price in the regression controls for unobserved marketing interventions that are coupled with discounts. Estimating the true price elasticity of demand is not important for the questions addressed in this paper, as I am not trying to prescribe the sizes of the discounts that firms should be giving to have a meaningful chance of affecting their review labels when they are close to a review transition. Arguably, a 100% discount is a powerful enough option to make this strategy viable, at least in principle. However, should the price coefficient be of primary interest, I would like to list some further factors that have an impact on its identification.

First, I believe that standard concerns about the endogeneity of prices are not directly applicable in my context. The reason for that is the high frequency nature of the data and the stickiness of posted prices. Using the classic notation, imagine that a discount is given on date t , and we observe a quantity-price pair (Q_t, P_t) , both of which are different from their yesterday's counterparts (Q_{t-1}, P_{t-1}) . The standard endogeneity concern is that firm's demand is subject to shocks, and that the firm would change its price precisely when those shocks occur. The two points then, roughly speaking, would belong to different demand curves, and one can not identify the slope of the demand curve. In my setting, however, this would require the firms to systematically give discounts *exactly on the days* of the demand shocks, which requires possessing a level of insight into one's demand condition that is unrealistic, especially for small independent studios. While a discount for a racing game on the day of a major F-1 race is not implausible, should the demand for the game go up with a lag of as little as one day, then both (Q_t, P_t) and (Q_{t-1}, P_{t-1}) would belong to the same demand curve, and, therefore, identification of the slope parameter would not be threatened⁶.

My second point is that, even though unobserved marketing interventions that are coupled with price promotions, would, undoubtedly, be an issue for identifying the causal price elasticity, this problem could be addressed with some extra data collection. One way to proceed would be to study if such coupled promotions in fact do take place. In particular,

⁶Assuming, as is usually done, that demand shifters lead to parallel shifts in demand

one could collect data on YouTube queries mentioning the games in the sample, and use spikes in such queries as a proxy for unobserved marketing campaigns.

A deeper problem for identifying the price elasticity of demand lies in defining precisely the elasticity of interest. A video game is a durable good, and some consumers could be purchasing it strategically, thinking about the probability and depth of discounts that they could get in the near future. In particular, Steam allows users to add any game to their wishlist, which means that they would be notified about promotions affecting that game. Later in the results section we will see that the lion’s share of sales on a discount takes place on the first day of the discount, which is consistent with users waiting for a discount and buying the game when the discount goes live. For studying counterfactual price policies one would need to specify a more complicated model of forward-looking consumers, and estimate the fundamentals of their behavior.

A related problem is the importance of salience in a marketplace that has many thousands of products. Any estimate of the price elasticity is subject to Steam keeping its algorithms unchanged. A game can give a 99% discount, but the quantity demanded will not change much if that promotion happens to not be reflected in Steam’s system. This has implications for the identification of the slope of the demand curve. For instance, consider the problem of finding good instrumental variables for prices. Curated discounts, created and managed by Steam, are suggested to firms, and not chosen by them. In principle, such promotions could serve as valid instruments for prices, as they are likely exogenous to immediate daily demand conditions. However, validity would be violated because such curated discounts also bring an immense boost in visibility, by the virtue of promoting the discounted game to the front page.

3.3 Results

An ideal set of covariates x_{it} that I would like to use in estimating (3), given by

$$A_{it} = \psi_i A_{it-1} + \lambda_i (1 + x'_{it} \beta) + u_{it},$$

would be price, review label, score, log reviews, age, and a set of day of the week and sample week time effects. However, this proved to be computationally infeasible. Even though I manage to concentrate out $2n = 1812$ game-specific parameters (λ_i, ψ_i) from the numerical optimization routine, estimation of β still relies on minimizing the sum of squared residuals in a $\dim(\beta)$ -dimensional space. The routine would fail to converge to a solution within reasonable boundaries, so I had to alter the specification.

The week effects contribute the most to the dimensionality of the problem, as there are 105 weeks in the sample. The reason to include these week effects is to account for platform-wide shocks. The biggest shocks shared by games on Steam are Seasonal Sales. These sales take place around major holidays, and, thus, blend together the increased platform-wide demand due to holidays with the higher quantity demanded caused by the plethora of price promotions (depicted in Figure 8). To capture these periods in a parsimonious way, I calculated the average daily discount in the sample, and labeled the days when the average discount exceeded 20% as days of the Seasonal Sale. Figure 8 in the Appendix shows that my definition tracks closely the spikes in the aggregate discounting behavior. I also tried using the raw value of the average discount in the sample instead of a dummy variable indicating the Seasonal Sale, but the results remained the same.

The substitution of the week dummies with a Seasonal Sale dummy proved to be sufficient for convergence. I mentioned in the Identification section that a more simplistic, yet more tractable, alternative to estimating the non-linear model (3) is given by equation (5)

$$\log A_{it} = \tilde{\lambda}_i + \tilde{\psi}_i \log A_{it-1} + x'_{it} \tilde{\beta} + u_{it}$$

To check whether the omitted week effects could play a crucial role, I report the results of estimating (5) both with the ideal set of covariates, and with the covariates used to estimate (3). The results are presented in Table 3.

Column (1) in Table 3 presents the estimates of the demand parameters in (3), and

Table 3: Estimates of the Demand Parameters

	(1)	(2)	(3)
Price	−0.342*** (0.010)	−0.064*** (0.005)	−0.075*** (0.005)
New Discount	0.230*** (0.013)	0.256*** (0.008)	0.284*** (0.008)
No Score	−0.128*** (0.019)	0.196*** (0.031)	0.215*** (0.030)
Negative	0.005 (0.020)	−0.051*** (0.018)	−0.048*** (0.018)
M. Positive	0.019** (0.008)	0.033*** (0.008)	0.038*** (0.008)
Positive	0.054*** (0.010)	0.039*** (0.012)	0.041*** (0.012)
V. Positive	0.052*** (0.010)	0.069*** (0.011)	0.072*** (0.011)
Ov. Positive	0.100*** (0.013)	0.084*** (0.013)	0.093*** (0.013)
Score	−0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)
Log Reviews	−0.091*** (0.002)	0.026*** (0.006)	0.031*** (0.006)
Age	0.000*** (0.000)	−0.000*** (0.000)	0.001*** (0.000)
Age ≤ 14	0.146*** (0.008)	0.144*** (0.006)	0.170*** (0.006)
Seasonal Sale	−0.006 (0.004)	0.071*** (0.003)	
Lag Players		0.798*** (0.002)	0.792*** (0.002)
Weekdays	✓	✓	✓
Week	×	×	✓
Observations	355900	355900	355983
R ²	0.850	0.707	0.708

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

column (2) presents the results of estimating the log-regression (5) with the same set of covariates; finally, column (3) reports the results of estimating the log-regression with the ideal set of covariates (using week effects instead of the Seasonal Sale dummy). It is readily checked that the difference between the preferred log-regression in column (3) and its restricted analog in column (2) is very small. The R^2 statistic confirms that week effects do not contribute much to explaining the data, and I conclude that the Seasonal Sale dummy captures the main time-specific shocks well. Thus, it is reasonable to say that the omission of the week effects from the nonlinear Model (3) comes at little cost. Now we can concentrate on the estimates of the demand parameters in (3), presented in the first column, and compare them with the heuristic regression results in the second or third column.

The semi-elasticities of sales with respect to review labels are fairly similar across the specifications. All specifications exhibit monotonicity in the effects of different review tiers on sales. The “Negative” label is at most as good as “Mixed” (the reference group), with the log-regression results finding a penalty of 5% that the “Negative” label entails; the “Mostly Positive” bin increases sales by 2-4% compared to the “Mixed” bin. At the top end we see that the “Overwhelmingly Positive” label increases demand by 8-10%. These magnitudes are economically sizable and reasonable. These results confirm that good reviews are important for sales in the Steam marketplace.

However, columns (1) and (2)-(3) differ substantially in the estimates of the effect of having no score on sales⁷. While the structural model (3) suggests that having no score is associated with a 13% slower customer arrival than having the “Mixed” review label, its simplified log-regression analogs find the opposite effect of a having 20-21% bonus associated with having not enough reviews. The latter effect can not possibly be taken at face value, given that the most exclusive review label, “Overwhelmingly Positive”, increases sales by at most 10% compared to the baseline. It would be possible to explain such large estimates

⁷In order to have both the score variable $\in [0, 100]$ and a dummy for the “No Score” label I set the former to be equal to 0 when a game has no review label (which happens when it has less than 10 reviews). Thus, the score variable really measures the effect of score once the score is defined.

by high levels of demand when the game is young and is likely to have less than 10 reviews, but I control for being “young” by including the dummy for being less than 14 days of age⁸. Thus, the coefficient on “No Score” is identified by games that exit this bin later than their first two weeks. Such games constitute 23% of the sample, and are probably different from the rest of the games. However, it is still hard to come up with omitted factors that would explain the positive effect of “No Score” on sales found in columns (2)-(3). One explanation that works is that the log-regression simplification (5) is simply misguided, and that one should only trust the coefficients from the structural model of the demand process (3). This latter model finds a penalty of 13% associated with having no score, which sounds much more reasonable. A negative effect could be attributed to customers’ reluctance to purchase products of unknown quality. If that is the case, this finding suggests a pretty serious cost of asymmetric information in this market.

The second important set of variables are the price variables. The first row of Table 3 suggests that the price elasticity of demand is very low, between 0.06 and 0.34 in absolute value. This magnitude is very similar to the price elasticity that Reimers and Waldfogel [2020] finds studying reviews for books on Amazon.com. That paper argues that Amazon has been known to prioritize growth over profit, thus charging low prices and operating on the inelastic part of the demand curve. However, individual firms on Steam choose which prices to charge for their product, and they should have a stronger preference for profit than an entity that is a marketplace, rather than a seller. The situation becomes more clear after examining the coefficient on “New Discount”, an indicator variable for the first day after the introduction of the discount⁹. As I have mentioned in the Identification section, the durable good nature of video games means that forward-looking behavior by consumers could lead to different price elasticities for short-lived promotions and long-term price changes, and the difference between the price coefficient and the indicator for the start of a discount could capture precisely that distinction. I find that, on average, a discount leads to a 23-28% spike

⁸A two weeks cutoff is inspired by high speed of review arrival documented in Figure 3

⁹I found the effect to be stronger if a one day lag is allowed

in sales upon introduction. An average discount in the sample is 45%, meaning that, based on the estimates in column 1, the effective change in quantity sold upon the introduction of the discount is $23\% + 15\% = 38\%$, with the implied elasticity of 0.84. This number looks more reasonable, especially given that the products I study have a zero marginal cost of production, and in a static world firms would set prices at which the demand for their products would be unit-elastic. Every additional day on sale is then associated with a 3-15% higher quantity demanded, where larger effect comes from the structural model. Discounts on Steam attract a considerable number of new users, and thus could be effective tools in generating new reviews needed for a review tier transition.

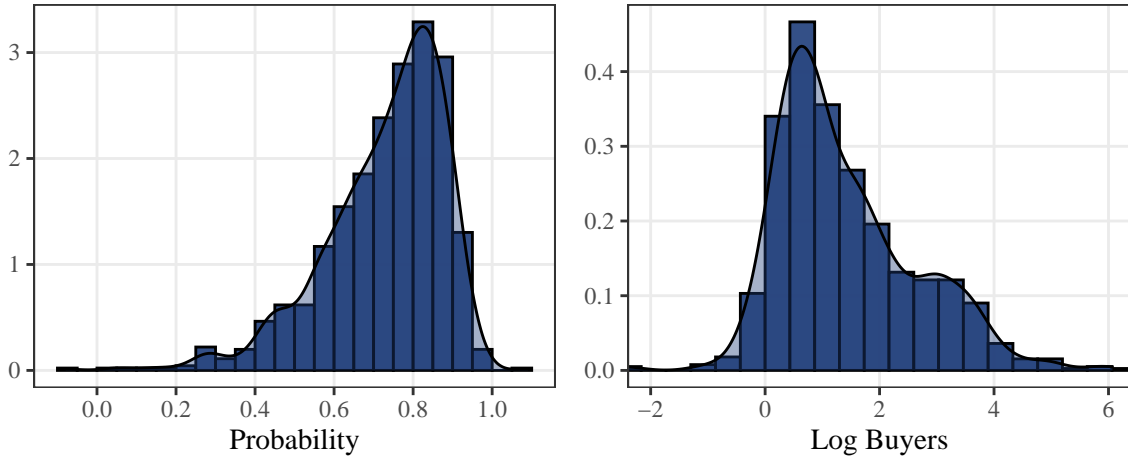


Figure 4: Distributions of the continuation probability ψ_i and average sales $\bar{\lambda}_i$.

I conclude the overview of the demand estimation results by going over the estimates of the game-specific customer arrival rates λ_i , and the probabilities ψ_i with which active players of game i continue playing the game from day to day. For the ease of interpretation, I present the estimates of $\bar{\lambda}_i := \lambda_i(1 + \mathbb{E}[x_{it}']\beta)$, the average sales for game i ¹⁰. Estimated distributions of these parameters are presented in Figure 4. The model produces very well-behaved estimates without any constraints imposed on the estimation: there are just 2 games with estimated values of ψ_i outside of the $[0, 1]$ interval, and only 4% of the observations in the sample predict negative sales. The average continuation probability, 0.73, is quite close

¹⁰This is not quite accurate, as I can only estimate the part of sales that contributes to the gaming activity during peak times

to the estimates of its log-regression counterpart in Table 3, line “Lag Players”, equal to 0.79.

4 Discounting and Review Bin Transitions

In the previous section I established the importance of favorable reviews for sales on Steam, and the important role that price promotions play in attracting new users. Now I turn attention to the main question of the paper: do firms use discounting to facilitate transitions to better review tiers, or to prevent transitions to worse review tiers? I start by presenting a stylized model of a product that is close to a review transition. The model predicts that the decision to discount or not depends on the balance between two forces: the selection effect and the variance effect. I then turn to data, and show that products are significantly more likely to sell at a discount when they are close to improving their review tier, while the effect for products that are close to sliding down a tier are much weaker. The extent to which the results are driven by the selection and the variance effects is then analyzed in the next section.

4.1 Stylized Model

Consider a firm that has a review status $s \in (0, 1)$. A firm can give a discount at cost c that can probabilistically change its status to exceed 1 (an upgrade), to go below 0 (a downgrade), or to remain within the $(0, 1)$ band (no change). The outcome depends on whether the additional customers who buy the product during the discount happen to leave more positive or more negative reviews, and on the volume of the new reviews relative to the existing ones. To reflect that, assume that with probability p firm’s status becomes $s + x$, and with probability $1 - p$ it becomes $s - x$. The symmetry is purely for expositional purposes. Firm’s utility $u(s)$ from having a review status $s \in (0, 1)$ is 0, while it earns u^H from the upgraded status and $u^L < 0 < u^H$ from the downgraded status. The expected utility from giving a

discount is

$$\mathbb{E}[U(s + X)] = pU(s + x) + (1 - p)U(s - x) - c \quad (6)$$

When would the firm be willing to give a discount? Consider a firm with $s > 1/2$. Such a firm could be regarded as one that is close to a review upgrade. The expected value to the firm of starting a price promotion is

$$\mathbb{E}[U(s + X)] = \begin{cases} -c & x < 1 - s \\ pu^H - c & x \in [1 - s, s) \\ pu^H + (1 - p)u^L - c & x \geq s \end{cases} \quad (7)$$

The firm would choose to start a price promotion if $\mathbb{E}[U(s + X)] > 0$. If $x < 1 - s$, the firm simply can not induce enough change to its review status, and would not give a discount, unless $c < 0$ (i.e, other factors make giving a discount attractive). This would be the case for products with many reviews and an established review score. My interest is mostly in the other type of products, the ones that potentially can transition. In those cases the decision to discount or not depends on the comparison between the value of the positive transition, u^H , the cost of giving a discount (c , or c and $(1 - p)u^L$), and the probability of a successful transition. If the probability of success is high, the firm is more likely to launch a discount. This is what I refer to as the (positive) *selection* effect. In the literature the selection effect of discounts refers to the idea that consumers buying during a price promotion could be different from regular consumers, and could therefore leave different reviews. A positive selection effect means that these reviews tend to be more favorable, while a negative selection effect implies the opposite. In the language of my stylized model, a positive selection effect means that $p > 0.5$, implying that the review score is more likely to improve from a promotion. The selection effect is negative if $p < 0.5$. Clearly, a positive selection effect improves the attractiveness of running a price promotion, and is thus an important factor to study.

Theoretically, the sign of the selection effect of buying during a discount on the favorability

of reviews is ambiguous. Consumers who buy when the price is low could be a worse match for the product and, thus, leave worse reviews. At the same time, such customers get a higher utility from paying less, which, together with a desire to reciprocate, can lead them to leaving better reviews [Cabral and Li, 2015, Ifrach et al., 2019, Acemoglu et al., 2019]. Empirical evidence on the sign of the selection effect is limited. Byers et al. [2012] and Li [2016] report that customers who choose to go to a restaurant because of a price promotion tend to leave worse reviews, while a positive effect was documented by Li [2016] and Zhu et al. [2019].

An interesting observation about the problem of the firm (7) is that the selection effect does not need to be positive for the firm to decide to launch a discount. Indeed, consider the case when the potential change to the score is moderate ($x \in [1 - s, s]$). The successful outcome of the discount leads to a transition and the prize of u^H , while the unsuccessful one leaves the firm with the current review status, which is not that costly. If the upside payoff u^H is high enough, then even if selection is negative ($p < 0.5$), the firm could still be willing to give a discount, with the sole purpose of gambling on the positive outcome. I refer to this idea as the “variance effect”. Even if the firm does not expect the new reviews to be favorable in expectation, they could still turn out to be positive, and if the reward is high enough, discounting could be worth the risk.

Note that variance could also be bad for the firm. First, if $x \geq s$, then a negative realization could push the product down one review tier. A more important case is the one of a firm that is closer to a review downgrade rather than an upgrade, i.e. when $s < 0.5$. Following the logic of (7), we can see that for intermediate values of the change in review status ($x \in [s, 1 - s]$) such a firm does not benefit from a discount. If the reviews left during the discount turn out to be negative, the firm will slide one tier down, while if they turn out to be positive, the firm will merely keep its current review status. The selection effect should be strongly positive for such a firm to find discounting profitable.

To summarize, it appears that firms close to a positive transition are more likely to benefit from the variance effect than the ones close to a negative transition. This makes them

more likely to run price promotions. Both types benefit from a positive selection effect, but whether the selection effect is positive or negative is an empirical question. A strong negative selection effect could overturn the appeal of discounting coming from the variance effect, ultimately meaning that the relationship between proximity to transitions and discounting behavior is theoretically ambiguous and should be studied empirically.

4.2 Empirical Analysis

Now that a simple theory of discounting close to a review threshold has been brought forward, I turn attention to the data. The departing point of the analysis is Figure 5, which shows that transitions between review labels on Steam are often preceded by discounts. The graph shows that two weeks prior to a transition only 15% of games are on a discount, essentially the sample average, but that this number more than doubles to around 36% one day before the label change. Regression analysis controlling for review label, number of reviews, age, day of the week and week time effects, as well as time since the previous discount, confirms that the association depicted in Figure 5 is robust (see Table 8 in the Appendix). A game is approximately 8 percentage points more likely to be on a discount on the day of the transition than it is two weeks prior to it, and it is 16 pp less likely to be on a discount two weeks following the transition. Given that the probability for a game to be on a discount on a random day is 16% in the sample, these effects amount to 3% and 7% changes in the daily probability of a discount, which is quite sizable.

Of course, this finding simply shows a correlation between firms' discounting behavior and review transitions. The same pattern could emerge if the causality between the two variables is reversed, i.e. the discounts cause transitions, and not the other way around. Imagine a hypothetical world in which games are only bought on discounts. In this world any action in the data would be preceded (and, to some extent, caused) by a discount. In other words, firms could be giving discounts for reasons unrelated to reviews, but transitions sometimes would follow as a result. As a matter of fact, suppose my theory is correct and

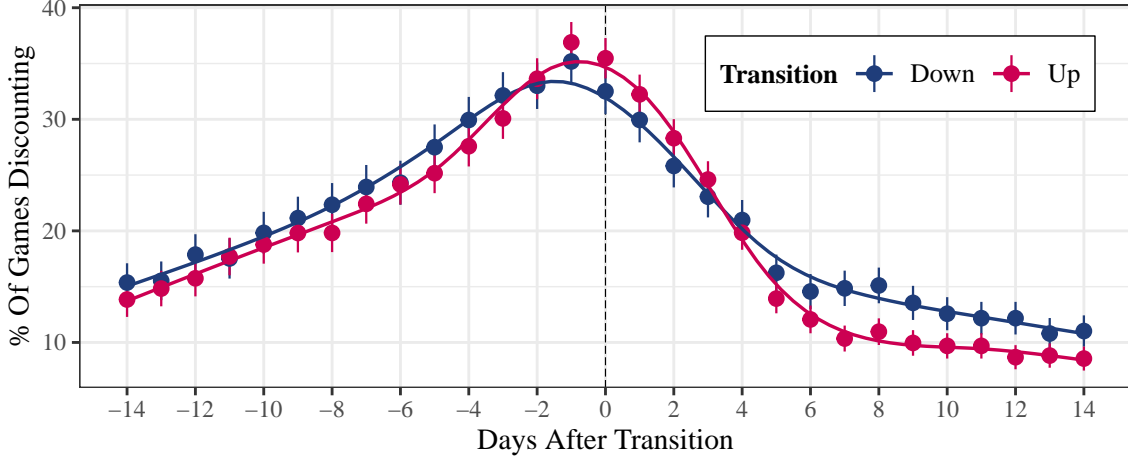


Figure 5: Discounting By Games Around The Transition.

firms try to achieve transitions via price promotions. In that case, discounts *have* to be able to aid transitioning, making the reverse causality inherent in this setting. The reverse causality would also explain why both review upgrades and downgrades are preceded by discounts.

The analysis behind Figure 5 remains imperfect for yet another reason. A discount given when a firm is close to a positive transition is not guaranteed to lead to a transition. Similarly, a firm trying to avoid a negative transition by giving a discount might succeed. In both cases, the behavior that we are interested in goes undetected if one only studies transitions that took place in the data.

The solution to both problems that I suggest is to study *potential* transitional situations instead of the realized ones. First, it solves the reverse causality problem. A discount given today can cause a transition tomorrow, but a decision to give a discount can not cause the proximity to a transition that chronologically precedes it. Second, it clearly solves the selection issue explained above, when the analysis considers solely the firms that transitioned.

While the merits of focusing on potential transitions are clear, defining proximity to a transition is not straightforward. One approach would be to use the raw review count, and to say that a game is close to upgrading its review label when it needs some fixed number of new positive reviews to transition. However, given the heterogeneity in the popularity

of different games (see Figure 7), five extra reviews could be nothing for a very popular game, and hard to acquire for a small game. For this reason I use a different approach, and measure proximity by the expected number of *days* that a game has to wait to accumulate the reviews necessary for a transition. In particular, for every game-date pair I first measure how many positive reviews that game needs at the moment to upgrade its review bin, and how many negative reviews that game needs to downgrade its review bin. In the next step, for each game I calculate the average speeds of review arrival, by simply dividing the number of positive (negative) reviews as of the last day in the sample by the age of the game on that day. Knowing the speeds of positive and negative review arrivals, and the number of reviews necessary for a transition, I then define the proximity to a positive (negative) transition to be the expected number of days needed for the game to accumulate the necessary number of positive (negative) reviews, assuming that it does not receive any negative (positive) reviews during that time.

To illustrate this definition, consider an example game with 19 positive and 6 negative reviews at some moment in time. This game has a review score of 76%, and the “Mostly Positive” review label. It needs 5 additional positive reviews to secure a score of 80%, the threshold that would earn it the “Positive” label. Similarly, 3 new negative reviews would be sufficient for the game to slide into the “Mixed” review category, as the score would become $19/(25 + 3) \times 100\% = 67\%$, which is less than the 70% required for the “Mostly Positive” bin. If this game ends up having 80 positive and 20 negative reviews at the age of 200 days, its last day in the sample, then, on average, it was receiving 0.4 positive and 0.1 negative reviews per day. Thus, for a positive transition it requires $5/0.4 = 12.5$ days of only good reviews arriving at this rate. Similarly, for a negative transition it requires $3/0.1 = 30$ days of only bad reviews arriving at this rate. Therefore, for this game I set its proximity to a potential positive transition to be 12.5 days, and its proximity to a potential negative transition to be 30 days.

As the example above shows, it is impossible to consider proximity to a positive transition

without taking into account proximity to a negative transition, at least for moderately sized games. My model of discounting had a similar message: while variance could be good for a game that does not risk a deterioration of its review tier, it could be bad for a game that does risk one. For this reason, I define two “treatment” variables of interest. Game i at date t is said to be close to a positive transition, $T_{it}^+ = 1$, if its proximity to a positive transition is less than or equal to 14 days. Similarly, game i at date t is said to be close to a negative transition, $T_{it}^- = 1$, if its proximity to a negative transition is less than or equal to 14 days. The 14 day cutoff is inspired by the maximum duration of custom discounts on Steam, and the patterns of discounting around successful transitions, depicted in Figure 5. I obtain the effect of being close to a review transition on the discounting behavior by estimating the following linear probability model:

$$disc_{it} = \beta^+ T_{it}^+ + \beta^- T_{it}^- + X_{it}\beta + f_i + \tau_t + \varepsilon_{it}, \quad (8)$$

$disc_{it} = \mathbf{1}\{Discount_{it} > 0\}$ is a dummy measuring if the game is on a discount or not, X_{it} is a set of control variables that includes log proximities to positive and negative transitions, log review count, score, review bin dummies, age, a set of game-level fixed effects f_i , and a set of day of the week and week time effects τ_t . These time effects are especially important to include in the regressions with discounting variables on the left hand side, because Steam’s curated discounts all start on predetermined days of the week, and seasonal sales affect a big number of games at the same time, as depicted in Figure 8. My model of discounting behavior highlighted that any sign of β^+ and β^- is possible, albeit suggesting that $\beta^+ > \beta^-$, as the variance effect is more likely to benefit a firm close to a review upgrade.

Notice that the inclusion of both the distance to a potential positive and a potential negative transition in the regression forces me to drop all observations from the lowest (“Negative”) and the highest (“Overwhelmingly Positive”) review bins, as for games in those bins only one direction of review transition is possible. However, given that the proportion of

Table 4: Discounts Close to Potential Transitions

	<i>Dependent variable:</i>	
	Discount Full	Probability 1 W. to Tr.
Close to Pos. Transition	0.015*** (0.003)	0.007** (0.003)
Close to Neg. Transition	-0.006** (0.003)	0.001 (0.004)
Log Days to Pos. Tr.	0.004*** (0.001)	0.001 (0.001)
Log Days to Neg Tr.	0.007** (0.003)	0.009*** (0.003)
Log Reviews	0.011** (0.004)	0.012*** (0.004)
Score	-0.001** (0.0003)	-0.001** (0.0003)
Mostly Positive	0.016** (0.008)	0.016** (0.007)
Positive	0.018** (0.009)	0.020** (0.009)
Very Positive	0.019* (0.010)	0.024** (0.010)
Age	0.0005*** (0.0001)	0.0005*** (0.0001)
Age ≤ 14	0.039*** (0.005)	0.039*** (0.005)
Time Effects: Weekdays, Week	✓	✓
Game Effects	✓	✓
Poly(t W/O Discount, $d = 2$)	✓	✓
Observations	295,228	295,228
R ²	0.428	0.428

Note:

*p<0.1; **p<0.05; ***p<0.01

such observations in the sample is quite small, this is not a big concern. The analysis also excludes the observations with a “No Review Score” label, as I do not want to take a stance on what constitutes an improvement or a deterioration of the review score for such games. My demand estimates presented in Table 3 suggest that having no reviews could be the worst review bin a game could be in, but then I would again have to exclude this label as one that does not admit review deterioration. In any case, games that spend a lot of time in “No Score” have, by construction, very few reviews, which makes them rather unlikely to ever be one one or two weeks away from a transition according to my measure.

The results of estimating Equation (8) are presented in Table 4. Column one is the preferred specification, while column two uses a more stringent definition of proximity to a transition, requiring a game to be 7, rather than 14, days away from a potential transition to be counted as being “close to a transition”. The results in Table 4 unequivocally support the hypothesis that proximity to a review bin upgrade increases firms’ willingness to run a price promotion. Measured against the 16% probability for a random game-day pair from the sample to feature a discount¹¹, the effects constitute a 4-9% increase in the daily probability of a discount. Proximity to a potential negative transition seems to have a negative effect on the probability of discount, albeit the effect is not very significant. I find that firms under a risk of deteriorating their review label are 4% less likely to go on a discount.

It is instructive to look at the results through the lens of the stylized model introduced earlier in the section. Both the variance and the positive selection effects could explain why products close to a review upgrade are more likely to run a price promotion. Weak significance of the effect for products close to a review slump could be explained by the clash between the variance effect and the selection effect (recall that we expect the variance effect to be negative for such products). As we see, even though my empirical investigation has resolved the theoretical ambiguity of the signs and the significance of the effects, it did not shed much light on the balance of the underlying forces behind these effects. I proceed to

¹¹The corresponding number for the games in the positive treatment group is 17.

investigating the two effects in greater detail now.

5 Selection and Variance Effects

In the previous section I established two key facts. First, I found that firms are more likely to go on a discount when their products are close to upgrading their review bin. This finding implies that the selection effect and the variance effect can not be both negative for games close to a review upgrade. While such a combination of the effects was theoretically possible, *a priori* it never seemed to be the leading case. Second, I found only limited change in discounting behavior for products that are close to downgrading their review bins. This finding indicates that the variance effect should be important, but is still consistent with both a positive, and a negative selection effect. In this section I investigate the signs of the selection and variance effects in greater detail.

The selection effect is easy to test directly. One would simply need to compare the fractions of positive and negative reviews left on and off discount. The variance effect is harder to test. The strongest prediction of my stylized model regarding the variance effect is that, in the absence of the selection effect, a moderate amount of variance is welcomed by the products that are close to an upward transition, and is disliked by the products close to a downward transition. I develop this idea into a test, described below. Another way to evaluate the importance of the variance effect is to look at the extent to which a discount speeds up the arrival of reviews. The idea here is that for the variance effect to be relevant, a discount should be able to bring in a significant number of new reviews.

I implement these ideas as follows. I complement my model of demand and gaming on Steam introduced in Section 3 with a model of reviewing behavior. A formal model of reviewing behavior is necessary for several reasons. First, a structural model allows me to explicitly introduce the parameters of interest—probabilities of leaving a good or a bad review, on and off a discount. With the estimates of these parameters at hand, I am able

to quantify the selection effect. I then control for the selection effect in the regression of discounting behavior on proximities to transition (8). Using my estimates, I can consider a counterfactual in which the selection effect is absent, and check if the resulting effect is consistent with the variance effect. I also use the model to estimate the effect of a price promotion on the probability to leave reviews. Since sales are unobserved, it is impossible to say if the surge in reviews during a discount is entirely due to the influx of new customers, or if part of it is due to an increased willingness to leave a review during a discount. I use my model to estimate sales, and therefore I can make inference on the propensity to leave a review, conditional on the good being discounted, gaining further insight into the origins of the variance effect.

As a preview of results, in my data I find no evidence of a positive selection effect. Table 9 in the Appendix shows that, following a discount, the review score goes down by half a point (out of 100), and does not recover in the week following the discount. With my estimates from the structural model I fail to reject the hypothesis that the selection effect is negative. When I control for the size of the selection effect at a game level, I find that the negative effect of the proximity to a negative transition on discounting increases substantially, and becomes more significant. This is my main evidence of the importance of the variance effect. Additional indirect evidence comes from analyzing the effect of discounts on the review arrival. My descriptive analysis finds that a discount is followed by around 7 extra reviews, on average, in the week after its introduction, or at least doubles the speed of review arrival (Table 9). My structural estimates show that consumers are 24-40% more likely to leave reviews during discounts. Both findings suggest that the variance effect could play an important role in firms' discounting decisions.

5.1 Review Model

Every buyer of game i is a potential reviewer. I assume that a buyer who buys the game at $t - k$ leaves a positive (negative) review for the game on day t with probability r_{it}^+ (r_{it}^-), and

no review otherwise. For the reasons I explain later, I will refer to r_{it}^+ as the *like rate*, and to r_{it}^- as the *dislike rate*. The focus of the analysis is on the difference between the reviews left on and off a discount. To that end, I parametrize the like and dislike rates to depend on the discounting behavior of the firm and other covariates as follows.

Assumption 5.1 *The like rate is a linear function of covariates $w_{it} = [1, disc_{it-k}, \dots]'$: $r_{it}^+ = w_{it}'\rho_i^+$. The dislike rate is a linear function of covariates: $r_{it}^- = w_{it}'\rho_i^-$.*

In the simplest specification I use $w_{it} = [1, disc_{it-k}]'$, $\rho_i^+ = \rho^+$ for all i , implying $r_{it}^+ = \rho_0^+ + disc_{it}\rho_1^+$. This simply says that buyers of all games leave a positive reviews with one of the two possible probabilities: ρ_0^+ off discount, and $\rho_0^+ + \rho_1^+$ on discount. Similarly, a probability of a negative review could be either ρ_0^- , or $\rho_0^- + \rho_1^-$. While the model with no heterogeneity in parameters is not very realistic (for one, it implies that two games with the same discounting behavior will have the same expected review score), I use this model to make inference on average parameter values. This is econometrically easier than estimating the entire distribution of (ρ_i^+, ρ_i^-) , and then testing hypotheses about the functions of the means of those distributions. The assumption here is that the model without heterogeneity reasonably approximates the aforementioned averages: $\rho^+ \approx \bar{\rho}_i^+$, $\rho^- \approx \bar{\rho}_i^-$.

In the introduction I implicitly defined two hypotheses I want to test. First, I want to know if consumers are more likely to leave a review when they buy on discount. This would constitute an indirect test of the variance effect. In order to do so, I need to test if $\rho_0^+ + \rho_1^+ + \rho_0^- + \rho_1^- > \rho_0^+ + \rho_0^-$, or, more formally:

$$H_0 : \rho_1^+ + \rho_1^- \leq 0 \tag{9}$$

$$H_1 : \rho_1^+ + \rho_1^- > 0 \tag{10}$$

The second hypothesis I am interested in is whether the selection effect is positive or not. The review score of a product measures the fraction of positive reviews among all reviews left, reflecting the probability to get a good review *conditional* on having a review. Indeed, if

$L_t \in \{0, 1\}$ is a random variable that takes the value of 1 if the t -th review is a like and 0 if it is a dislike, then the review score based on T reviews would simply be $Score(T) = \frac{1}{T} \sum_{t=1}^T L_t$, which converges to $\mathbb{E}[L_t] = \mathbb{P}(L_t = 1)$ as T grows to infinity. For a user buying a game off discount we have

$$\mathbb{P}(L_t = 1) = \mathbb{P}(\text{"Like"} \mid \text{"Review"}) = \frac{\rho_0^+}{\rho_0^+ + \rho_0^-} \quad (11)$$

For a user buying the same game on discount we have

$$\mathbb{P}(L_t = 1) = \mathbb{P}(\text{"Like"} \mid \text{"Review"}) = \frac{\rho_0^+ + \rho_1^+}{\rho_0^+ + \rho_1^+ + \rho_0^- + \rho_1^-} \quad (12)$$

I can then test the sign of the selection effect by testing the following hypothesis: the expected review score on a discount is not higher than the expected review score off a discount. Formally, we have

$$H_0 : \frac{\rho_0^+ + \rho_1^+}{\rho_0^+ + \rho_1^+ + \rho_0^- + \rho_1^-} - \frac{\rho_0^+}{\rho_0^+ + \rho_0^-} \leq 0 \quad (13)$$

$$H_1 : \frac{\rho_0^+ + \rho_1^+}{\rho_0^+ + \rho_1^+ + \rho_0^- + \rho_1^-} - \frac{\rho_0^+}{\rho_0^+ + \rho_0^-} > 0 \quad (14)$$

5.2 Identification

To make my identification argument I will focus on the case of two regressors, $w_{it} = [1, disc_{it-k}]$, and heterogeneous parameters ρ_i^+, ρ_i^- . From this analysis it will become clear how I am able to identify ρ_i 's on a game-by-game basis, and how more regressors could be accommodated. Intuitively, the propensity to leave a review on and off a discount (the like and dislike rates r_{it}^+, r_{it}^-) are identified by the differences in reviews left on and off a discount. In fact, this intuition becomes a rigorous proof, as will be shown shortly. The proof will also highlight the importance of having estimated the arrival rates of consumers.

Following Assumption 3.1, the arrival of buyers for game i on day t follows a Poisson distribution with arrival rate $\lambda_{it} = \lambda_i(1 + x'_{it}\beta)$. I assumed that, k days after the purchase is

made, each consumer leaves a positive review with probability r_{it}^+ , a negative review with probability r_{it}^- , and no review otherwise. Then, the following proposition is true:

Proposition 5.1 *The number of good reviews G_{it} for game i on day t is distributed Poisson with rate $r_{it}^+ \lambda_{it-k}$. The number of bad reviews B_{it} for game i on day t is distributed Poisson with rate $r_{it}^- \lambda_{it-k}$. Moreover, G_{it} and B_{it} are independent.*

Proposition 5.1 allows one to easily write down the likelihood for positive and negative reviews separately. I will use the positive reviews as the leading example here, but all the findings automatically translate to the case of negative reviews as well. For every game I observe the history of the review arrivals $\{(g_{it}, b_{it})\}_{t=1}^{T_i}$. Since $\mathbb{P}(G_{it} = g_{it}) = \frac{(r_{it}^+ \lambda_{it-k})^{g_{it}}}{g_{it}!} e^{-r_{it}^+ \lambda_{it-k}}$, the log-likelihood of the history of likes is given by

$$\ell(g_i; r_{it}^+, \lambda_{it-k}) = \sum_{t \geq k} g_{it} \log r_{it}^+ \lambda_{it-k} - r_{it}^+ \lambda_{it-k} - \log g_{it}! \quad (15)$$

For simplicity, I treat $\lambda_{it} = \lambda_i(1 + x'_{it}\beta)$ as known, rather than acknowledging the fact that I only have estimates $\hat{\lambda}_{it} = \hat{\lambda}_i(1 + x'_{it}\hat{\beta})$. By the virtue of observing the majority of the games for many periods, the estimates of λ_i should be relatively precise; estimates of β were estimated using all the observations in the sample, and should be relatively precise as well. For now I also take the lag k between the time when the user buys the game and leaves a review for the game to be known. The parameters of interest are $(\rho_{0i}^+, \rho_{1i}^+)$, which parametrize the like rate as $r_{it}^+ = \rho_{0i}^+ + \rho_{1i}^+ disc_{it-k}$. The likelihood is concave in the parameters, so the parameters are identified as the maximizer of the likelihood. Restricting attention to a binary discount variable $disc_{it}$ allows me to derive a closed-form estimator for $(\rho_{0i}^+, \rho_{1i}^+)$, which is convenient and illustrates the source of identification better. During a discount the like rate of game i is $\rho_{0i}^+ + \rho_{1i}^+$, which is consistently estimated as

$$\hat{\rho}_{0i}^+ + \hat{\rho}_{1i}^+ = \frac{\sum_{t \geq k} disc_{it-k} g_{it}}{\sum_{t \geq k} disc_{it-k} \lambda_{it-k}}, \quad (16)$$

As we can see, the ML-estimator of the like rate during a price promotion is just the ratio of

the number of good reviews left on discount over the expected sales on those days. Similarly, the like rate off discount can be estimated by

$$\hat{\rho}_{0i}^+ = \frac{\sum_{t \geq k} (1 - disc_{it-k}) g_{it}}{\sum_{t \geq k} (1 - disc_{it-k}) \lambda_{it-k}}, \quad (17)$$

with the interpretation that the like rate in the absence of a discount is just a ratio of the reviews left outside of a discount and the expected number of customers who paid the full price. Notice that identification of these parameters requires the knowledge of the demand parameters λ_{it} , highlighting the point I made earlier: seeing more good reviews arriving during a discount is not sufficient to infer that customers are more likely to leave positive reviews when purchasing during a discount. One needs to know how many more customers are buying because of the aforementioned discount.

My second hypothesis, formulated in (13), involves the sentiment of the reviews left on and off price promotion. Intuitively, one should be able to detect a change in the valence of reviews during a price promotion by simply comparing the ratios of good to bad reviews on and off discount. In particular, the answer should not depend on any information on the absolute sales. Indeed, if we plug in the estimators from (16)-(17) into the expression for the expected review score off discount from (13), we can see that our inference here will not depend on the estimates of the demand parameters in λ_{it} :

$$\frac{\hat{\rho}_{0i}^+}{\hat{\rho}_{0i}^+ + \hat{\rho}_{0i}^-} = \frac{\sum_{t \geq k} (1 - disc_{it-k}) g_{it}}{\sum_{t \geq k} (1 - disc_{it-k}) g_{it} + \sum_{t \geq k} (1 - disc_{it-k}) b_{it}} \quad (18)$$

The estimator of the review score off discount simply calculates the average of the positive reviews left off discount among all the reviews left off discount, confirming our intuition.

The analysis above is complicated by two facts. First, for each game I observe all reviews ever left, but I only have information on a fraction of buyers, as my measure of the number of buyers is based on the peak-time data only. To see what the problem is, consider the following example. Assume that every buyer of the game leaves a review. Suppose also that

every day there are two new buyers who play in the morning, and three new buyers who play in the evening. Both types abandon the game after just one day of playing. In such a world my data would register 3 players leaving 5 reviews on a daily basis, implying that one user leaves more than one review. This is the reason why I refer to the values of (r_{it}^+, r_{it}^-) as rates, rather than probabilities (which they are in the model). Proposition 5.1, stating that the number of good reviews on day t for game i is a Poisson random variable with rate $r_{it}^+ \lambda_{it-k}$, could be treated as an assumption, rather than a result. In that case, nothing constrains the r_{it}^+ parameter to be less than 1, and the identification argument goes through in exactly the same way.

Another problem stems from the fact that the purchase date behind a review is not available to the researcher. For that reason, the model features an additional parameter k , the lag between purchasing the game and leaving a review, that I assumed to be known. To estimate this parameter I leverage the fact that spikes in player activity on the first day of a sale represent new users, and study the review response during the following week in order to uncover the modal lag for leaving a review (see Table 9 in the Appendix). This number turns out to be one day. Even though time to posting a review has a non-degenerate distribution, modeling it in a more nuanced way would significantly contribute to complexity: the likelihood of receiving a positive review on day 100 would depend on the entire history of buyer arrivals prior to that date. Therefore, I proceed using $\hat{k} = 1$.

5.3 Estimation and Inference

I start off by estimating the average probabilities of leaving a good and a bad review in the population, and by using the estimates to test hypotheses formulated in (9) and (13). I use two sets of covariates to explain these probabilities. The baseline set of covariates, $w_{it} = [1, disc_{it-k}]'$, only allows the probability of a like or a dislike to change if a discount is in place. The theory behind this choice is that reviews reflect the average perception of quality by the consumers. As my sample was selected to only include products with

Table 5: Estimates of the Review Rates

	(1)	(2)
Constant (Like)	0.074*** (0.002)	0.067*** (0.001)
Discount (Like)	0.032*** (0.003)	0.017*** (0.001)
Constant (Dislike)	0.015*** (0.001)	0.013*** (0.000)
Discount (Dislike)	0.004*** (0.001)	0.002*** (0.000)
Young (Like)		0.296*** (0.015)
Old (Like)		-0.045 (0.001)
Young (Dislike)		0.067*** (0.006)
Old (Dislike)		-0.008 (0.000)
Observations	355077	355077

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$

constant quality, this perception should only change if different types of consumers are buying the product. Consumers buying on a discount are paying a lower price, and thus could be different from consumers paying the full price. This is the selection effect of discounts on reviews. I also acknowledge the possibility that there could be a selection over time, when early buyers are very different from late buyers. Similarly, the perception of quality might be deteriorating over time as new games, with potentially better characteristics, are coming out. For this reason, the second set of covariates also adds two variables measuring age: $w_{it} = [1, disc_{it-k}, young_{it}, old_{it}]'$, where $young_{it}$ is a dummy indicating that the game is less than 2 weeks old, and old_{it} is a dummy indicating that the game is more than a year old. The estimates are presented in Table 5.

Given the constraints on my data, it is hard to tell if the estimates that I obtain over- or underestimate the true probabilities of leaving a review. On the one hand, I can only estimate sales for a fraction of players, and thus the probability to leave a review could be severely overestimated. On the other hand, my model assumes that a user can only leave a

review on the next day after she purchases the product, so my estimates do not reflect the possibility that a user can leave a review at a later date. This, potentially, can underestimate the probability with which consumers leave reviews, especially during a discount¹². Thus, my estimates are better suited for making relative statements about the probabilities of interest.

Setting aside the aforementioned issues, my results suggest that a typical user on Steam leaves a positive review with a probability around 7.4%, and leaves a negative review with a probability around 1.5%. Discounts appear to increase both probabilities, implying that users seem to be more likely to leave a review during a discount. I formally test this hypothesis, stated in (9), and I strongly reject the null that users are less likely to leave reviews when buying on discount in favor of the alternative that they are more likely to leave reviews when buying on discount (the values of the t -stats are 9.16 and 12.48 in the two specifications). The estimated change in the probability of leaving a review due to a discount is between 24% and 40%, where the smaller estimate is probably more reliable, as it takes into account the effects of age on discounting. These results indirectly confirm that the variance effect could be playing a big role in the discounting decision of firms on Steam, as discounts are really effective at bringing new reviews.

My second hypothesis, stated in (13), is that reviews are less positive during a discount than they are outside of a discount. I fail to reject the null in favor of the alternative that reviews left during a price promotion are more positive (the p -values are 0.14 and 0.18). As we can see, there is no disagreement across specifications on both hypotheses, giving me confidence that this result is robust. This is also in accord with the results from the regression of the score on the lags of discounts and controls, presented in Table 9, which finds a small negative effect of discounts on the review score. Together these findings suggest that the selection effect of discounts is not big, and can not be the decisive force driving the discounting behavior close to review transitions.

To strengthen this point, I use the difference in the valence of reviews on and off discount

¹²Reviews arriving more than one day after the discount is over will not be counted towards reviews left during discount

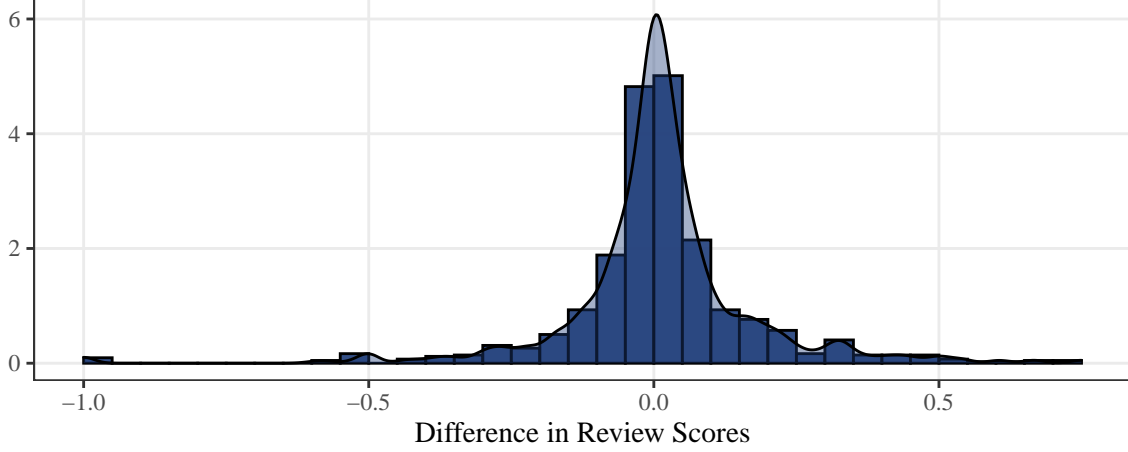


Figure 6: Distribution of the Difference Between Estimated Average Score on and off Discount.

to measure the selection effect at the game level. With such estimates at hand, I can revisit my regression of discounts on proximities to transitions (8), controlling for the selection effect. In the identification section I showed how one can estimate the like and dislike rates on the product level. Using these estimates, the selection effect can be measured as the difference between the expected review scores on and off a discount:

$$\Delta score_i := \frac{\rho_{0i}^+ + \rho_{1i}^+}{\rho_{0i}^+ + \rho_{1i}^+ + \rho_{0i}^- + \rho_{1i}^-} - \frac{\rho_{0i}^+}{\rho_{0i}^+ + \rho_{0i}^-} \quad (19)$$

Figure 6 depicts the distribution of $\widehat{\Delta score_i}$, the sample counterpart of this measure. In line with the hypothesis test of the sign of the selection effect in the sample, the histogram is centered around zero. Estimates of $\Delta score_i$ vary in their precision from game to game, as games that spent less time in the sample and/or on discounts would naturally have noisier estimates. However, the distribution appears to be well-behaved.

I use these estimates to modify (8) by including the interaction between the treatment variables of interest (proximities to potential transitions) and $\Delta score_i$, leading to

$$disc_{it} = (\beta_0^+ + \Delta score_i \beta_1^+) T_{it}^+ + (\beta_0^- + \Delta score_i \beta_1^-) T_{it}^- + X_{it} \beta + \Delta score_i + f_i + \tau_t + \varepsilon_{it} \quad (20)$$

As $\Delta score_i$ does not change over time, its effect on discounting probability can not be estimated by the within-estimator, but the interaction term could still be included. Notice that β_0^+ and β_0^- measure the effect of proximity to transition on discounting probability if $\Delta score_i = 0$, i.e. *if there is no selection effect*. The signs and significance of these coefficients offers another test of the importance of the variance effect. In the absence of the selection effect, a moderate amount of variance is good for a game close to a review upgrade, and is bad for a game close to a review downgrade. Thus, we expect to find $\beta_0^+ > 0$ and $\beta_0^- < 0$. The remaining coefficients β_1^+ and β_1^- measure how the selection effect changes the probability of running a price promotion close to a transition. Selection effect is good for the firms, and the prediction of my stylized model is $\beta_1^+ > 0$ and $\beta_1^- > 0$.

Table 6: Discounts Before Potential Transitions And Firm Heterogeneity

	<i>Dependent variable:</i>	
	$\mathbb{P}(Discount)$	
	OLS	FE
Pos. Transition	0.010*** (0.002)	0.014*** (0.003)
Neg. Transition	-0.010*** (0.002)	-0.009*** (0.003)
S. Diff \times Pos. Tr.	0.020** (0.008)	0.009 (0.013)
S. Diff \times Neg. Tr.	0.011 (0.009)	0.019 (0.015)
Score Difference	0.007 (0.007)	
Observations	290,628	290,628
R ²	0.437	0.434
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01	

I estimate (20) by OLS (without the fixed effect term f_i) and by the within-estimator. The results are presented in Table 6 (the controls are omitted for brevity). The least squares estimates in column (OLS) are mainly presented to check the correlations in the data. Games that tend to receive better reviews on discounts discount more. However, the selection effect

is not significant in both specifications. Crucially, in the preferred specification (FE) I find that $\beta_0^- < 0$, and $\beta_0^+ > 0$. Once the variation in the selection effect is controlled for, I obtain the coefficients measuring the impact of the variance effect on discounting decisions, and the signs align with theory. Their sizes are also quite economically meaningful, suggesting that, in the absence of the selection effect, games close to a positive transition are 8% more likely to go on a discount, and games close to a negative transition are 6% less likely to go on a discount. Recall that my estimates of (8) that did not control for the selection effect, found only a limited effect of the proximity to a negative transition on the probability of discount.

6 Conclusion

In this paper I document a novel mechanism through which firms use pricing to manage their online reviews. Using data from Steam, a major digital marketplace for computer video games, I show that firms discount their products in order to facilitate transitions between review tiers. My estimates of the demand process indicate that: (i) better review tiers substantially increase sales, (ii) that discounts are effective at attracting new users and, as a consequence, reviewers. Therefore, using price promotions to improve the review label of a game appears to be a plausible strategy. I find that products that are close to upgrading their review tier are 4-9% more likely to be on sale, while products close to deteriorating their review tier are up to 4% less likely to discount.

I identify two channels that can make such behavior profitable: the *selection* effect and the *variance* effect. Customers who purchase the good during a discount are likely to be different from the ones who pay the full price, and could potentially leave different reviews. If such reviews tend to be more favorable, then the induced positive selection effect creates incentives for the product to go on discount, regardless of whether the product is about to upgrade or downgrade its review tier. There is some disagreement in the literature about the sign of the selection effect. I contribute by showing that on Steam the effect is not positive.

The variance effect reflects the asymmetry between a successful and an unsuccessful discounting campaign, as measured by how favorable the new reviews are. A product close to moving one review tier up benefits greatly from a good realization, while a bad one might not change its review tier. On the other hand, a product that is close to sliding into a worse tier is averse to a bad outcome, and would benefit little from a good outcome. While I manage to test the selection effect directly, I have to test the variance effect by trying to falsify the predictions that follow from it. My strongest test is given by the prediction that, in the absence of a positive selection effect, the variance effect has to make a product on the verge of deteriorating its reviews less likely to discount. When I control for the size of the selection effect, I find that the aforementioned decline in probability of discounting goes from being 4% and weakly significant to 6% and highly significant. I also leverage my data and structural model to document that consumers are more likely to leave reviews when they buy during a discount, which indirectly supports the importance of the variance effect.

A Additional Tables and Figures

A.1 Descriptive Statistics

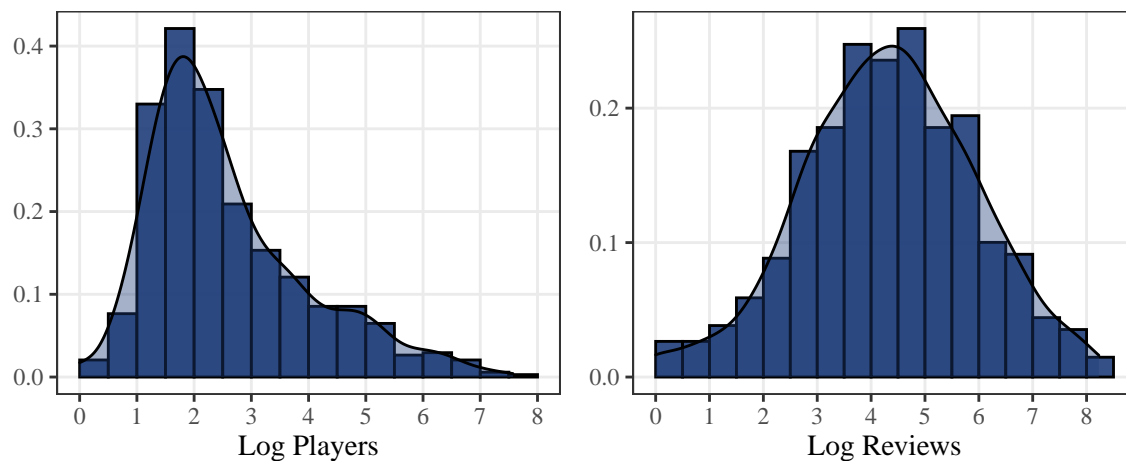


Figure 7: The Player Count on, and the Number of Reviews Received by, the Age of 180 Days.

Table 7: Average Discount, Age, and Review Count Two Weeks Before A Transition

Statistic	N	Mean	St. Dev.	Pctl(25)	Pctl(75)
Discount	1,225	10	14	0	16
Age	1,225	142	157	8	226
Reviews	1,225	130	343	26	81

A.2 Seasonal Sales

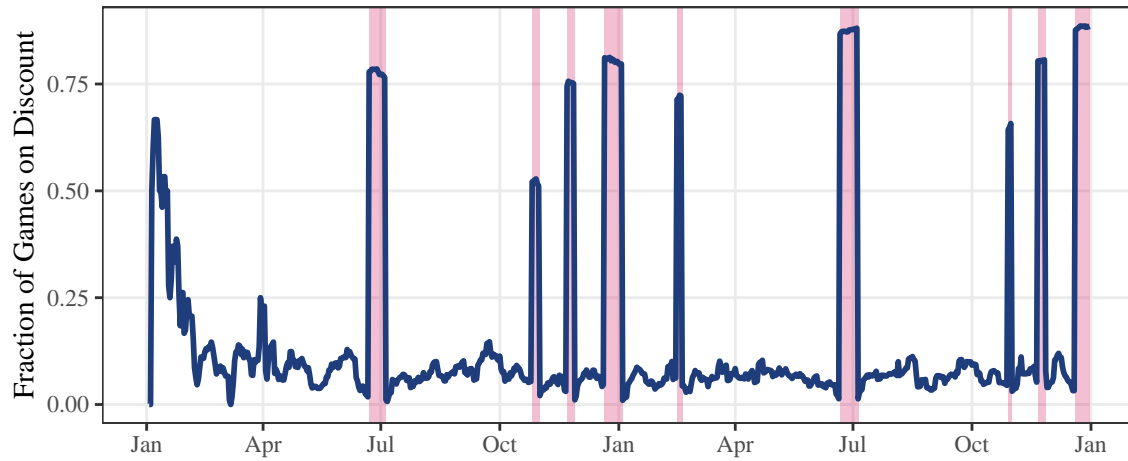


Figure 8: Seasonal Sales and Discounting Activity. Shaded regions are my definition of Seasonal Sales periods. The first spike on the graph is excluded because the sample starts with just one game, and I exclude the first 2 weeks from the calculations.

A.3 Discounts Around Transitions

Table 8: Discounts In The Days Around Transition

	<i>Dependent variable:</i>
	Discount Probability
Days to Transition	0.005*** (0.0005)
Days After Transition	-0.011*** (0.0005)
Negative	-0.007 (0.011)
Mostly Positive	0.005 (0.006)
Positive	0.044*** (0.009)
Very Positive	0.014 (0.009)
Ov. Positive	0.020 (0.016)
Score	-0.0005* (0.0003)
Log Reviews	-0.010*** (0.002)
Age	0.0003*** (0.00001)
Const	0.903*** (0.021)
Observations	32,112
R ²	0.396
Time Effects	Weekdays, Week
Game Effects	×
Polynomial(t W/O Discount)	$d = 2$
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

A.4 Response Of Reviews and Score to Discounts

Table 9 presents the results of estimating

$$y_{it} \sim \sum_{l=0}^7 \mathbf{1}(\text{Disc. } l \text{ days ago}) + \text{Controls}_{it} + f_i + \tau_t + u_{it}$$

for y_{it} being the number of new reviews for game i on day t , the ratio of new reviews to the average speed of review arrival, and game i 's review score. I use all the same controls as usual: review bin, score (in the first regression), log reviews, age, dummy for being less than 14 days old.

Table 9: Changes in Review Flows and Score After a Discount

	<i>Dependent variable:</i>		
	New Reviews	New Reviews(%)	Score (0-100)
Discount 0 Days Ago	1.094*** (0.122)	1.801*** (0.135)	-0.159 (0.152)
Discount 1 Day Ago	3.254*** (0.276)	4.282*** (0.228)	-0.557*** (0.184)
Discount 2 Days Ago	1.377*** (0.139)	1.969*** (0.100)	-0.470*** (0.176)
Discount 3 Days Ago	0.743*** (0.110)	1.197*** (0.072)	-0.370** (0.170)
Discount 4 Days Ago	0.622*** (0.096)	0.847*** (0.066)	-0.401** (0.166)
Discount 5 Days Ago	0.418*** (0.104)	1.023*** (0.125)	-0.395** (0.161)
Discount 6 Days Ago	0.071 (0.086)	0.396*** (0.063)	-0.368** (0.159)
Discount 7 Days Ago	-0.120* (0.072)	0.254*** (0.076)	-0.336** (0.155)
Weekdays + Week Effects	✓	✓	✓
Game Effects	✓	✓	✓
Observations	350,819	350,819	350,819
R ²	0.059	0.105	0.154

Note:

*p<0.1; **p<0.05; ***p<0.01

B Mathematical Appendix

B.1 NLLS estimator of β

In this appendix I develop the estimator for the parameters of the model

$$y_{it} = \psi_i y_{it-1} + \lambda_i(1 + x'_{it}\beta) + u_{it}, \quad (21)$$

with the moment condition $\mathbb{E}[u_{it} | y_{it-1}, x_{it}] = 0$. This model has $2n$ i -specific parameters (λ_i, ψ_i) and $\dim(\beta)$ parameters that are shared by all entities in the panel. Estimation is via Non-linear Least Squares. The F.O.C. of the problem could be reduced so as to concentrate out all (λ_i, ψ_i) parameters. Thus, estimation of $2n + \dim(\beta)$ parameters reduces to solving a system of non-linear equations for $\dim(\beta)$ parameters (or solving a NLLS problem of that dimension).

The estimator is defined as the minimizer of the sum of *weighted* squared errors:

$$\hat{\theta} := \underset{\lambda_i, \psi_i, \beta}{\operatorname{argmin}} \sum_{i=1}^n \sum_{t=2}^{T_i} w_i (y_{it} - \psi_i y_{it-1} - \lambda_i(1 + x'_{it}\beta))^2 \quad (22)$$

The reason for using weights w_i comes from the large disparities in the sizes of games in my sample. A on observation with $y_{it} = 12$, when the model predicts 10, contributes 2^2 to the sum of squares, while an observation with $y_{it} = 120$ instead of 100 contributes 20^2 . I use $w_i = 1/\max_t y_{it}$ as weights, forcing the dependent variable to be between 0 and 1. The F.O.C. are

$$\psi_i : \sum_{t=2}^{T_i} (y_{it} - \psi_i y_{it-1} - \lambda_i(1 + x'_{it}\beta)) y_{it-1} = 0 \quad (23)$$

$$\lambda_i : \sum_{t=2}^{T_i} (y_{it} - \psi_i y_{it-1} - \lambda_i(1 + x'_{it}\beta))(1 + x'_{it}\beta) = 0 \quad (24)$$

$$\beta : \sum_{i=1}^n \sum_{t=2}^{T_i} w_i (y_{it} - \psi_i y_{it-1} - \lambda_i(1 + x'_{it}\beta)) \lambda_i x_{it} = 0 \quad (25)$$

Notice that the weights drop out of the first two conditions: as I scale all observation within a product by the same weight, it has no effect on the identification of the product specific parameters λ_i and ψ_i (conditional on β). As for the last condition, observe that the moment condition $\mathbb{E}[u_{it} | y_{it-1}, x_{it}] = 0$ implies $\mathbb{E}[x_{it}u_{it}] = 0$ under the true parameter values. Thus, the sum in the F.O.C. converges to 0 as $(\min_i T_i) \rightarrow \infty$ under any set of weights:

$$\frac{1}{n} \sum_{i=1}^n \frac{1}{T_i} \sum_{t=2}^{T_i} w_i (y_{it} - \psi_i y_{it-1} - \lambda_i (1 + x'_{it} \beta)) \lambda_i x_{it} = \frac{1}{n} \sum_{i=1}^n \lambda_i w_i \left(\frac{1}{T_i} \sum_{t=2}^{T_i} u_{it} x_{it} \right) \xrightarrow{p} 0$$

In other words, the true value of β is identified by the weighted moment conditions.

Conditional on β , the F.O.C. for (ψ_i, λ_i) are F.O.C.'s of an OLS problem of the form

$$\min_{\psi_i, \lambda_i} \sum_{t=2}^{T_i} (y_{it} - \psi_i y_{it-1} - \lambda_i z_{it})^2,$$

where $z_{it} = (1 + x'_{it} \beta)$. Defining, in the standard way, $\tilde{\mathbf{X}}_i(\beta)$ to be the matrix with row t given by $[z_{it}, y_{it-1}]$, and y_i to be the vector of y_{it} observations, we get that the values of $\hat{\lambda}_i, \hat{\psi}_i$ that solve (23)-(24) are given by

$$\begin{bmatrix} \hat{\lambda}_i \\ \hat{\psi}_i \end{bmatrix}(\beta) = (\tilde{\mathbf{X}}'_i(\beta) \tilde{\mathbf{X}}_i(\beta))^{-1} \tilde{\mathbf{X}}'_i(\beta) y_i \quad (26)$$

The estimator $\hat{\beta}$ of β is now obtained as a solution to

$$\sum_{i=1}^n \sum_{t=2}^{T_i} w_i (y_{it} - \psi_i(\beta) y_{it-1} - \lambda_i(\beta) (1 + x'_{it} \beta)) \lambda_i(\beta) x_{it} = 0 \quad (27)$$

In practice I solve the concentrated minimization problem using (27) as the gradient.

Supplying an analytic expression for the Hessian matrix has proven to significantly expedite and improve convergence. The concentrated out sum of squared errors is $SSE(\beta) = SSE(\beta, \lambda(\beta), \psi(\beta))$, and while the expression for the gradient simplifies to $\partial SSE / \partial \beta$ due to

the envelope theorem, the expression for the Hessian is more complex:

$$\frac{dSSE}{d^2\beta} = \frac{\partial SSE}{\partial^2\beta} + \sum_i \frac{\partial SSE}{\partial \lambda_i \partial \beta} \left(\frac{d\lambda_i}{d\beta} \right)' + \sum_i \frac{\partial SSE}{\partial \psi_i \partial \beta} \left(\frac{d\psi_i}{d\beta} \right)' \quad (28)$$

The expressions for $\frac{d\lambda_i}{d\beta}$ and $\frac{d\psi_i}{d\beta}$ are obtained by applying the inverse function theorem to (26), and the remaining derivatives could be obtained directly.

The asymptotic covariance matrix for $\hat{\theta}$ in a nonlinear cross-section LS problem $\sum_i (y_i - m(x_i, \theta))^2$ is given by the sample analog of

$$\mathbf{V}_\theta = (\mathbb{E}[m_{\theta i} m'_{\theta i}])^{-1} \mathbb{E}[m_{\theta i} m'_{\theta i} e_i^2] (\mathbb{E}[m_{\theta i} m'_{\theta i}])^{-1}, \quad (29)$$

where $m_{\theta i} = \frac{\partial}{\partial \theta} m(x_i, \theta_0)$, $e_i = y_i - m(x_i, \theta_0)$ [Hansen, 2020, p.751]. I extend this approach to my panel setting as follows¹³. For product i the concentrated conditional expectation function and its derivative are given by

$$m(y_{it-1}, x_{it}, \beta) = \psi_i(\beta) y_{it-1} + \lambda_i(\beta) (1 + x'_{it} \beta) \quad (30)$$

$$m_{\beta it} := \frac{\partial}{\partial \beta} m(y_{it-1}, x_{it}, \beta) = \frac{\partial \psi_i(\beta)}{\partial \beta} y_{it-1} + \frac{\partial \lambda_i(\beta)}{\partial \beta} (1 + x'_{it} \beta) + \lambda_i(\beta) x_{it}. \quad (31)$$

I then make the following analogy between the (sample analogs of) cross-section expressions in (29) and their panel data analogs

$$\frac{1}{n} \sum_{i=1}^n m_{\theta i} m'_{\theta i} \sim \frac{1}{n} \sum_{i=1}^n \frac{1}{T_i} \sum_{t=2}^{T_i} m_{\beta it} m'_{\beta it}$$

The resulting expression for the estimator of the variance of $\hat{\beta}$ is

$$\left(\frac{1}{n} \sum_{i=1}^n \frac{1}{T_i} \sum_{t=2}^{T_i} m_{\hat{\beta} it} m'_{\hat{\beta} it} \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^n \frac{1}{T_i} \sum_{t=2}^{T_i} m_{\hat{\beta} it} m'_{\hat{\beta} it} \hat{u}_{it}^2 \right) \left(\frac{1}{n} \sum_{i=1}^n \frac{1}{T_i} \sum_{t=2}^{T_i} m_{\hat{\beta} it} m'_{\hat{\beta} it} \right)^{-1}, \quad (32)$$

where $\hat{u}_{it} = y_{it} - \hat{\psi}_i y_{it-1} - \hat{\lambda}_i (1 + x'_{it} \hat{\beta})$.

¹³I rescaled the data by the weights w_i for estimation, so the expressions below do not feature the weights

B.2 Proofs

Proposition 5.1. Let the total number of reviewers during a day, R , be distributed Poisson with rate λ , $R \sim P(\cdot; \lambda)$. A reviewer leaves a good review with probability π , and a bad review with probability $1 - \pi$. Then, the numbers of good and bad reviews, G and B , are independent Poisson random variables with rates $\pi\lambda$ and $(1 - \pi)\lambda$.

In the main text I have a Poisson arrival of customers, and only a fraction of them become reviewers, but the same proof goes through with that modification as well. The independence between the numbers of good and bad reviews is the surprising part of the proposition, and it is better highlighted in the form presented here.

Proof: We start by showing that G is Poisson.

$$\mathbb{P}(G = g) = \sum_{n=g}^{\infty} \mathbb{P}(G = g \mid R = n) \mathbb{P}(R = n) = \sum_{n=g}^{\infty} \frac{n!}{g!(n-g)!} \pi^g (1-\pi)^{n-g} \frac{\lambda^n}{n!} e^{-\lambda}$$

We leave only the parts that depend on n within the sum:

$$e^{-\lambda} \frac{\pi^g}{g!} \sum_{n=g}^{\infty} \frac{(1-\pi)^{n-g} \lambda^n}{(n-g)!} = e^{-\lambda} \frac{\pi^g \lambda^g}{g!} \sum_{n=g}^{\infty} \frac{(1-\pi)^{n-g} \lambda^{n-g}}{(n-g)!} = e^{-\pi\lambda} \frac{\pi^g \lambda^g}{g!} \sum_{k=0}^{\infty} \frac{[(1-\pi)\lambda]^k}{k!} e^{-(1-\pi)\lambda}$$

The sum we have is just $\sum_{k=0}^{\infty} P(k; \lambda) = 1$, so the answer is

$$\mathbb{P}(G = g) = \frac{[\pi\lambda]^g}{g!} e^{-\pi\lambda} = P(g; \pi\lambda)$$

Now the independence part. We are interested in $\mathbb{P}(G = g, B = b)$, which could be written as

$$\mathbb{P}(G = g, B = b) = \mathbb{P}(G = g, B = b \mid R = g + b) \mathbb{P}(R = g + b) = \frac{(g+b)!}{g!b!} \pi^g (1-\pi)^b \frac{\lambda^{g+b}}{(g+b)!} e^{-\lambda}$$

Simply write $e^{-\lambda} = e^{-\pi\lambda} e^{-(1-\pi)\lambda}$, and collect the terms with g and with b to get

$$\mathbb{P}(G = g, B = b) = \frac{\pi^g \lambda^g}{g!} e^{-\pi\lambda} \frac{(1-\pi)^b \lambda^b}{b!} e^{-(1-\pi)\lambda} = P(g; \pi\lambda) P(b; (1-\pi)\lambda) \triangleleft$$

References

- Daron Acemoglu, Ali Makhdoui, Azarakhsh Malekian, and Asu Ozdaglar. Learning from reviews: The selection effect and the speed of learning. June 2019. [Cited on page 33.]
- Michael Anderson and Jeremy Magruder. Learning from the crowd: Regression discontinuity estimates of the effects of an online review database. *The Economic Journal*, 122(563): 957–989, 2012. [Cited on page 6.]
- Christina Aperjis and Ramesh Johari. Optimal windows for aggregating ratings in electronic marketplaces. *Management Science*, 56(5):864–880, May 2010. [Cited on page 5.]
- James Batchelor. Steam removes hundreds of games after publishers abuse steamworks, 2019. [Cited on page 10.]
- John W. Byers, Michael Mitzenmacher, and Georgios Zervas. Thegroupon effect on yelp ratings. *Proceedings of the 13th ACM Conference on Electronic Commerce - EC '12*, 2012. [Cited on pages 8 and 33.]
- Luís Cabral and Lingfang (Ivy) Li. A dollar for your thoughts: Feedback-conditional rebates on ebay. *Management Science*, 61(9):2052–2063, Sep 2015. [Cited on pages 8 and 33.]
- Yeon-Koo Che and Johannes Hörner. Recommender systems as mechanisms for social learning. *The Quarterly Journal of Economics*, 133(2):871–925, 2018. [Cited on page 5.]
- Judith A. Chevalier and Dina Mayzlin. The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, 43(3):345–354, Aug 2006. [Cited on pages 5 and 12.]
- Judith A. Chevalier, Yaniv Dover, and Dina Mayzlin. Channels of impact: User reviews when quality is dynamic and managers respond. *Marketing Science*, 37(5):688–709, Sep 2018. [Cited on page 7.]
- Edwards Cliff. Valve lines up console partners in challenge to microsoft, sony, 2013. [Cited on page 10.]
- Bart de Langhe, Philip M. Fernbach, and Donald R. Lichtenstein. Navigating by the stars: Investigating the actual and perceived validity of online user ratings. *Journal of Consumer*

- Research*, 42(6):817–833, 09 2015. [Cited on page 12.]
- Utpal M. Dholakia. How businesses fare with daily deals: A multi-site analysis ofgroupon, livingsocial, opentable, travelzoo, and buywithme promotions. *SSRN Electronic Journal*, 2011. [Cited on page 9.]
- George W. Downs and David M. Rocke. Conflict, agency, and gambling for resurrection: The principal-agent problem goes to war. *American Journal of Political Science*, 38(2):362–380, 1994. [Cited on page 9.]
- Entertainment Software Association. Breakdown of u.s. computer and video game sales from 2009 to 2017, by delivery format, May 2019. [Cited on page 10.]
- Kristopher Floyd, Ryan Freling, Saad Alhoqail, Hyun Young Cho, and Traci Freling. How online product reviews affect retail sales: A meta-analysis. *Journal of Retailing*, 90, 06 2014. [Cited on page 5.]
- Bin Gu and Qiang Ye. First step in social media: Measuring the influence of online management responses on customer satisfaction. *Production and Operations Management*, 23(4):570–582, 2014. [Cited on page 7.]
- Bruce H. Hansen. Econometrics. Manuscript, 2020. [Cited on page 59.]
- Bar Ifrach, Costis Maglaras, Marco Scarsini, and Anna Zseleva. Bayesian social learning from consumer reviews. *Operations Research*, 67(5):1209–1221, September 2019. [Cited on page 33.]
- Xitong Li. Could deal promotion improve merchants’ online reputations? the moderating role of prior reviews. *Journal of Management Information Systems*, 33(1):171–201, Jan 2016. [Cited on pages 8 and 33.]
- Michael Luca. Reviews, reputation and revenue: The case of yelp.com. *Working Paper*, 2016. [Cited on page 6.]
- Michael Luca and Georgios Zervas. Fake it till you make it: Reputation, competition, and yelp review fraud. *Management Science*, 62(12):3412–3427, Dec 2016. [Cited on page 7.]
- Dina Mayzlin, Yaniv Dover, and Judith Chevalier. Promotional reviews: An empirical

investigation of online review manipulation. *American Economic Review*, 104(8):2421–55, August 2014. [Cited on page 7.]

Mike Minotti. Npd: U.s. game sales hit a record \$43.4 billion in 2018, 2019. [Cited on page 10.]

PC Games. With \$4.3 billion in sales, 2017 was steam’s biggest year yet. PSGames, 2017. [Cited on page 10.]

PC Games. Valve shuts down another way to estimate steam sales, and indie devs aren’t happy. PSGames, 2018. URL <https://www.pcgamesn.com/steam-sales-estimates>. [Cited on page 12.]

Davide Proserpio and Georgios Zervas. Online reputation management: Estimating the impact of management responses on consumer reviews. *Marketing Science*, 36(5):645–665, Sep 2017. [Cited on page 7.]

Imke C Reimers and Joel Waldfogel. Digitization and pre-purchase information: The causal and welfare impacts of reviews and crowd ratings. Technical report, National Bureau of Economic Research, 2020. [Cited on pages 6 and 29.]

Dmitry Sorokin and Ryan Louis Stevens. The impact of online reviews on product sales: Evidence from video games. February 2020. [Cited on pages 6, 7, 13, and 22.]

Steamworks Documentation. Positive "review bombs", May 2019. [Cited on page 12.]

Steamworks Documentation. Discounting, 2020a. [Cited on page 16.]

Steamworks Documentation. Visibility on steam, July 2020b. [Cited on page 11.]

Jean Tirole. *Economics for the common good*. Princeton University Press, 2017. [Cited on page 2.]

Nikhil Vellodi. Ratings design and barriers to entry. *Working Paper*, 2020. [Cited on page 5.]

Karen L. Xie, Kevin Kam Fung So, and Wei Wang. Joint effects of management responses and online reviews on hotel financial performance: A data-analytics approach. *International Journal of Hospitality Management*, 62:101 – 110, 2017. [Cited on page 7.]

Dainis Zegners. Building an online reputation with free content: Evidence from the e-book

market. *SSRN Electronic Journal*, 2017. [Cited on page 9.]

Dong Hong Zhu, Zhi Jie Zhang, Ya Ping Chang, and Shichang Liang. Good discounts earn good reviews in return? effects of price promotion on online restaurant reviews. *International Journal of Hospitality Management*, 77:178 – 186, 2019. [Cited on pages 8 and 33.]

Feng Zhu and Xiaoquan (Michael) Zhang. Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing*, 74(2): 133–148, 2010. [Cited on page 6.]