Could it be Better to Refund than Recommend? The Role of Ex-Post Match Values in Digital Entertainment Markets

Hans Zhu*

November 1, 2024

Abstract

Digital entertainment markets are characterised by many niches and the 'long tail' of consumption. I argue that the size of these tails may be dramatically underestimated if ex-post match values - match values realised only after consumption - are not taken into account. I collect a novel individual-level dataset on the Steam platform and interpret playtime data as directly informative about ex-post match values. The playtime data reveal patterns of consumption starkly different to those inferred from purchases. I then examine the policy relevance of ex-post match values, by studying a unique series of policy changes on the Steam platform. The introduction of personalised stores followed by the introduction of refunds. Personalised stores should improve consumer matches on average, but are unable to recommend on the basis of realised ex-post match values. Refunds allow consumers to realise ex-post match values before committing to the purchase. I find that sales increase following the introduction of refunds by around 5 times the increase following the introduction of personalised stores. In addition, the sales patterns following the introduction of refunds are dramatically different compared to those with the introduction of personalised stores. These results suggest that the effects of personalised stores alone are limited. Whether personalised stores matter or are complementary to refunds will be investigated using a structural model which I describe.

*Northwestern Univeristy; hanszhu2024@u.northwestern.edu

Many thanks to my advisors Rob Porter, David Dranove, Gaston Illanes and Vivek Bhattacharya for their generous guidance and advice. I also thank Hyein Cho, José D. Salas, Jin Yang, Jason Zhao and participants in Northwestern IO lunch.

1 Introduction

Neiman and Vavra (2023) show that niche consumption is surprisingly ubiquitous. Consumers of all demographics and all geographies in the US are spending more and more on their own 'niches'. In IO and marketing, this is often described in terms of the 'long tail' benefit of online retail Anderson (2004). Evidently, horizontal match values are highly relevant for consumer welfare. However in markets for durable products, since we only observe purchase decisions, we are unable to learn much about the role of *ex-post* match values - match values learned only after consumption. If ex-post match values play a significant role, then they can make a big difference to consumer welfare and findings about the long tail could be quite different when these are taken into account. I study the role of ex-post match values on the Steam platform, the dominant platform for digital computer games. To do so, I construct a novel individual-level dataset, which allows me to infer individual purchases and track the playtime of purchased games[1].To the best of my knowledge, I have not seen papers which have been able to explore durable product usage in leisure good markets[2] I find extraordinary variation in playtime which would seem impossible to rationalise without ex-post match values. The 90th and 95th percentile average playtime are more than ten and twenty-five times the median average playtime. And even when finding pairs of games that appear to be ex-ante similar, and both purchased by the same player, it is common to find highly lopsided playtime. These results suggest that playtime also exhibits long tail behaviour, and if we were to take into account ex-post match values the long tail in benefits from product variety might be even larger than previously understood.

Ex-post match values are also policy relevant. Given the long tail benefit of online retail, policy or tools that improve matches could increase consumer welfare substantially. In the context of digital platforms, recommendation algorithms are the predominant focus of the literature. While recommendation algorithms play an important role, they can only recommend based on observables that can be elucidated as data. In settings such as mine, it's impossible to fully characterise products, including features that could potentially be first-order important in determining a product's utility. For example, in games with a lot of action, how the game 'feels' to play is important. This depends on many aspects. Do players have sufficient freedom and complexity in movement? Are interactions with inanimate objects satisfying? Do the sound effects and visuals adequately convey the impact of movement and collisions? I regard these as 'captured' by ex-post match val-

[1] While I need to make assumptions to infer individual purchases which I will describe and verify, compared to other empirical papers studying digital platforms that use publicly available data such as Lee and Musolff (2023), I do not infer sales using a rank-based extrapolation

[2] Crawford et al. (2018) and Yurukoglu and Crawford (2012) have access to individual viewership data for television markets. However, there consumers do not purchase particular shows thus the framework in this paper does not apply to their context.

ues. Since recommendation algorithms cannot recommend on this basis, policies and tools that make it easier for consumers to discover their matches can also be important. I utilise a unique policy setting involving two sequential policy changes on the Steam platform to study these ideas. One policy introduces personalised stores, while a subsequent policy introduces a generous refund policy. I model refunds as allowing consumers to explore and realise their ex-post match values before committing to a purchase. All else equal, a sequential search model suggests that consumers should search products with higher variance of ex-post match values for a higher chance of getting a good draw. I show that sales in aggregate increase with the introduction of refunds and by around 5 times that of the increase in sales following the introduction of personalised stores. Most importantly, the patterns of sales following the introduction of refunds are also consistent with this model. The increase in sales are predominantly from games with the highest uncertainty in playtime. This is quite distinct from the sales patterns following the introduction of personalised stores which are much noisier. Taken together, the effects of personalised stores alone seem limited. Since recommendations could also change with the introduction of refunds - a recommendation algorithm could take into account the variance of ex-post match values (via variance in playtime), reduced-form evidence is insufficient to determine whether personalised stores are limited in their effects entirely, or if they are complementary to refunds.

In addition to determining the value of personalised stores in this setting, a structural model is also required to quantify the extent of consumer welfare arising from ex-post match values and construct the 'long tail' of benefits. Given the importance of a structural model, I provide a sketch and leave estimation for future work and drafts. I intend to model consumer demand as a discrete choice problem subject to limited consideration as in Goeree (2008). The period before either of the policy changes allows me to identify and estimate parameters of demand that do not change throughout the policy regimes. I do not directly observe individual personalised stores, but I know factors that influence personalised recommendations and observe the data for many of those. This means that with an assumption of linear separability, the effect of the personalised stores can be captured in a reduced-form way in how they shift consideration set probabilities. Finally to model the period following the introduction of refunds, I intend to follow the ideas of Moraga-González et al. (2023) who leverage the results of Armstrong (2017) and Choi et al. (2018) to estimate a discrete choice demand model with sequential search[3]. To model ex-post match values I intend to draw on the ideas of Bhat (1995) and allow for the variance of the logit errors to be parameterised instead of normalised.

To fix ideas for the rest of the paper, consider a simple discrete choice model of demand with

[3] The key contribution of Armstrong (2017) and Choi et al. (2018) is to provide a tractable characterisation of demand with sequential search even with heterogeneous products and consumers. In particular, they are able to characterise a consumer's demand by a simple set of inequalities in the flavour of a standard discrete choice demand model.

risk-neutral consumers[4] and ex-ante and ex-post match values. Risk-neutral consumer i is faced with alternatives j which provide utility:

$$U_{ij} = \underbrace{\delta_{ij}}_{\text{Observed utility}} + \underbrace{\varepsilon_{ij}}_{\text{Ex-ante match value}} + \underbrace{\omega_{ij}}_{\text{Ex-post match value}}$$

Both, ε_{ij} and ω_{ij} are mean zero, type 1 EV shocks with variance σ_j^{ε} and σ_j^{ω} respectively. ε_{ij} is the ex-ante match value that consumers form when they view the game and can always be inferred by revealed preference. ω_{ij} is the ex-post match value which consumers only learn after purchasing and playing the game. Conditional on δ_{ij}, ε_{ij} and ω_{ij} are independent of each other and everything else. σ_j^{ε} and σ_j^{ω} are both observed by consumers. Independence is important for tractability. Allowing for arbitrary correlation dramatically increases the number of parameters to estimate. Critically, the key results of Armstrong (2017) and Choi et al. (2018) also depend on the independence of the match values. Dependence is likely to make the consumer's optimal search strategy path-dependent, making such a model no longer empirically tractable. For ω_{ij} which are most of interest, independence seems reasonable if interpreted as something difficult to predict as described previously.

If σ_j^{ω} is relatively large, ω_{ij} could be highly welfare relevant, yet also completely irrelevant to decisions. Since it is mean 0 and consumers do not observe it ex-ante, they purchase a product based on δ_{ij} and ε_{ij} instead. Sellers cannot price differently either. They know that some consumers could realise a very large draw of ω_{ij}, but even so could never price this as long as consumers are risk-neutral. In an entertainment goods market j is very large, thus it is natural to incorporate limited consideration. With limited consideration, questions about platform design and recommendation systems become relevant. In this model however, if ω_{ij} is independent, then the recommendation system is just a consumer who can observe all products. It can recommend consumers towards products with higher ex-ante utility, which should lead to higher ex-post utility on average, however there could still be large gains left on the table. As ω_{ij} are independent and mean 0, the recommendation system shouldn't 'bias' consumers systematically away from the best matches, but it does not have the information to help consumers match on ω_{ij}.

With the introduction of refunds, consumers have the ability to 'search' or try games before committing to a purchase. That is, they can undertake a sequential search, incurring a search cost c_i each time to learn an ω_{ij}. Although ω_{ij} are mean 0, σ_j^{ω} will influence a consumer's optimal search sequence, since it affects the expected marginal benefit of searching the game. Following the introduction of refunds, we should expect consumers to shift their final purchases towards games with higher σ_j^{ω}.

[4] Since nothing in my data would allow me to separately identify risk preference I will not address this

2 The PC Games Market and Steam

The video gaming industry is the largest entertainment industry in the world. In 2022, global gaming industry revenues were estimated to be $184.4 billion. This is more than triple the combined global music and movie industry revenues (Arora, 2022)[5]. Steam is an online platform which operates in the digital PC games segment of the global gaming market. The total (including both physical and digital) PC games segment is just over 20% of the global games market (Wijman, 2013)[6] and the proportion of PC games sales which are digital is 98% (Batchelor, 2022) [7]. As Steam is a privately owned company, estimates of their market share vary significantly. One article from 2011 suggested a market share in digital computer games from half to 70 percent (Chiang, 2011) [8]. More recently Tim Sweeney, CEO of Epic Games which also operates a direct competitor - Epic Games Store - suggested a market share of 85% in 2023. Redditors compiling older comments from Sweeney infer that this market share has not varied much for a long time[9]. Given Steam's apparent market dominance and the lack of data availability I will ignore the role of all other platforms and stores where consumers may purchase digital PC games throughout the paper.

Steam as a digital store

For structural estimation, I have to deal with or take advantage of aspects of a typical consumer's experience, so in this section, I will describe some relevant features. First, like in many retail markets consumers have a very large space of potential products to deal with. I observe data from the beginning of 2014 to March 2018. At the start of the sample, I observe 1,911 games, while at the beginning of 2018 I observe 13,083 games[10]. To help consumers navigate the extensive catalogue there are several store features. When consumers first access the Steam Store they encounter the front page. The front page displays games with their titles and a cover picture, their price and the current discounts if there are any[11]. Games make it onto the front page for a variety of reasons. They may be trending, new or on sale. In my data, on average 43.7 games are displayed on the front page each day. As shown later in section 3, being on the front page is associated with signif-

[5] Accessed at https://www.forbes.com/sites/forbesagencycouncil/2023/11/17/the-gaming-industry-a-behemoth-with-unprecedented-global-reach/?sh=17abfce512fb

[6] Accessed at https://newzoo.com/resources/blog/video-games-in-2023-the-year-in-numbers

[7] Acessed at https://www.gamesindustry.biz/gamesindustrybiz-presents-the-year-in-numbers-2022

[8] Accessed at https://www.forbes.com/forbes/2011/0228/technology-gabe-newell-videogames-valve-online-mayhem.html?sh=541fd8523ac0

[9] See the reddit thread here at https://www.reddit.com/r/fuckepic/comments/17iypv0/egs_has_no_market_share_growth_in_three_years/

[10] This is not quite the universe of games, but is the majority. According to SteamDB - at https://steamdb.info/stats/releases/ - at the beginning of 2014 there were 2,118 games and 18,251 games at the beginning of 2018.

[11] For convenience, here's a link to the current Steam frontpage https://store.steampowered.com/

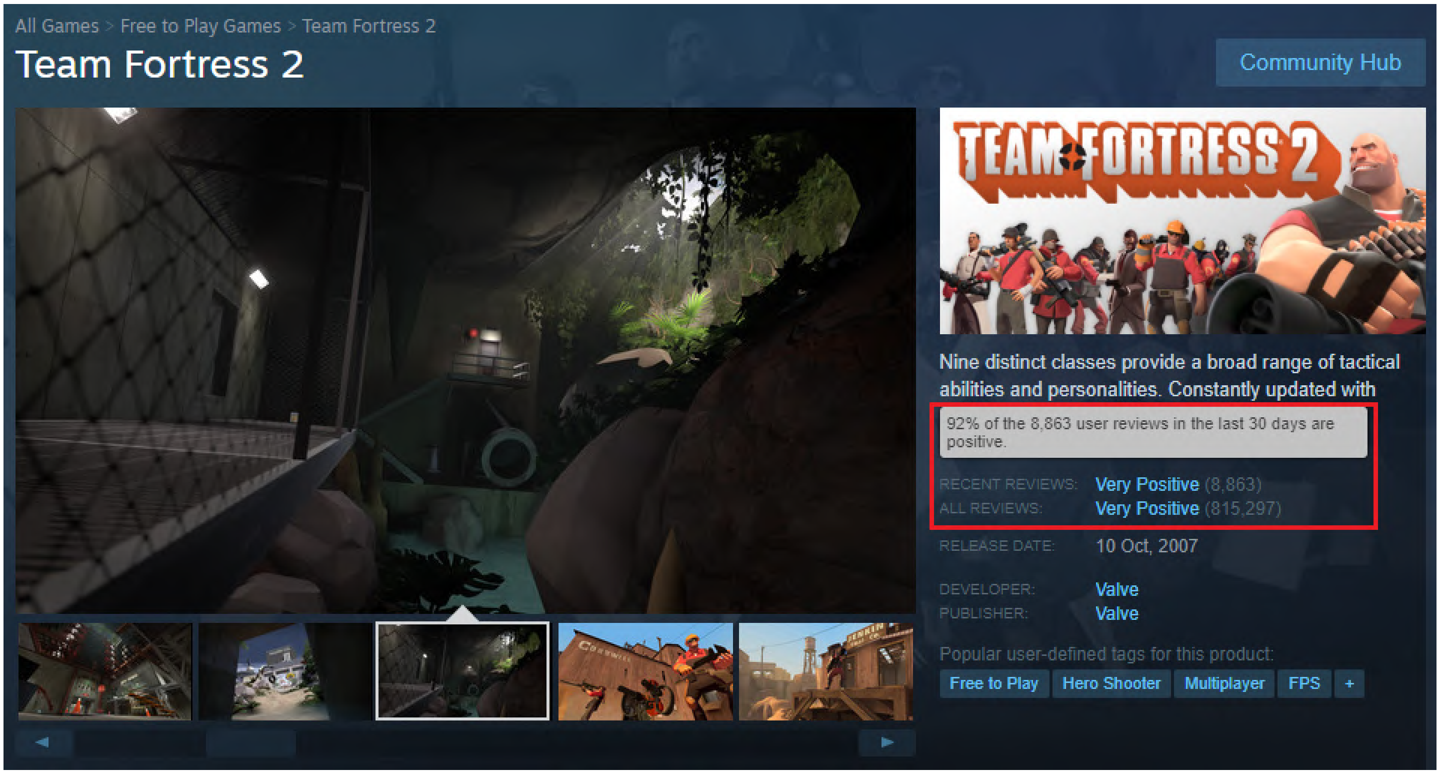

Figure 1: Example of review score and label presentation

icantly larger sales which will provide useful variation for the structural model. Consumers can use the search bar or click through pages featuring games in particular genres. Whatever the case, when a consumer clicks on a particular game they are taken to a game's individual page which presents information about the game in a consistent and accessible way[12]. I collect many game characteristics from these pages.

[12]One example https://store.steampowered.com/app/255710/Cities_Skylines/

'Review Label'	Review Score	Reviews Needed	# Games 2018-02-01
Overwhelmingly Positive	95+	500	337
Very Positive	85+	50	3264
Positive	80+	10	1681
Mostly Positive	70-80	10	1982
Mixed	40-70	10	2644
Mostly Negative	20-40	10	485
Negative	0-20	10	63
Very Negative	0-20	50	30
Overwhelmingly Negative	0-20	500	4

Note: Games without 10 reviews do not receive a review label. On 2018-02-01 there were 2881 games in my data with no label

Table 1: Steam's review label 'function'

An important characteristic is the game's 'review label' and 'review score'. Steam allows users to submit written reviews for games along with a 'recommended' or 'not recommended' binary rating[13]. A review score can then be calculated as the proportion of reviewers who recommend the game. A 'review label' is then constructed according to the rules specified in table 1. Review labels are a function of both the review score and volume of reviews submitted. A game with a review score of 99% but only 5 reviews would receive a label of positive. With 60 reviews it would receive a very positive label. Figure 1 provides an example of how the review score and review label are displayed. Here, we are looking at the top of the game page for Team Fortress 2. In addition to pictures and videos, consumers are immediately shown the review label (in the highlighted red box). Hovering over the label with the mouse reveals the underlying review score. The difference between 'recent reviews' versus 'all reviews' is that only reviews submitted in the last 30 days are used for the calculation of the recent reviews score and label. I will not be working with the recent review score and label as it was only introduced on the 3rd May 2016, whereas the all reviews score and label were available throughout my data. Also, the recent reviews score and label can be quite noisy given that review activity can be quite lumpy. The last column of table 1 gives the number of games with each type of review label on 1st February 2018. As can be seen, the worst labels have very few games. On the other hand, most games achieve one of the positive labels, although overwhelmingly positive is quite difficult to get.

Steam is well known for the frequency and size of promotions available. Throughout 2017 in my data, games went on sale on average almost 15% of the days in the year they were available. Promotions come in many forms. Most notoriously are the large-scale Steam sales where the vast

[13] There are no requirements for the written component of the review, thus they can be as short as a sentence.

majority of games are discounted at rates that can be as high as 80 per cent. These typically coincide with common occasions for sales more generally, such as for the summer break, Thanksgiving and Christmas. Even outside of these periods, a sale is featured on the front page every day. These might arise from specific games, or for an entire developer or publisher's catalogue. Occasionally, there are also free weekends where games are free-to-play for a weekend and then require purchase for a consumer to keep playing.

There are several other consumer relevant features about the Steam platform that are not closely related to its storefront. For my purposes, the most important of these is that Steam tracks much of each individual's activity - specifically time spent on each game, and unlike other platforms makes much of these data accessible. With permissions, this data can also be made visible to the public which is conducive to being collected - a feature which I make extensive use of in this paper.

Steam as a distributor for developers

I now briefly turn attention to describing important features about the supply side. One of the most transformative features about Steam as a digital storefront is how it lowered barriers to entry for game developers.

2018-02-01	All	Indie	Non-Indie
Average Price ($USD)	9.72	8.07	13.33
25th, 50th, 75th pct Price ($USD)	3.99, 7.99, 11.99	2.99,5.99,9.99	4.99,9.99,19.99
2017 (My individual level data)			
Total Sales	93,548	55,102	38,446
Average sales per game	0.027	0.024	0.033

Note: The individual-level data mentioned here capture only sales, when there was no discount, for reasons I describe in the data section. I use this data as the basis for structural estimation.

Table 2: Price and Sale summary statistics

Steam eliminated the dependency on game developers for a publisher and allowed them to directly sell to consumers. The fees that developers pay to release games on Steam come in two parts. To make a game available for sale on Steam, developers need to pay a fixed cost of $100 USD. Second, for the duration of my data, Steam took a 30 percent cut of all revenues made by developers[14]. Developers are responsible for setting prices, including across all currencies, though Steam does provide tools to suggest conversions for prices in other currencies. They are also ultimately

[14] At the end of 2018, Steam introduced a 'tiered' system where Steam takes a revenue cut of 30 percent on revenues

responsible for sales, including the duration and size. Large scale sales events are on an opt-in basis. I present summary statistics about prices of all games in table 2 as at the 1st February 2018. On average games are quite cheap, although there is heterogeneity in prices across games and across time. I split games by indie versus non-indie games which is an important delineation in entertainment goods markets. Indie games come in all different kinds but are typically released by smaller developers - which can be as small as a single person and without publishers. Without the need to meet a publisher's requirements, indie games often try to do something different compared to their non-indie counterparts. Steam likely contributed significantly to the growth of indie games, due to the lowering of barriers to entry. Table 2 shows that indie games tend to be less expensive than their non-indie counterparts, having significantly smaller budgets to work with than a non-indie developer. In terms of sales, for the year of 2017 we see in aggregate that indie games are an important segment representing the majority of purchases. However, average sales are less for indie games indicating that there are more indie games than non-indie games. Also, average sales are very small. This is emblematic of the long tail of sales in entertainment good markets, and also that my sample is a very small fraction of the universe of Steam accounts. Heterogeneity in pricing across time does not mean anything sophisticated. In the appendix, I show that developers tend to set a price that is fixed for long periods, changing infrequently, and over time appear to trend downwards. The fact that it seems hard to characterise seller behaviour as optimal in a standard sense, leads me to avoid modelling a supply side for this paper.

Two policy changes - the 'Discovery update' and introduction of refunds

Having described the features of the Steam market, I now describe in detail the two policy changes of interest in this paper. First is the 'Discovery update' which launched on 22nd September 2014. This update changed the consumer's experience of the store from being the same for everyone to one where there was partial personalisation. Specifically, I previously described the front-page on a pre Discovery update basis, where everyone saw the same featured games. Following the discovery update, any Steam account with sufficient activity was shown a front-page where some of the featured games are recommendations based on the account holder's data. No one really knows the specific details of how recommendations are made, however, as the reasoning is provided, one can infer that recommendations are made on the basis of a few things such as genres, tags, play-time of games played and games played by friends. Still, many recommendations continue to be consumer invariant and are made on the basis of being popular or on sale. As the name of the update suggests, Valve implemented this change to help consumers navigate the increasing size of

up until $10 million USD, 25 percent on revenues between $10 million USD and $50 million USD and 20 percent on any revenues made from $50 million USD and above. This change was likely made to placate large game developers who had been complaining that the cut was too large and were looking for alternative options.

the store. Alden Kroll a Valve Designer says in an interview by Yin-Poole (2014):

"In the past nine months over 1300 new titles have been added to Steam and we see no signs of that volume slowing down. With so many new titles coming, the old format did not offer enough placements on the front door and throughout Steam to expose the growing breadth of offerings to customers."

About a year later, Valve released a blog post describing an assessment of the update[15], and showed that the update did seem to have driven increased exposure to more games as well as sales.

The second policy change of interest is the introduction of a refunds policy. For a long time, Steam did not have a refunds policy. In general, this meant that consumers could not request refunds even if the game did not run on the consumer's computer. In 2014, the Australian Competition and Consumer Commission launched legal action against Valve challenging this policy ACCC (2014). Following this, on 2nd June 2015, Valve introduced a global refunds policy and refund request tool. This policy entitled consumers to a refund provided that the player had not spent more than 2 hours playing the game, and that the refund request was made within 2 weeks of purchase. Noticeably, these terms are somewhat generous and allow for consumers to refund games simply because they don't like the game they bought. The policy alone may not be so extraordinary given refund policies on other platforms, but in practice it is worth noting that the refund process is very easy and costless. To initiate a refund, a consumer submits a refund request which only requires a few clicks and a note which can be as short as a few words. If the refund request falls within the conditions, the turnaround on approving the refund request is usually within the hour. Consumers have the option of electing the refund to be either credited to a bank account or to credits which can be used for future Steam purchases. The approvals are not automatic because Steam reserves the right to deny requests if they are being 'abused'. In any case, in contrast to packaging a product to ship back to the retailer and potentially spending time on support, this is still very easy.

I do not directly observe refunds, however figure 2 shows a graph and some statistics which are publicly available [16]. We see that almost 85% of support requests are for refund requests. Concurrent usage is also publicly available[17] and at the time of this draft suggests that Steam has anywhere between 20-35 million concurrent users a day. Extrapolating the 85% proportion to the chart against time which plots daily support requests it's clear that consumers take significant advantage of re-

[15] Which can be found in this reddit post https://www.reddit.com/r/Games/comments/309469/steam_platform_analysis_the_discovery_update/

[16] from https://store.steampowered.com/stats/support/

[17] from https://store.steampowered.com/charts/

Figure 2: Steam support tickets including refund requests as at 07/06/2024

fund requests which can reach over 600,000 a day and even higher during a major sale. Ideally we would want to know something about how refund requests evolved from the introduction of the policy, however this data was not collected at that time (even using the Wayback Machine), and I do not know of any other statistics.

3 Data

I begin by summarising my data sources. While I am able to directly collect some data from Steam through crawling or publicly-exposed APIs, I also collect some data through third-parties. At the individual-level, I use two distinct datasets. One, which I collect myself and one collected by O'Neill et al. (2016). O'Neill et al. (2016)'s data allow me to directly observe purchases without further assumptions. Unfortunately, this data only spans a short window of about a month in the post discovery update and pre refunds regime, thus I cannot use this data to study policy effects. To address this gap, I collected my own individual-level data. However in this data I do not directly observe quantity data and infer it by assumption. I use O'Neill et al. (2016)'s data to show that the assumption is reasonable for purchases bought when there is no discount available.

Price Data

Steam's publicly-exposed APIs only provide data at the time of the download, thus I turn to third-parties who have built historical price data by repeatedly and consistently calling from the API. I triangulate pricing data across two sources as there are some games or periods where one or the other source may not be collecting data for. The sources are IsThereAnyDeal (ITAD) [18] and SteamDB [19]. As the name suggests, ITAD's main purpose is to track prices of games across time and across stores. On the other hand, SteamDB is exclusively focused on the Steam platform and collects an extensive amount of data on many facets of the Steam platform. Neither have datasets avaliable for download, but both allow crawling [20]. In both cases, the data which I am actually able to collect are instances when prices ($ USD) change. That is, the prices and discounts if any when a change happens[21], and the associated timestamp. From these I can compile a daily level panel of prices.

[18] https://isthereanydeal.com/

[19] https://steamdb.info/

[20] For SteamDB, specific permissions are required as per https://steamdb.info/faq/#can-i-scrape-steamdb-for-academic-purposes

[21] As noted in section 2, Steam Developers set prices and discounts if any. In the data this is reflected as a 'regular' price and 'actual' price. The actual price only differs from the regular price if the developer has set a discount and I observe both from both ITAD and SteamDB.

Game Characteristics and Reviews Data

I collect most game characteristics data from Steam's publicly-exposed APIs[22]. This gives me the game characteristics listed on a game's Steam page including genres and release dates. All individual Steam reviews can be accessed on Steam. I crawl through these pages to build up a dataset of individual steam reviews for each game (and thus the review score and label). Although, I could directly collect review score data, the main advantage of collecting individual reviews is that I also collect the associated individual Steam account user's ID by doing this. This is important as to collect individual level data I need Steam account IDs to know where to collect from[23].

Individual level data and inferring quantities

With a list of steam account IDs in hand from the reviews data, I can use publicly exposed APIs to collect information about individual accounts. From the reviews data I collect 8,934,891, steam account IDs. From this set of accounts I randomly collected 'basic' account details of 39,607 US accounts. Importantly by default, Steam accounts are set to private which means that most of their data is not publicly accessible. I am then able to collect more detailed individual level data from 10,178 accounts, that is, roughly 1/4 of US Steam accounts make their data publicly available. This opt-in public rate might seem high, but this sample was already selected from users who have submitted at least one Steam review. I don't have a good sense of why users might opt-in to a public account. Speculatively, one reason could be that some third-party Steam services require a public Steam account. For example, in some of the biggest multiplayer games such as DOTA2 and CS:GO/2, there are markets for in-game items where players can trade, purchase and sell 'skins' - custom designs of in-game weapons. Services which facilitate trading may require public accounts. Whatever the case, the sub-sample of consumers who submit a review and make their accounts public are likely to be the more active Steam users and gamers. Table 3 presents summary statistics about the accounts collected. The accounts sampled are quite mature and have more games than you might expect. As a reference point, at the time of writing I own 162 games myself. Regarding playtime, first note that the data does not include free-to-play games such as DOTA2, CS:GO/2 and TF2. These are some of the most popular games, attracting the most playtimes. In this light, playtimes of accounts are also relatively large. Focusing on more active and high usage accounts limits could obscure some interesting heterogeneity in the analysis, but do not affect answering the main questions of interest in this paper.

[22] Although they are publicly exposed, there is no official documentation. I found `https://wiki.teamfortress.com/wiki/User:RJackson/StorefrontAPI#packagedetails` and `https://steamapi.xpaw.me/#ISteamApps` useful references.

[23] I only discovered O'Neill et al. (2016)'s data after doing this collection. Future drafts might use the accounts collected from their data.

	Mean	10pct	Median	90 pct
Account Age (yrs)	11.3	7.5	10.8	15.9
# Games Owned	453.4	90	287	878.3
Total Playtime (hrs)	115.3	5.4	89.6	241.6

Note: The data does not include data about free games such as DOTA2, CS:GO and TF2.

Table 3: Individual account data summary statistics

In addition to an individual's library of owned games and their playtime in each owned game, I also collect information about achievements earned in each owned game. Game achievements are set by game developers. They are unlocked by the player when they meet the conditions of the achievement, and thereafter show up on their steam account profile as having been completed. Achievements can come in many different forms. Some are hard to achieve. For a single player game, an example could be 100% completion of the game. Others are trivially easy to achieve, for example, complete the first level or tutorial of the game. Achievements serve many roles. They are relatively easy for a developer to implement, yet on the other hand, can help improve a player's engagement with the game - some players are 'completionists' and like to try and achieve everything. They also provide measurable data for a developer. This is also important for me. I observe which achievements were unlocked and the timestamp of each unlocked achievement. This 'unlocks' more possibilities with the individual level data. Specifically, I do not directly observe purchases. Since I crawl each account once, I only observe which games the account owns at the time of crawling. To make progress, I make use of the achievements data and assume that a game was purchased on the date for which the first achievement was unlocked. As noted before, games typically have achievements which are trivial to obtain and do not require much playtime. The time from opening a game for the first-time to first achievement is unobserved[24], however the 75th percentile time from first to second achievement is about 1 hour and well within a day. To this extent, the first achievement could be a reasonable proxy to the date of purchase. In the next section, I confirm this for when a game is purchased with no discount.

[24]Even O'Neill et al. (2016)' data would not allow me to observe this as the timestamp of when a game is opened for the first time is not publicly available.

Game release year	Share of games with ach.	Share of in-data owned games with ach.	Share of in-data owned games with non-zero playtime with ach.
2014	0.63	0.79	0.83
2015	0.67	0.80	0.82
2016	0.70	0.84	0.86
2017	0.72	0.85	0.86
2018	0.74	0.93	0.95

Note: ach. = achievements. The data does not include data about free games such as DOTA2, CS:GO and TF2.

Table 4: Shares of games with achievements

In the meantime, there are a few other limitations with this approach that need to be stated. First, not all games have achievements. Table 4 shows various measures of the proportion of games with achievements. In the first column we see that it is indeed the case that the greater majority of games do have achievements. Whereas the first column is unweighted, the second column weights by the number of owners, showing that when focusing on more relevant games the share is even larger. The last column shows that focusing on the number of owners with any playtime (which does not require anything about achievements for me to observe) increases the shares further still. I take table 4 as suggesting that requiring games to have achievements for inferring sales is a minor limitation. The second limitation is that even if a game does have achievements, if a player never achieved an achievement then I also cannot infer the purchase date. This could arise if the player never plays the game at all or if they play the game insufficiently. Of around 3.3 million player-game cases, I drop almost 60% or almost 2 million player-game cases. Of the almost 2 million player-game cases, around 73% are cases where the game has zero playtime with the remainder being games with insufficient playtime for an achievement. While these numbers may sound confronting at first, going back to table 3 this is not that surprising given the relatively large number of games which consumers own[25].

To validate my individual data and supplement it for structural estimation I draw upon data from. O'Neill et al. (2016). In this data, O'Neill et al. (2016) collect individual account information every day for most of November 2014 for around 178,000 accounts. Since each account's library is observed on a daily basis, actual purchases are directly observed. In this sample, 373,716 purchases are observed. Since November 2014 is entirely within the the post discovery update and pre refunds regime I cannot use it to study policy effects.

There are other potential sources of quantities data that could potentially be used. Given the

[25] There could be many different reasons giving rise to the large numbers of games never being played or being played insufficiently. I don't consider this important for the research questions at hand.

limitations introduced by my method, it is useful to briefly discuss my choice of data. The most established source that provides quantity estimates is SteamSpy. As I understand, over my sample period Steamspy collected individual account information in a similar way as I did but on a much larger scale, sampling around 800,000 accounts a day Galyonkin (2018). Ownership estimates are then computed by using the proportion of accounts owning the game in a 3-day sample and then extrapolating using a measure of the total number of accounts on Steam. I avoid using Steamspy data because for individual games the estimates can be quite noisy. The 3-day moving average in particular can induce undesirable persistence that makes sales implausibly high on dates outside of sales. There is also not much documentation about this data, so I am not familiar with the details of the methodology [26]. In recent years, a few other parties have also come to provide quantity estimates, for example VG Insights[27], Gamalytic [28] and PlayTracker [29]. These parties appear to extend the methodology of SteamSpy, by using other sources of information and machine learning algorithms. There are three reasons why I do not use these data. First, none of the sources provides precise documentation about the methods used. Second, the quantities sold estimates for individual games derived from these sources are quite noisy. To these websites' credit, this fact is not obscured. VG Insights and Gamalytic publish comparisons with actual known sales. While the error is relatively small for some games, there are many games for which the error is around or greater than 10%. Finally, individual-level data and especially playtime data will be crucial for my reduced-form and structural model as I show throughout the paper.

Validating quantity data

Given the strong assumptions used in inferring purchase dates, I now provide checks to show that the quantity data appears to be reasonable. I begin by turning to the daily sample from O'Neill et al. (2016). I augment this data and extract corresponding achievements data for purchases observed in this sample. In this sample, 373,716 purchases are observed. I collected achievements data for around 43% of these purchases. The results are plotted in figure 3. With both the true purchase date and the date for when the first achievement was unlocked in hand, I can compute the difference between the two. We observe that while there are many cases around the 0 mark, there is a long tail. Across all purchases this means that only around 17% of cases have a 0 day difference, while just about 30% have a difference of no more than 7 days. However, if we focus on only purchases made when there is no discount, accuracy improves dramatically. In particular, 36% of cases have a 0 day difference and about 80% have a difference of no more than 7 days. 7 days is a reasonable

[26] In fairness, I accept that my datasets and code are not publicly avaliable. I commit to doing so, if I get a research job and this work is pushed towards publication.

[27] https://vginsights.com/insights/article/steam-sales-estimation-methodology-and-accuracy

[28] https://gamalytic.com/blog/how-to-accurately-estimate-steam-sales

[29] https://playtracker.net/info/how_insight_works/

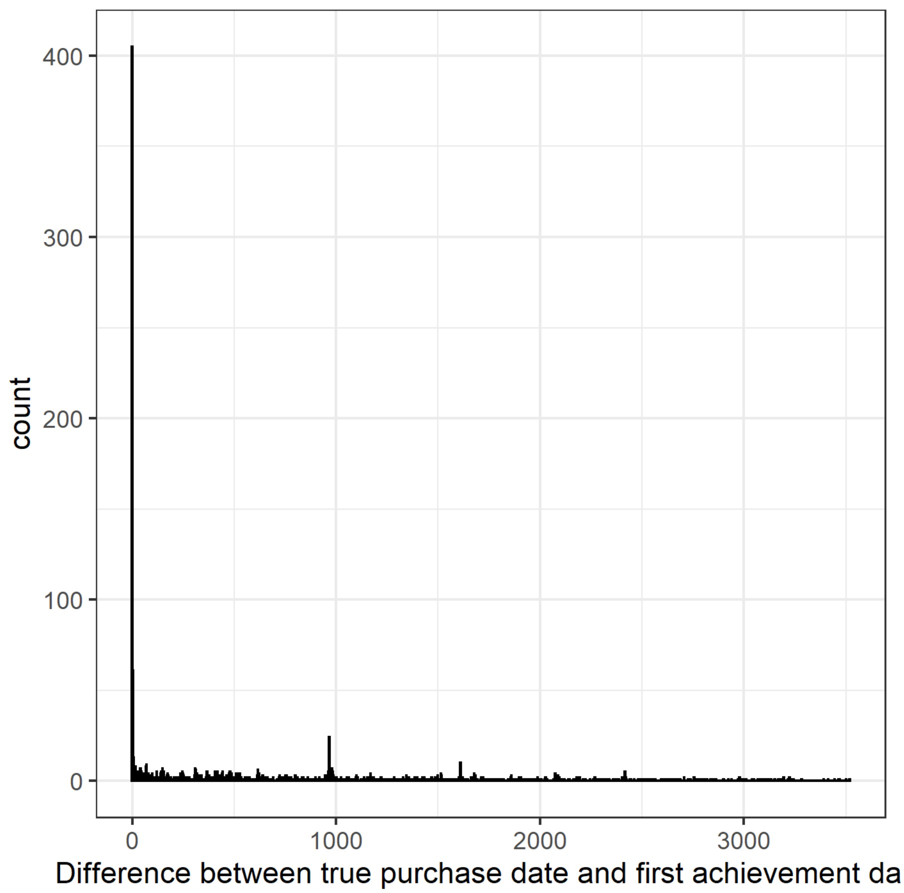

Figure 3: Difference in true purchase date and date of unlocking first achievement

benchmark for my purpose. The only observables which may vary with time significantly at a daily level are price and being on the Steam front page. As shown in the appendix, prices outside of sales periods very rarely change as shown in the appendix, thus analysis which ignores sales periods already removes this problem[30]. Regarding the front page, there are two points. First, the accuracy improves even further with almost 90% of purchases falling within 7 days and 74% within 3 days for games bought when featured on the front-page, but not on sale. Next I regress total purchases of games in my data at a daily level with an indicator for whether the game was featured on the front-page with game and day fixed effects. I get an estimate of 1.34 with std. error of 0.10. This is a very large increase in sales relative to the mean of 0.027 which makes sense given the dramatic exposure a game gets for being on the Steam front-page. Overall I regard the time-accuracy of my individual level data in game-days where there is no discount as sufficient for my analysis. In some of the data shown in section 2 and for some of the motivating evidence where timing of purchase is not relevant I use the entire individual-level data sample.

4 Motivating Evidence

Playtime varies wildly across games consumers have purchased

Hours/day	Average	10th pct	25th pct	50th pct	75th pct	90th pct	95th pct	99th pct
All Player-game av. playtime	0.014	0.00003	0.0005	0.002	0.007	0.025	0.054	0.22
Within-player av. playtime st dev.	0.052	0.002	0.011	0.026	0.058	0.12	0.18	0.43
Within-game av. playtime st dev.	0.009	0.0003	0.0009	0.002	0.006	0.017	0.033	0.12

Average playtime for each player-game pair is calculated by dividing total playtime by the difference of the date of extraction minus the first achievement date.

Table 5: Playtime: Overall, within-player and within-game

Intuitively, product usage should be informative about ex-post match values. Observing a consumer having purchased two products, but using one much more than the other can be attributed to differences in ex-post match values. Similarly, distributions of product usage should also be informative about the variance on ex-post match values. While a structural model is required to

[30]For demand estimation, variation in prices from sales may be important for correctly identifying price elasticities, one option is to useO'Neill et al. (2016)'s sample to identify price elasticities.

fully distinguish ex-post match values, I begin by presenting the playtime data to show that ex-post match values are likely to be significant. In fact, we have already seen to some extent that ex-post match values may be important - the sheer number of games that were purchased but never played. But the fact that the game was never played might suggest something other than the utility of the game is at play, for example, shocks to leisure time. In table 5 I show summary statistics about average playtime. Average playtime is calculated as total playtime divided by the difference between the extraction date (when I collected the data) and the first achievement date - essentially how long the player 'actively' owned the game for. The first row shows summary statistics about average playtime for every game-player pair in my data. In the second row, I compute the standard deviation of average playtime for all games owned by a given player. Similarly, the third row is the standard deviation of average playtime across players for a given game. The summary statistics are then the distribution of standard deviations. I present a table instead of a histogram because the tails are extreme, but important. The first row shows that even conditional on a given game being ever played, there is still extreme variation in playtime. Most games register minimal playtime, while some register a lot. The 90th percentile game has ten times the playtime of the median game, while the 99th percentile game has 100 times the playtime of the median game. The first row could simply reflect differences in types of gamers - with the right tail being composed of 'heavy' gamers. The second row shows that this is far from the case. Even the 10th percentile standard deviation is the same as the median average playtime from the first row. Although the within-player statistics remove time-invariant player heterogeneity, it is still possible that the within-player standard deviations are large from shocks to leisure time[31]. Since I don't observe playtime as a flow, it is not possible to rule this out, however, I note that the average time difference between the first and last achievement unlocked is 358 days, while the 75th percentile is 404 days. This could suggest that players will play games over long periods, or return to a game after a long period, as long as they like the game enough. Finally, we might worry that the variation in the first row is predominately driven by differences in game length. The final row allows us to focus on variation in playtime within games. Again we see that there is significant variation remaining. For example, the 25th percentile standard deviation of average playtime is about half of the median average playtime. Overall, these data suggest the presence of a long tail in consumer welfare when including ex-post match values which is likely to be even larger than one without.

Thinking of our own consumption preferences, it may not be that surprising that there is such variation in playtime. To highlight the incredible variation in playtime, I go one step further and present a series of 3 case-studies on pairs of games that are ex-ante similar, but still have large variation in playtime. I present one case study here - Europa Universalis 4 and Crusader Kings

[31] For example a player plays a particular game each month, but in some months they have more gaming time than others

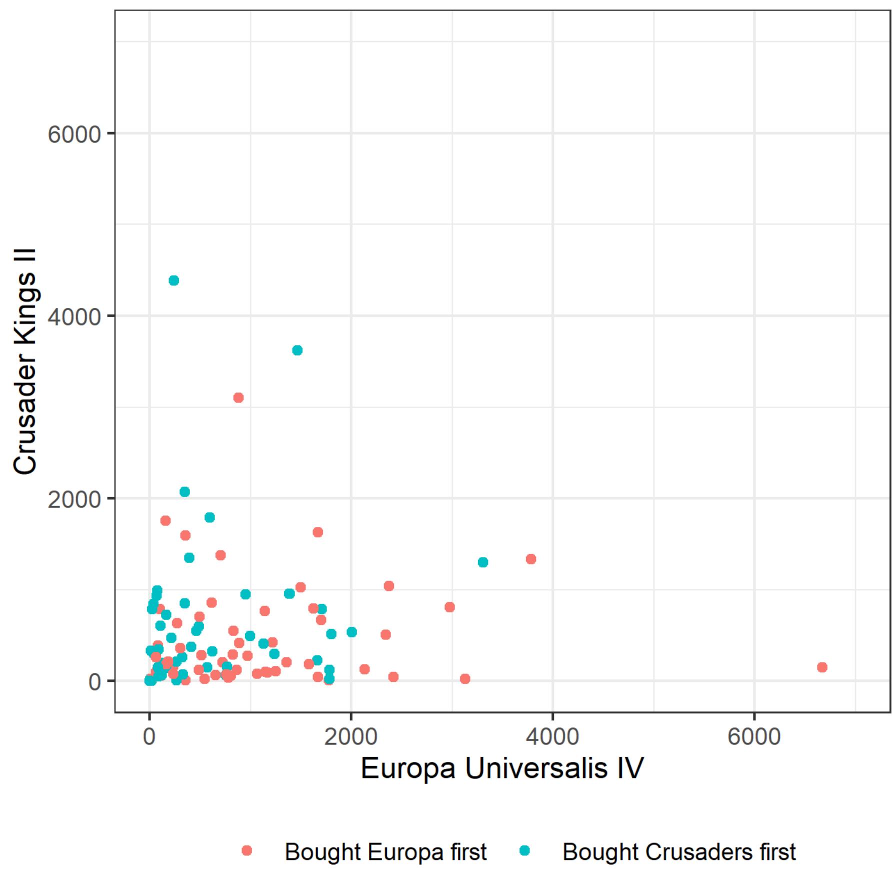

Figure 4: Total playtimes of the two games for individuals owning both

2. Snapshots of the Steam store pages for these games are included in the appendix as well as the remainder of the case studies. Ideally, I would present a comprehensive assessment. However, it is difficult to fully capture game heterogeneity with just genres and tags[32]. I chose Europa Universalis 4 and Crusader King 2 deliberately for the following reasons. The games are produced by the same developer and publisher, they are in the same genre and graphically they look similar, as do their user interfaces. Although there is some time apart from their releases, they are not so far apart that we would expect their characteristics to be dramatically different. Also one is not a sequel of the other, so it's not as if players may have additional reasons to play both or only one because one is a strict improvement on the other. The key distinction between the games according to reddit is that Europa Universalis 4 is about managing nations, while Crusader Kings 2 is about managing people and having a larger role-play aspect [33]. In short, according to the observables on the Steam store page, these games should be ex-ante similar for players, yet their ex-post match values might be quite different. In figure 4, I plot the total playtimes of people who own both games, with different colours for people who bought whichever game first. Since the games are ex-ante similar, if the role of ex-post match values is limited then we would expect most observations to lie close to the 45-degree line. In fact, the vast majority are not. If the assertion that the games are ex-ante similar is not quite correct, we might still find it informative to examine the points by which game was bought first and infer that that game is preferred. While there are many observations of people buying Crusader Kings 2 first in the upper triangle and many observations of people buying Europa Universalis 4 in the lower triangle, there are still quite a number who lie in the opposite triangle. I interpret the case studies as lending further support to the significant role played by ex-post match values.

The policy changes are associated with increases in average playtime but through changes in the games bought

Having established that ex-post match values are likely to be an important determinant in consumer welfare, I now turn to thinking about how they can also be policy-relevant. In table 6, I regress average playtime on policy regime indicators. In the first row, I include no fixed effects and in the second I use game and player fixed effects. The first row tells us that average playtime increases with each policy regime, which might be suggestive that the policy changes improved matches. Due to the inclusion of the fixed effects and the negative or null estimates, the second

[32]In other settings, one could consider using machine learning tools such as an image recognition model as in Korganbekova and Zuber (2023) or Quan and Williams (2021) to find ways of measuring ex-ante similarity of games. However, as I argued previously there are important characteristics in games that are hard to elucidate in data.

[33]See for example https://www.reddit.com/r/paradoxplaza/comments/emagld/which_is_better_ck2_or_eu4/

row tells us that the increase in average playtime in the first row is likely due to a compositional effect. That is, the increased average playtime in the first row is likely due to players buying games with better matches and not playing more of their existing games for whatever reason.

	Intercept	Between	After
No fixed effects	0.0094	0.0031	0.0076
	(0.0004)	(0.0005)	(0.0004)
Game and player fixed effects		-0.0004	-0.0031
		(0.0004)	(0.0004)

Average playtime for each player-game pair is calculated by dividing total playtime by the difference of the date of extraction minus the first achievement date. As this table studies playtime across policy regimes, I only include games bought when there is no discount. The mean average playtime in this sample is 0.015

Table 6: Playtime across policy regimes

Sales increased by more following the introduction of refunds compared with the introduction of personalised stores

Since the playtime effects from policy changes are likely due to compositional effects - from players buying new games with better matches, it is especially interesting to study whether the policy changes have any effect on sales. Since the policy effects for playtime likely arose from changes in the games that consumers bought, I did not pursue a causal effect. For sales, we can, and I begin by outlining the empirical strategy to do this. As both policy changes are immediate and global there are no obvious control groups that can be used [34].

Instead, I adopt a before-after approach, similar to that taken by Bhattacharya et al. (2024). In short, I compare daily sales outcomes before and after the policy, controlling for games, their trends in sales and whether the game is new on the day (no more than 30 days since the release date). I show in the appendix that much of a game's sales occur when new, thus it is important to account for this effect. I implement the estimator in two steps. In the first step, I estimate the following specification using only data in a policy before period:

$$y_{jt} = \lambda_j + \xi_j^1 d_{jt} + \xi_j^2 N_{jt} + \epsilon_{jt} \tag{1}$$

[34]If I had true sales data from Valve, a temporal regression discontinuity design with the cutoff being at the introduction of the policies could be a useful approach, however, it also seems reasonable to expect these policies would take some time for their effects to kick in as players may need time to become familiar with the changes. In the case of refunds especially, there were no official announcements on Valve's behalf, thus it would've taken time for the average consumer to become aware of the introduction of refunds.

Here, j is a game, t is a day, so λ_j are game fixed effects while ξ_j^1 allows those game fixed effects to have varying slopes depending on the days since release of game j on day t and ξ_j^2 allows the fixed effects to vary by whether the game is a new game (no more than 30 days since release) or not. Since this regression is estimated using data in the before policy period, I interpret it as being able to provide a reasonable prediction of the counterfactual in the post policy period were the policy not implemented. I am not able to allow for seasonality because my before period data are never long enough. For the discovery update the before-period would be the almost 10 months before the policy, while for refunds, the before-period would be the window between the policy changes of about 9 months. On the other hand, I am dropping all observations associated with discounts. Seasonality in this market is likely to be largely driven by seasonal sales, so without observations during sales, the role of seasonality is likely to be limited. I account for game heterogeneity by allowing for game fixed effects which cover time-invariant unobserved differences across games. I also allow for linear game-specific time trends, where the time is based on the days since the game's release to deal with time-variant unobserved differences. The choice of using days since release is justified by strongly decreasing sales from release as shown in the appendix. The appendix also shows that linear time trends appear most suitable when allowing for a discontinuous 'new game' effect. In future drafts, I will consider robustness about the simple linear time trends assumption[35].

With the first step in hand, the second step is to then simply estimate:

$$y_{jt} - \hat{y}_{jt} = \alpha + \varepsilon_{jt} \tag{2}$$

Here $\hat{y}_{jt}$ are the predicted sales for game j on day t in the post-policy period based on the model from equation 1. I use a two-step estimator so that trends in the before policy period are not contaminated by the post policy period. α gives the average difference in sales for all games between sales in the post policy period and its prediction using pre-policy period data.

The results of these estimators are presented in table 7. The first column shows that there is a small but insignificant increase in sales of on average 1.1 units per day following the introduction of personalised stores. In the second column, we see that following the introduction of refunds there is a sales effect of around 5 times the size of the effect of the introduction of personalised stores. Importantly, the interpretation of the second column is not that the effects of personalised stores are small compared to the effects of refunds. We might expect that a recommendation algorithm would take into account that a consumer's optimal search in the refunds regime depends on σ_j^ω (the variance in ex-post match values) and change accordingly. Thus we can only conclude that the combined effect of the introduction of refunds and personalised stores is much larger than just the introduction of personalised stores. A structural model is needed to further disentangle the two.

[35]For example, one suggestion I've received is to consider a 'dampening' effect on the linear trends.

	Before-Between	Between-After
Average Sales Effect	1.10	5.40
	(0.85)	(0.97)

Note: Before refers to the before discovery update period, between is the period between the discovery update and introduction of refunds, after is the period after the introduction of refunds (which is after the discovery update as well). Standard errors are bootstrapped using 5000 iterations.

Table 7: Before-After Estimator Results

Mostly positive non-indie games have the highest variability in average playtime

Since the theory framework suggests that the effect of refunds should vary by game through σ_j^ω, I now turn to understanding the playtime of particular groups of games. In table 8, I show how average playtime varies across review label categories. It is also important to further divide between indie and non-indie games because the information environment is likely to be different. Players can search for information about non-indie games from many other sources such as trailers and reviews. On the other hand, for most indie games the Steam store page and Steam reviews are the primary source of information. Also, in addition to the mean and percentiles, I include standard deviation as it relates most closely to σ_j^ω and is expected to play an important role in policy regime effects. A few facts stand out. First, focusing only on non-indie games, except for the 10th percentile, Mostly Positive games record higher summary statistics than every other review group. Also, mixed games, although generally lower than mostly positive and very positive still have game-player pairs with high average playtime. For example, the 75th percentile mixed average playtime is still larger than the median very positive average playtime, and only just under the 75th percentile very positive average playtime. The lack of 'monotonicity' in playtime versus review label may seem odd. However monotonicity was never guaranteed in the first place. Review labels only reflect the proportion of players giving a positive rating, but some mixed or mostly positive games might attract particularly good matches that are high enough to yield higher average playtime despite relatively fewer reviewers liking the game. Regarding the positive label in particular, the market share of positive games is very small. On the corresponding date for table 1 the share of sales from 1681 positive games is only 1.1%[36]. For indie games, the patterns are similar but less exaggerated. There notable differences are that while mostly positive indie games have

[36]The reason for this is not entirely clear, but from my experience games rarely stay in positive for long as they either get enough reviews to make it into 'Very Positive' or the rating becomes low enough that it becomes 'Mostly Positive'

roughly the same average, average playtime as very positive indie games, from the median to at least the 90th percentile, average playtime is less than for very positive indie games.

Comparing non-indie versus indie games, we see that across the board, non-indie games tend to have higher average playtime. In the tails for mostly positive and very positive, this is particularly striking. For example, the 90th percentile non-indie mostly positive game is more than 5 times that of their indie counterparts. Some caution might be warranted when thinking about playtimes in the context of indie versus non-indie games. One issue is that the average indie game might be 'shorter' than the average non-indie game. Drawing on data from a site called How long to beat (HLTB) I observe that the average non-indie game has a length of 27.5 hours while the average indie game has a length of 12.2 hours[37]. To try and deal with this issue, I attempted to construct two measures of game lengths and then used those to compute playtimes relative to game length. I did not find this fruitful as many strong assumptions and choices had to be made in constructing the measures which rendered them uninformative [38]. More importantly, if the goal of studying playtime is to infer something about match values, then in a standard utility and demand framework it seems difficult to justify how a shorter playtime could be preferred to a longer playtime because a game is 'shorter'. To do so would have to imply that the gamer received a higher utility per unit time playing the shorter game than the longer game, and also that the utility per unit time of the shorter game is no longer larger than the longer game such that they wouldn't want to replay the shorter game. To have a chance at teasing this out, one would have to have 'neuroeconomic' data. Given the conceptual and practical challenges of attempting to account for game length, I will ignore this issue for the remainder of the paper and interpret longer playtime as indicative of a better match.

[37]HTLB allows players to log completion times or hours in a game which can be used to obtain a game length measure. The website can be found at `https://howlongtobeat.com/`. I sourced this data from Baara Zaid who had already scraped the data and posted their dataset on Kaggle, see here `https://www.kaggle.com/datasets/baraazaid/how-long-to-beat-video-games`

[38]In practice, measurement of game length is very difficult. In the first place, there is no 'official' measurement of game length which is why I attempted two alternative constructions. One is based on the achievements data I have and relies on strong, heavy-handed extrapolations. The other is based on the HLTB data. As the HLTB data relies on players making individual reports, the data are likely highly selected. Also for many games, there are relatively few logs making estimates of game lengths particularly uncertain. Another key challenge in constructing a game length measure is defining what might constitute a 'game length'. For games with large replayable components such as competitive multiplayer games, this seems particularly unclear. Nowadays, even games that are single-player with a somewhat clearer 'start' and 'end' might be quite replayable as well. For example, there might be alternative 'paths' that could be taken to get to the end or alternative endings.

Indie	Review Label	10th Pct	25th Pct	Median	Average	St. Dev.	75th Pct	90th Pct	99th Pct
No	Mixed	0.000012	0.00064	0.0028	0.0139	0.0502	0.0100	0.0294	0.1853
No	Mostly Positive	0.000040	0.00071	0.0034	0.0253	0.0932	0.0151	0.0552	0.3567
No	Positive	0.000000	0.00031	0.0014	0.0053	0.0257	0.0039	0.0100	0.0662
No	Very Positive	0.000051	0.00067	0.0030	0.0179	0.0750	0.0109	0.0342	0.2655
Yes	Mixed	0.000000	0.00018	0.0007	0.0083	0.0680	0.0021	0.0070	0.1471
Yes	Mostly Positive	0.000006	0.00024	0.0010	0.0102	0.0820	0.0030	0.0092	0.1915
Yes	Positive	0.000000	0.00019	0.0008	0.0037	0.0320	0.0023	0.0062	0.0446
Yes	Very Positive	0.000023	0.00035	0.0013	0.0101	0.0690	0.0043	0.0137	0.1545

Average playtime for each player-game pair is calculated by dividing total playtime by the difference of the date of extraction minus the first achievement date.

Table 8: Playtime summary statistics across groups of games

The pattern of sale changes with the introduction of refunds is quite different compared to the introduction of personalised stores, and is associated with an increase in sales for games with the most uncertainty

Finally, with an understanding of playtime data and which games are likely to exhibit the most uncertainty in ex-post match values, I present the main motivating evidence for this paper. While the discovery update may have been associated some games receiving modest increases in sales. The introduction of refunds on the other hand are associated with large increases in sales for non-indie games with relatively high uncertainty in matches. At the same time, indie sellers tend to be adversely affected by the introduction of refunds. I largely follow the same empirical strategy as when studying overall sales effects. The changes are that in the first step I instead estimate:

$$y_{jt} = \sum_{r} \left(\beta_r^1 R_{rjt} + \beta_r^2 R_{rjt} I_j\right) + \lambda_j + \xi_j^1 d_{jt} + \xi_j^2 N_{jt} + \epsilon_{jt} \qquad (3)$$

Thus I allow the level of sales to systematically differ with R_{rjt} - indicators for whether game j has the review label r on day t and I_j - an indicator for whether the game is indie or not. I include only games which are Very Positive, Positive, Mostly Positive or Mixed. As shown in table 1, most games fall into these categories.

	Indie x Time varying review labels		Indie x Time fixed review labels			Indie x Playtime Uncertainty Quartiles	
	Before-Between	Between-After	Before-Between	Between-After		Before-Between	Between-After
Intercept	5.01	7.69	0.25	5.44	Intercept	0.058	0.744
	(2.20)	(2.15)	(0.11)	(1.58)		(0.035)	(0.228)
Indie	-5.13	-3.90	10.02	-1.49	Indie	-0.029	-0.114
	(2.20)	(2.57)	(10.32)	(2.56)		(0.061)	(0.290)
Mostly	-3.74	11.42	2.96	10.83	Var 2	0.006	8.286
Positive	(1.97)	(7.42)	(4.24)	(8.60)		(0.049)	(5.140)
Positive	-4.79	-3.48			Var 3	1.737	1.826
	(2.20)	(5.38)				(2.116)	(0.700)
Very	-1.99	0.41	1.51	3.86	Var 4	3.310	27.802
Positive	(2.69)	(1.74)	(0.95)	(2.56)		(1.940)	(6.479)
Indie x	12.63	-13.38	-13.17	-12.82	Indie x	0.387	-6.545
M. Positive	(9.11)	(8.02)	(11.23)	(9.50)	Var 2	(0.558)	(5.244)
Indie x	4.96	-0.3			Indie x	-3.658	0.317
Positive	(2.21)	(5.55)			Var 3	(2.698)	(2.694)
Indie x	2.25	-3.81	-12.44	-7.06	Indie x	4.009	-27.073
V. Positive	(2.69)	(2.35)	(10.39)	(3.28)	Var 4	(7.565)	(6.489)

Note: Before refers to the before discovery update period, between is the period between the discovery update and introduction of refunds, after is the period after the introduction of refunds (which is after the discovery update as well). The baseline (intercept) are non-indie, mixed games in the first four columns and non-indie var 1 (lowest uncertainty quartile) games in the last two columns. Likewise indie captures indie, mixed games and indie, var 1 (lowers uncertainty quartile) games. Standard errors are bootstrapped using 5000 iterations.

Table 9: Before-After Estimator Results

The second step is then modified to:

$$y_{jt} - \hat{y}_{jt} = \alpha I_j + \sum_r \left(\alpha_r^1 R_{rjt} + \alpha_r^2 R_{rjt} I_j\right) + \varepsilon_{jt} \tag{4}$$

Here $\hat{y}_{jt}$ are the predicted sales for game j on day t in the post-policy period based on the model from equation 3. I use a two-step estimator so that trends in the before policy period are not contaminated by the post policy period. Each of the coefficients in this specification give the average difference in sales for a group of games between sales in the post policy period and its prediction using pre-policy period data. The results are presented in table 9. I begin with the first two columns which reflect the main specification that I just covered. The first column of the pair captures the effects of the discovery update, while the second column of the pair captures the effects of the introduction of refunds, following the discovery update. Beginning with the effect of the discovery update, non-indie games tended to benefit modestly. In my sample, non-indie, mixed games saw around 5 more sales following the discovery update, while the effects for other non-indie games are smaller but still positive, as the coefficients are negative, but less than 5. For indie games, we see that mixed, indie games are largely unaffected with the coefficients (5.01 - 5.13) almost exactly cancelling out. Mostly positive indie games see around 9 more purchases following the introduction of refunds (although noisy), while other indie games are mostly unaffected. Although noisily estimated, I am generally inclined to believe positive estimates given the public findings from Valve cited earlier, who of course observe all data. Overall, I interpret these results as suggestive that the discovery update was modestly beneficial for sellers and maybe more so for mostly-positive indie sellers.

Next, I turn to the introduction of refunds. Broadly, non-indie games see relatively large increases in sales. Non-indie mixed games have increased sales of almost 8 in my sample following the introduction of refunds, while mostly positive non-indie games see an even larger boost (although noisy). Other non-indie games also seem to fare better from the policy change compared to the discovery update. Indie games see modest effects, including mostly positive indie games which see a small increase of around 2.

Table 9 also provides two alternative specifications. First, the fact that review scores vary over time means that the underlying groups are also changing over time. Correlation between sales and these changes could bias estimates. To address this concern, I run a specification where I use fixed review labels. I construct the fixed review label as the review label that the game spent the most time in. In the first step of estimation, I drop the review label and indie variables in constructing the pre-policy trends as now there is no more within-game variation in observables, but otherwise the steps are the same. Also except for one indie game, there are no positive games under the fixed review labels, hence in these results there are no positive group results. In this specification I find some differences with the introduction of the discovery update. The estimates are noisier and mixed indie games seem to benefit in place of mostly positive indie games. Effects associated with the introduction of refunds are similar, but noisier. The use of fixed review labels is not without loss, since there will inevitably be periods where a game that was actually mostly positive at a particular time could be very positive for example. The increase in noise could come from these timing mismatches.

Noisiness in the first specification may not be surprising given that not all non-indie mostly positive games need have significant playtime uncertainty. While I think this measure is important as it directly corresponds to what players see on the store, in the third specification I turn to a direct measure of uncertainty. In particular, I use my playtime data to construct quartiles of playtime uncertainty. I compute the standard deviation of average playtime for each game and then construct quartile buckets. Like with the second specification there is no within-game variation in observables, besides that the estimation strategy is analogous to the first and second specifications. Var 1 is the lowest uncertainty quartile bucket with var 4 being the highest. Under this specification the introduction of the discovery update seems to have similar effect for indie games with the largest uncertainty indie games potentially seeing some benefit. For the introduction of refunds there is a very large increase in sales for non-indie games in the highest uncertainty quartile. This estimate seems reasonably precise consistent with the use of a direct uncertainty measure. Overall I interpret these results as further support for the role of ex-post match values and their policy implications.

5 Structural Model Outline

The goal of a structural model is two-fold. First, I will decompose consumer surplus and utility into observable utility, ex-ante match values and ex-post match values to understand the relative importance of ex-post match values and how much 'bigger' or 'longer' the long tail of product variety could be. Second, the structural model will allow us to coherently disentangle the effects of personalised stores from the effect of refunds when refunds are introduced. We do not observe in-sample the effects of only introducing refunds, so it's hard to disentangle the effects of refunds and personalised stores without more structure. Disentangling the policy effect will make clear the extent to which ex-post match values can be policy relevant. For example, if most of the benefits in the post refunds and post personalised stores regime are attributable to consumers being able to learn their ex-post match values before purchase then these findings would suggest that platform policy and tools that make it possible and easy for consumers to discover their ex-post match values may be more important than recommendation systems.

The structural model will focus on modelling demand. Modelling supply is not required for disentangling the policy effects on demand. Also, while it would be interesting to explore counterfactuals where the supply side responds, it is not clear how best to model supply. As noted previously, sellers appear to 'set and forget' prices with most changes being a temporary discount for a promotional period and almost never increase prices permanently. This also suggests that it is possible to consider counterfactuals where the demand environment changes while keeping prices fixed. For example, it might be interesting to consider the effects of refunds without personalised stores. In other settings, it would be more appropriate to allow supply to respond.

5.1 Demand prior to the policy changes

In period t, each individual in our sample i chooses alternative j to maximise their utility which is:

$$U_{ijt} = \underbrace{\delta_{jt}}_{\text{Observable mean utility}} + \underbrace{\sum_k \alpha_k x_{ikt}}_{\substack{\text{Individual-varying preferences}\\ \text{on observable utility}}} \underbrace{\sum_k \beta_k^u z_{ikt}}_{\substack{\text{Systematic}\\ \text{Ex-ante Match}}} + \underbrace{\varepsilon_{ijt}}_{\substack{\text{Unsystematic}\\ \text{Ex-ante Match}}}$$

$$\varepsilon_{ijt} \sim \text{Independent, type 1 EV with variance } \sigma_j^{\varepsilon}$$

Consumers can also simply opt to choose the outside option which is to not purchase a game with utility normalised to 0. The first two terms are intended to capture preferences for vertical attributes, most notably prices, the overall review label, age of the game and unobserved demand shocks. Preferences are allowed to vary at the individual-level with observed demographics. I do not observe standard demographics, however variables such as account age, total playtime and total number of games owned could provide relevant variation. The systematic ex-ante match component is intended to capture horizontal preferences. This would include interactions between game genres and the individual's total time spent playing games within that genre. Observing an individual spending more time playing action games should suggest that the individual would prefer purchasing action games irrespective of the vertical attributes. Finally, the unsystematic ex-ante match value will be independent type 1 EV shocks with heteroskedasticity. That is, we do not normalise the variance and allow for it to vary by game with parameter σ_j[39]. Ex-post match values do not play a role here. Prior to the introduction of refunds, consumers do not have the opportunity to realise their ex-post match values before purchase. Thus revealed preferences from purchases provide no identifying power for ex-post match values.

Of course, consumers do not consider all alternatives. In a market with thousands of alternatives, it is important to introduce limited consideration. I follow Goeree (2008) and model ϕ_{ijt}, the probability that a consumer considers alternative j:

$$\phi_{ijt} = \frac{exp(\rho a_{ijt} + \sum_k \beta_k^c z_{ikt} + \kappa_{ijt})}{1 + exp(\underbrace{\rho a_{ijt}}_{\substack{\text{utility excluded}\\ \text{consideration shifters}}} + \underbrace{\sum_k \beta_k^c z_{ikt}}_{\substack{\text{Systematic}\\ \text{Indiv. Heterogeneity}}} + \underbrace{\kappa_{ijt}}_{\substack{\text{Unsystematic}\\ \text{Indiv. Heterogeneity}}})}$$

Just as tastes for genres captured by z_{ikt} govern ex-ante match values, they should also matter for awareness. Thus consideration depends on z_{ikt}. Conditional on observables, I do not think that consideration should vary significantly as utility might, hence κ_{ijt} are normalised to be standard normal shocks. a_{ijt} contains utility excluded consideration shifters. These are important for separately identifying consideration versus utility parameters given the results of Agarwal and Somaini

[39] Following the ideas in Bhat (1995)

(2024). Most importantly whether a game is on the Steam store front page would be included here, given its influence as documented. An example of a shifter that would vary at the individual level could include how many friends of i own game j on day t. I have not collected friend list data yet, but this can be done with Steam's APIs, including when the relationship was made.

5.2 Demand following the introduction of personalised stores

I do not observe an individual's personalised store. However, as I know the basis on which general recommendations are made, it is still possible to capture the effects arising from personalised stores in a reduced form way. I assume that the effect of the personalised stores can be captured in a linearly separable way in ϕ_{ijt}. Specifically:

$$\phi_{ijt} = \frac{exp(\rho a_{ijt} + \sum_k \beta_k^c z_{ikt} + \sum_k \beta_k^{'c} z'_{ikt} + \kappa_{ijt})}{1 + exp(\rho a_{ijt} + \sum_k \beta_k^c z_{ikt} + \underbrace{\sum_k \beta_k^{'c} z'_{ikt})}_{\text{policy effect}} + \kappa_{ijt})}$$

Since playtime of games is information the personalised store recommendation uses, z'_{ikt} includes z_{ikt}. However z'_{ikt} may include other variables such as how many friends of i own j on day t. I assume all other parameters are fixed, or in other words, preferences and consideration not related to a personalised store are the same. This also means they could potentially be identified and estimated using only data prior to the introduction of the personalised store and refunds.

5.3 Demand following the introduction of refunds

With the introduction of refunds, consumers can now learn their ex-post match values before purchasing the game. I assume that the process of purchasing a game, trying it and then deciding whether to commit to it or ask for a refund can be modelled as sequential search with the search cost c_i representing the cost of time spent in trying the game. Accordingly, utility now becomes:

$$U_{ijt} = \underbrace{\delta_{jt}}_{\text{Observable mean utility}} + \underbrace{\sum_k \alpha_k x_{ikt}}_{\substack{\text{Individual-varying preferences} \\ \text{on observable utility}}} \underbrace{\sum_k \beta_k^u z_{ikt}}_{\substack{\text{Systematic} \\ \text{Ex-ante Match}}} + \underbrace{\varepsilon_{ijt}}_{\substack{\text{Unsystematic} \\ \text{Ex-ante Match}}} + \underbrace{\omega_{ijt}}_{\text{Ex-post Match}}$$

$$\omega_{ijt} \sim \text{Independent, type 1 EV with variance } \sigma_j^\omega$$

Again we maintain the assumption that preferences are fixed, so this utility is almost identical to utility in the previous periods except for the addition of the ex-post match value. Crucially this means that the variance of the ex-ante match value is already identified from demand in the before refunds and before personalised store policy regime. Moraga-González et al. (2023) shows how a model of sequential search to learn ex-post match values can be embedded in a discrete choice demand model which is tractable for identification and estimation. This discussion largely follows their paper. I can collapse utility into $U_{ijt} = M_{ijt} + \omega_{ijt}$ to highlight that conditional on j being in i's choice set at time t, consumers know everything about M_{ijt} and only need to 'search' to learn ω_{ijt}. Conditional on j being in the choice set, the conditional distribution of i's utility is: $F_{ij}(z) = F_j(z - M_{ij})$, where F_j is distributed as Gumbel with location parameter M_{ij} and scale parameter $\theta_j = \frac{\sqrt{6}}{\pi}\sigma_j^\omega$. Independence of ω_{ij} means Weitzman (1979)'s optimal search rule applies. To apply this rule we need the reservation values - the value that makes a consumer indifferent between wanting to search j and not. Each product's reservation value r_{ij} is the r which satisfies:

$$\int_r^\infty (z - r)dF_{ij}(z) - c_i = 0$$

Defining $H_{ij}(r) = \int_r^\infty (z - r)dF_{ij}(z)$, Moraga-González et al. (2023) show that under the gumbel distribution assumptions:

$$r_{ij} = M_{ij} + H_j^{-1}(c_i),\ H_j = \int_r^\infty (z - r)dF_j(z)$$

Moraga-González et al. (2023) make use of results from Armstrong (2017) and Choi et al. (2018) to solve for market shares without needing to explicitly consider all possible search paths that each consumer could follow. In particular defining $w_{ij} = \min\{r_{ij}, U_{ij}\}$, Armstrong (2017) and Choi et al. (2018) show that the outcome of a consumer's optimal search is equivalent to choosing the product with the highest w_{ij}. Moraga-González et al. (2023) show with some further work and a clever choice of distribution of search costs that in fact the market shares take on a form very similar to the usual logit market shares.

Our motivating evidence shows that the full extent of ex-post match values are best observed through the playtime data. Accordingly, I plan on using this data to assist in identifying σ_j^ω. Here I briefly sketch out how playtime could map to the demand parameters:

$$P_{ij} = h(\tilde{U}_{ijt} + \omega_{ijt}) + \Omega_{ijt}, h > 0, h' > 0, h'' > 0$$

Conditional on having purchased the game, playtime P_{ij} is an increasing, convex function of the ex-post match value plus experience utility $\tilde{U}_{ijt}$, where experience utility includes the components of utility which relate to product experience. Most notably, this would exclude prices. Ω_{ijt} captures IID measurement error. This simple representation captures that the more a consumer likes the game, the more they will play it. And, considering the long tail of playtime observed in the data, a convex relationship seems appropriate. With a choice of h, observed playtime could be matched with simulated playtime to assist identification of σ_j^ω. Although Steam knows nothing about an individual's draw of ω_{ijt}, it does know σ_j^ω. Thus the personalised store algorithm may change to take into account that consumers can now search and realise ex-post match values. Accordingly, maintaining the assumption that the platform policy effects are linearly separable, we can again model the effect of any changes to personalised recommendations by:

$$\phi_{ijt} = \frac{exp(\rho a_{ijt} + \sum_k \beta_k^c z_{ikt} + \sum_k \beta_k^{\prime\, c} z'_{ikt} + \kappa_{ijt})}{1 + exp(\rho a_{ijt} + \sum_k \beta_k^c z_{ikt} + \sum_k \beta_k^{\prime\prime c} z'_{ikt}) + \underbrace{\sum_k \beta_k^{\prime\prime c} z''_{ikt})}_{\text{policy effect}} + \kappa_{ijt})}$$

To summarise, the goal is to estimate demand in each policy regime. The only parameters that differ with the introduction of the policy regimes are indeed the ones directly related to the policies. In the case of the personalised stores, it is the $\beta_k^{\prime\, c}$ which characterise in a reduced form way the effects of the personalised store and with the introduction of refunds, the $\beta_k^{\prime\prime c}$ which characterise the effects of the personalised store in the presence of refunds, the σ_j^ω and parameters which govern the search cost distribution. With estimates of the parameters, it would be straightforward to compare the relative contributions and easy to simulate the effect of the refunds on demand without the personalised store with the assumption that the supply side is fixed, by setting $\beta_k^{\prime\, c}$ and $\beta_k^{\prime\prime c}$ equal to zero. There is of course, considerable work still in taking this model to the data, but the aim of the section was only to sketch out the structural model and how it relates to the research questions at hand.

6 Conclusion

In this draft, I document and describe patterns in the data that suggest that ex-post match values play a large role in the welfare that players derive from the games they play. Accordingly, the long tail of benefits from product variety may be even larger than in previous findings in markets for durable products. Underlying this analysis I collect unique and novel data on the Steam platform which allows me to infer individual purchases of games with reasonable accuracy while also observing the playtime of these games. The usage of playtime data is particularly novel as I regard it as directly informative of ex-post match values. The data reveal that purchases of games are not particularly informative of their playtime. Even amongst games that are played there is incredible heterogeneity in playtime, including amongst pairs of games that are both played and appear to be ex-ante similar.

To highlight that ex-post match values can have material implications for policy, I study two platform policy changes. One which introduces personalised stores, and one which introduces refunds. I interpret the latter as allowing consumers to purchase games after realising ex-post match values. This unique setting allows me to think about comparing the role of platforms in matching consumers to products versus the role of platforms in allowing consumers to explore products themselves. I find that the effects on sales arising from the introduction of refunds are five times those of the effects arising from the introduction of personalised stores. In addition, following the introduction of refunds, consumers tend to purchase non-indie games with the highest playtime uncertainty (often mostly positive games) which is consistent with a model of refunds as sequential search. The effects of personalised stores alone seem limited. On the other hand allowing for refunds and the realisation of ex-post match values before purchase are more important. It is also possible that the effects of a personalised store and the introduction of refunds are complimentary. However, these explanations can only be distinguished with a structural model of demand. I provide a sketch of a model. Although I do not directly observe the personalised stores themselves, I will be able to characterise their effects as the causes of personalised recommendations are known, and the unique data I have collected can be mapped to those causes. Thus drawing on the consideration set literature I can capture the effects of personalised recommendations in a reduced form way in how they shift choice set formation probabilities. I intend to capture the effect of refunds as allowing consumers to search for games and discuss how Moraga-González et al. (2023)'s work can be applied to my setting.

At first, it might seem that the illustration of policy relevance for ex-post match values could be too niche. To conclude, I briefly discuss how I think this framework can be used to explain other phenomena and think about other policy questions that are common to digital entertainment markets. First, especially in the past few years, 'remakes' have become especially popular in movies and games. One interpretation through the lens of this framework is that if the ex-post match value of a remake is correlated with the ex-post match value of the original, then the remake could be a way of expropriating the ex-post match value that the original couldn't. The rise of free-to-play and gaming-as-a-service models can also be thought of in a similar way. Instead of requiring the consumer to pay upfront before their ex-post match values have been realised, they pay in-game costs throughout their play. This explanation squares well with so-called 'whales'[40], that most of the revenue from free-to-play games comes from a small proportion of the player-base who are willing to spend a lot - those with large ex-post match values. Finally, whether this all matters for innovation and the development of new works also seems interesting. Do different models of pricing and seller interactions result in differing incentives for sellers for investment and development based on their ability to expropriate match value? Subscription models which have become ubiquitous throughout digital entertainment markets could be of particular concern if they further remove the ability for sellers to expropriate surplus that consumers receive compared to standard price-setting.

[40] See for example https://www.gamemarketinggenie.com/blog/market-to-whales-dolphins-minnows

References

ACCC, "ACCC institutes proceedings against Valve for making alleged misleading consumer guarantee representations," *ACCC Media Release*, 2014.

Agarwal, Nikhil and Paulo Somaini, "Demand Analysis Under Latent Choice Constraints," *Working paper*, 2024.

Anderson, C., "The Long Tail," *Wired Magazine*, 2004, *12*, 170–177.

Armstrong, Mark, "Ordered Consumer Search," *Journal of the European Economic Association*, 2017.

Arora, Krishan, "The Gaming Industry: A Behemoth with Unprecedented Global Reach," *Forbes*, 2022.

Batchelor, James, "GamesIndustry.biz presents... The Year In Numbers 2022," *GamesIndustry.biz*, 2022.

Bhat, Chandra R., "A Heteroscedastic Extreme Value Model of Intercity Travel Mode Choice," *Transportation Reearch Part B*, 1995.

Bhattacharya, Vivek, Gastón Illanes, and David Stillerman, "Merger Effects and Antitrust Enforcement: Evidence from US Consumer Packaged Goods," *Working paper*, 2024.

Chiang, Oliver, "The Master of Online Mayhem," *Forbes*, 2011.

Choi, Michael, Anovia Yifan Dai, and Kim Kyungmin, "Consumer Search and Price Competition," *Econometrica*, 2018.

Crawford, Gregory S., Robin S. Lee, Michael D. Whinston, and Ali Yurukoglu, "The Welfare Effects of Vertical Integration in Multichannel Television Markets," *Econometrica*, 2018.

Galyonkin, Sergey, "Steam in 2017," *GDC 2018*, 2018.

Goeree, Michelle S., "Limited Information and Advertising in the U.S. Personal Computer Industry," *Econometrica*, 2008.

Korganbekova, Malika and Cole Zuber, "Balancing User Privacy and Personalization," *working paper*, 2023.

Lee, Kwok Hao and Lee Musolff, "Entry into Two-sided Markets Shaped by Platform-Guided Search," *Working paper*, 2023.

Moraga-González, José Luis, Zsolt Sándor, and Matthijs R. Wildenbeest, "Consumer Search and Prices in the Automobile Market," *Review of Economic Studies*, 2023.

Neiman, Brent and Joseph Vavra, "The Rise of Niche Consumption," *American Economic Journal: Macroeconomics*, 2023.

O'Neill, Mark, Elham Vaziripour, Justin Wu, and Daniel Zappala, "Condensing Steam: Distilling the Diversity of Gamer Behavior," *USA:ACM*, 2016.

Quan, Thomas W. and Kevin R. Williams, "Extracting Characteristics from Product Images and its Application to Demand Estimation," *Working paper*, 2021.

Weitzman, Martin L., "Optimal Search for the Best Alternative," *Econometrica*, 1979, *47* (3).

Wijman, Tom, "Newzoo's year in review: the 2023 global games market in numbers," *Newzoo*, 2013.

Yin-Poole, Wesley, "Steam Discovery update personalises video game shopping," *Eurogamer*, 2014.

Yurukoglu, Ali and Gregory S. Crawford, "The Welfare Effects of Bundling in Multichannel Television Markets," *American Economic Review*, 2012.

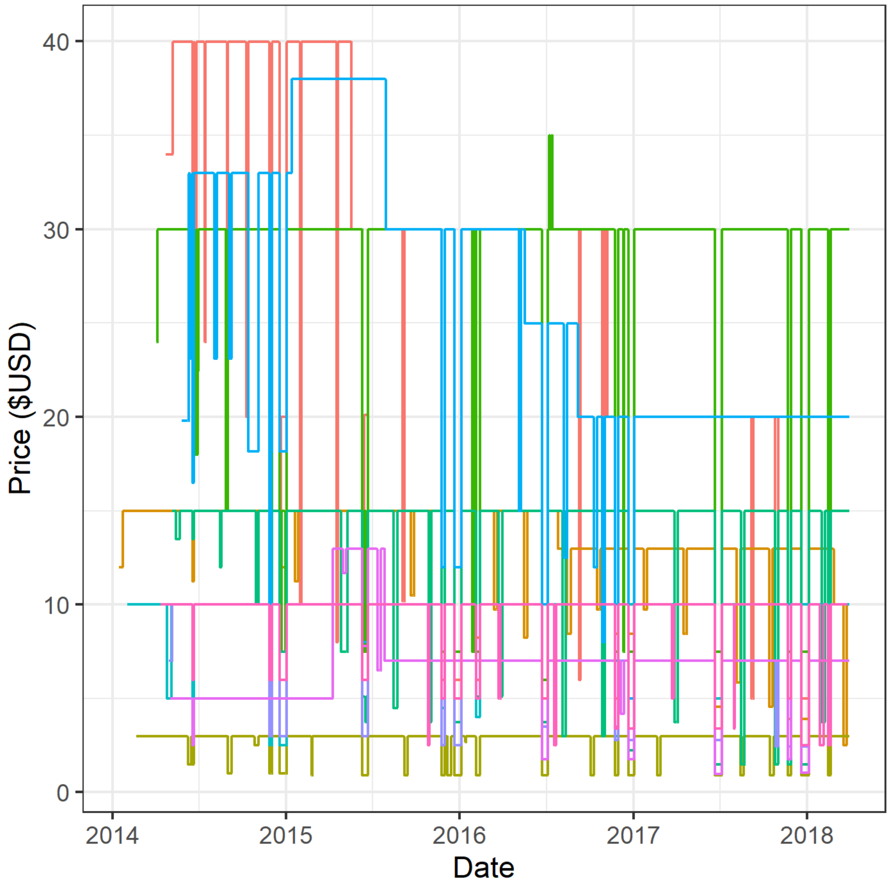

Figure 5: Prices of 10 randomly chosen games

Appendix

A Addditional Descriptives

A.1 Pricing across time

In this section I show that developers set prices in a relatively unsophisticated fashion. They set a ’regular’ price for a relatively long span of time allowing for short sales within. Over a long period of time they may occasionally adjust their ’regular’ price. When they do, it will tend to be a decrease. Figure 5 shows the prices of 10 randomly chosen games which released within the first half of 2014. The ’step’ shape of the pricing curves is apparent, as is how prices tend to decrease in the long run. Plotting individual price curves is necessary to see the ’steps’. Figure 6 demonstrates that the trend of decreasing prices over time holds more generally. I plot average and quartile prices by days since release of the game. Except for the first quartile (which is already quite low),

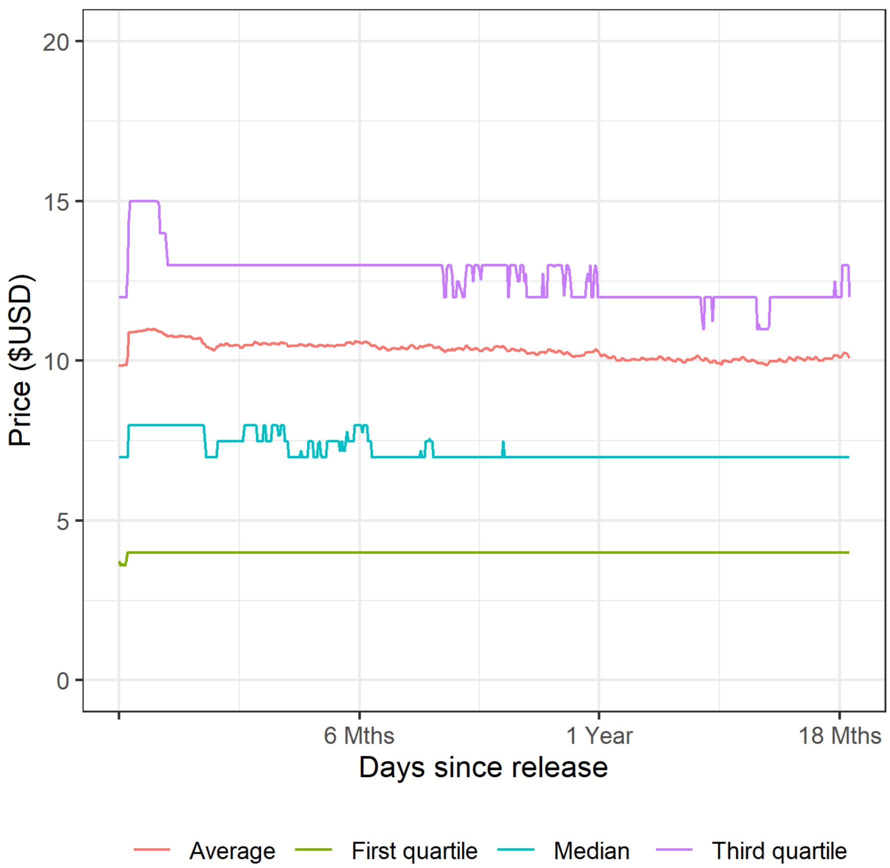

Figure 6: Average and quartile prices since release by day

prices decline given enough time. That the average and third quartile show a small uptick is likely because the panel is unbalanced. The early increase represents the strategy whereby sellers often offer a small discount on release to drive early sales.

A.2 Sales since release

Figure 7 shows a binned scatterplot of sales for games since their release date, each point is a bin covering 30 days. The figure shows allowing for a large discontinuity in the first month, game sales decrease over time in a linear fashion.

A.3 Pair-wise individual playtime case studies

Figure 8 and 9 present the Steam store pages for Europa Universalis 4 and Crusader Kings 2.

Next I present 2 more case studies to show that the case study I presented is not particularly unique.

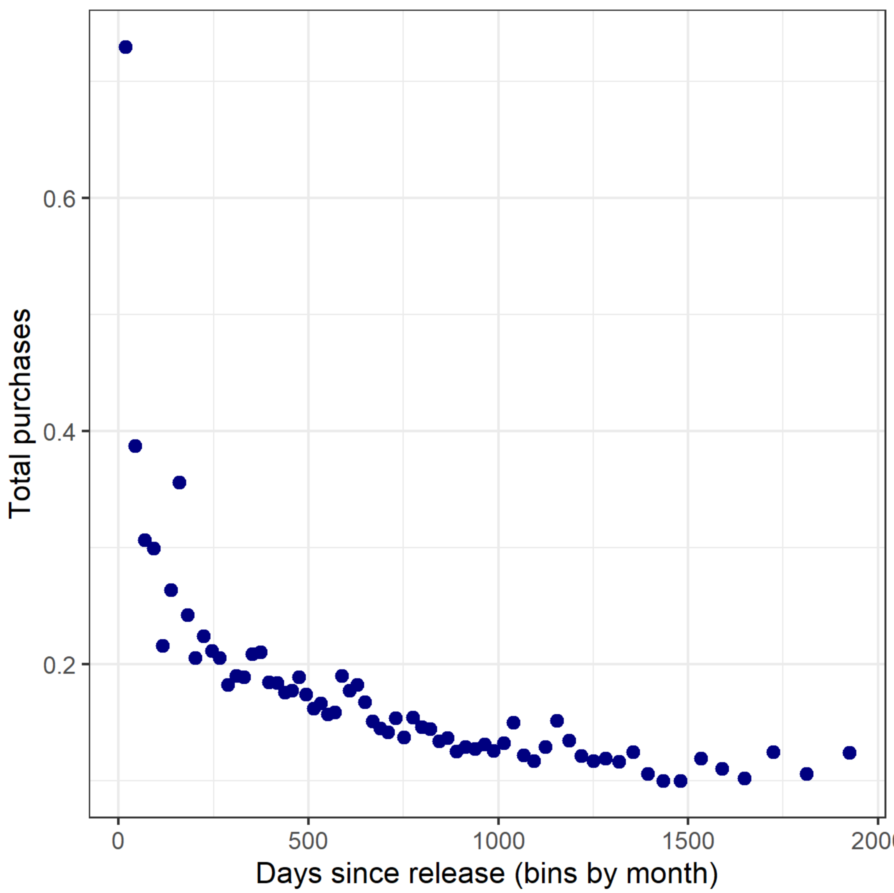

Figure 7: Binscatter of sales since release

Figure 8: Europa Universalis 4 on the Steamstore

Figure 9: Crusader Kings 2 on the Steamstore (went free-to-play in 2019)

A.3.1 Golf it! and Golf with your Friends

This case-study looks at two casual mini-golf games whose Steam store pages are displayed in figures 10 and 11. Again, I argue that these games are ex-ante similar. They have similar aesthetics, are both rated very positive and were released roughly 1 year apart[41]. Figure 12 plots individual playtime in each game, for consumers who own both games. The same analysis as for the Europa Universalis 4 and Crusader Kings 2 can be applied. Many observations are far from the 45 degree line. And while Golf With Your Friends seems to be a particularly good match for more people than Golf It! There are also a number of people who despite buying Golf With Your Friends first, still play more of Golf It! A reasonable question to ask is what could be different about the games that could lead to such playtime variation? I picked this case study as I have personally played both games. I purchased Golf With Your Friends first, and purchased Golf It! because my friends and I wanted to new mini-golf content to play with. We found that the physics engines of the games were noticeably different and strongly preferred Golf With Your Friends.

A.3.2 Dishonored and Styx: Master of Shadows

In the last case-study I look at two stealth-based action games whose Steam store pages are displayed in figures 13 and 14. Compared to the other two case-studies there are some more differences such as Dishonored having better reviews, but the games have similar aesthetics and do come up in the same threads on reddit. Figure 15 plots individual playtime in each game, for consumers who own both games. Given the better reviews, observations tend to skew towards more playtime in Dishonored, but there is much variation in that spread. And despite Styx: Master of Shadows being released 2 years later than Dishonored, there are quite a few consumers who bought Styx first, but still played Dishonored considerably more.

[41] The release dates in figures 10 and 11 are the 'full' release dates. The Golf It! and Golf With Your Friends were initially launched in 'early access' on 17th February 2017 and 29th January 2016 respectively. Early access games can be thought of as 'working papers', they are fully playable, but have more content to be added.

Figure 10: Golf it! on the Steamstore

Figure 11: Golf with your friends on the Steamstore

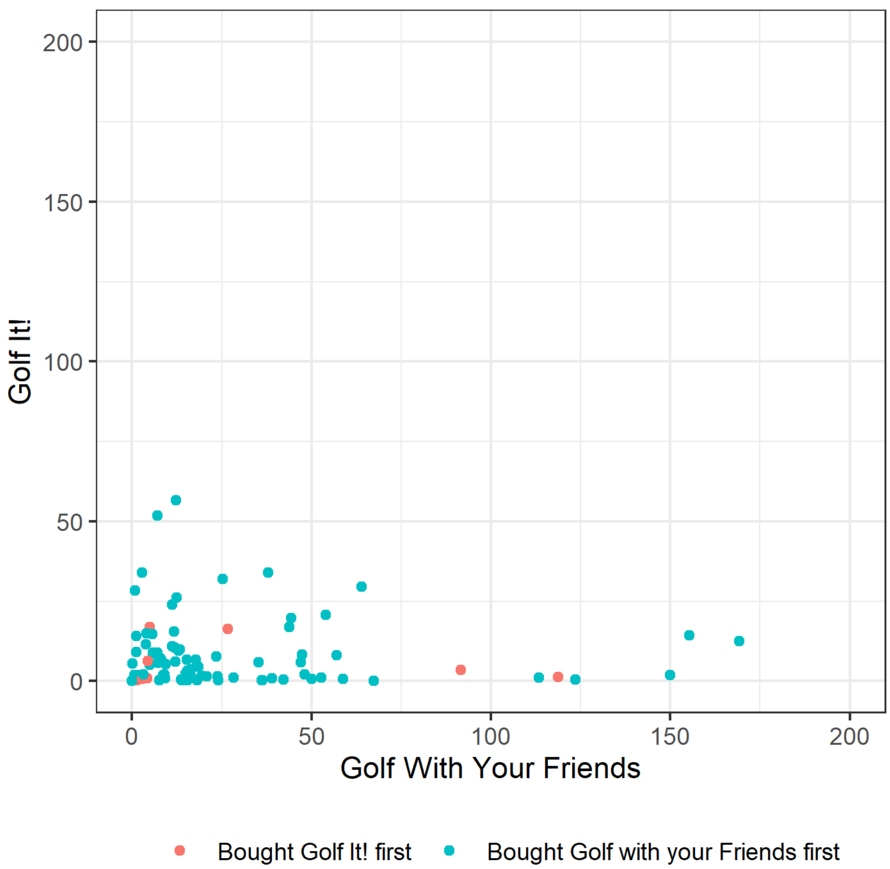

Figure 12: Total playtimes of the two games for individuals owning both

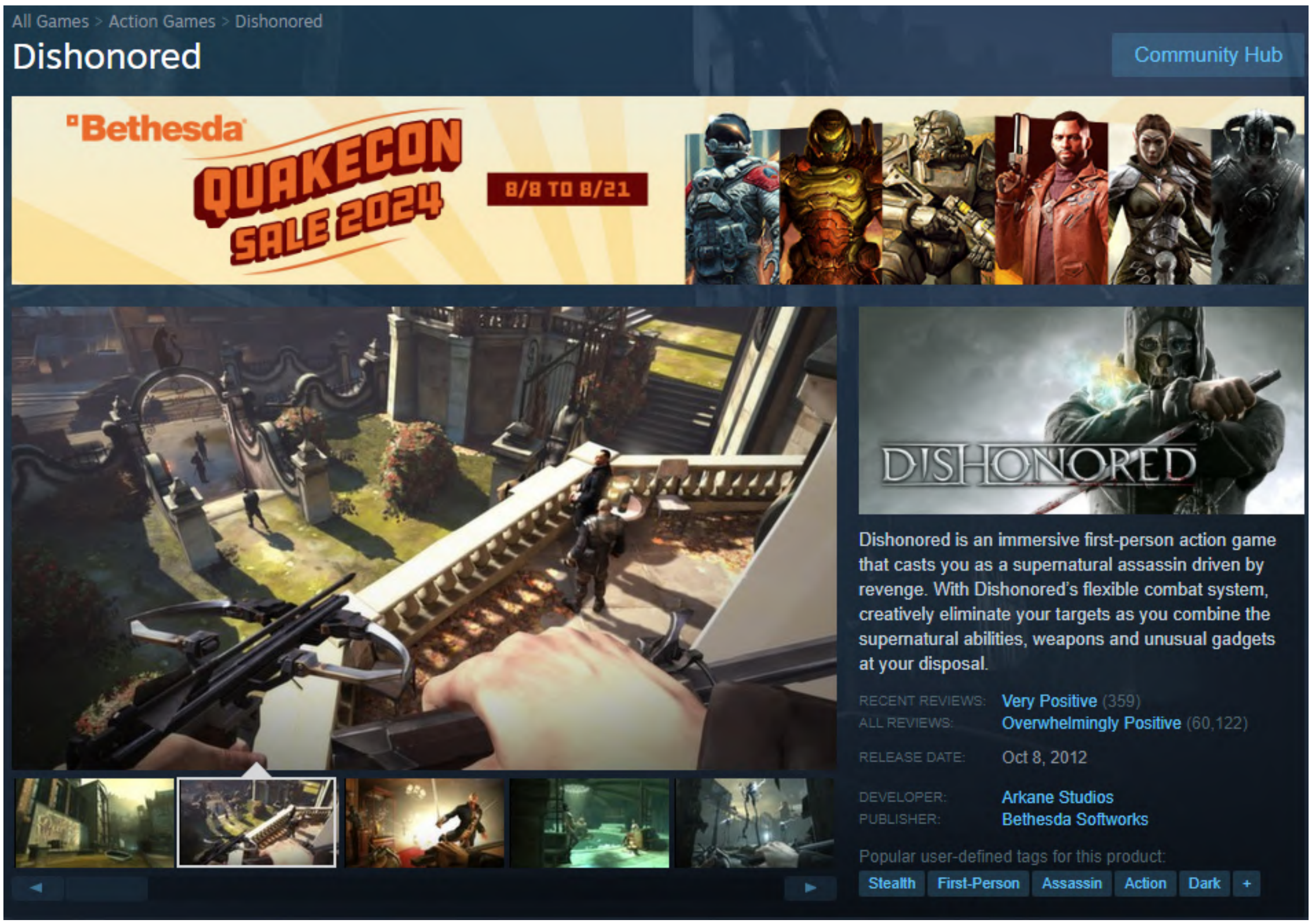

Figure 13: Dishonored on the Steamstore

Figure 14: Styx: Master of Shadows on the Steamstore

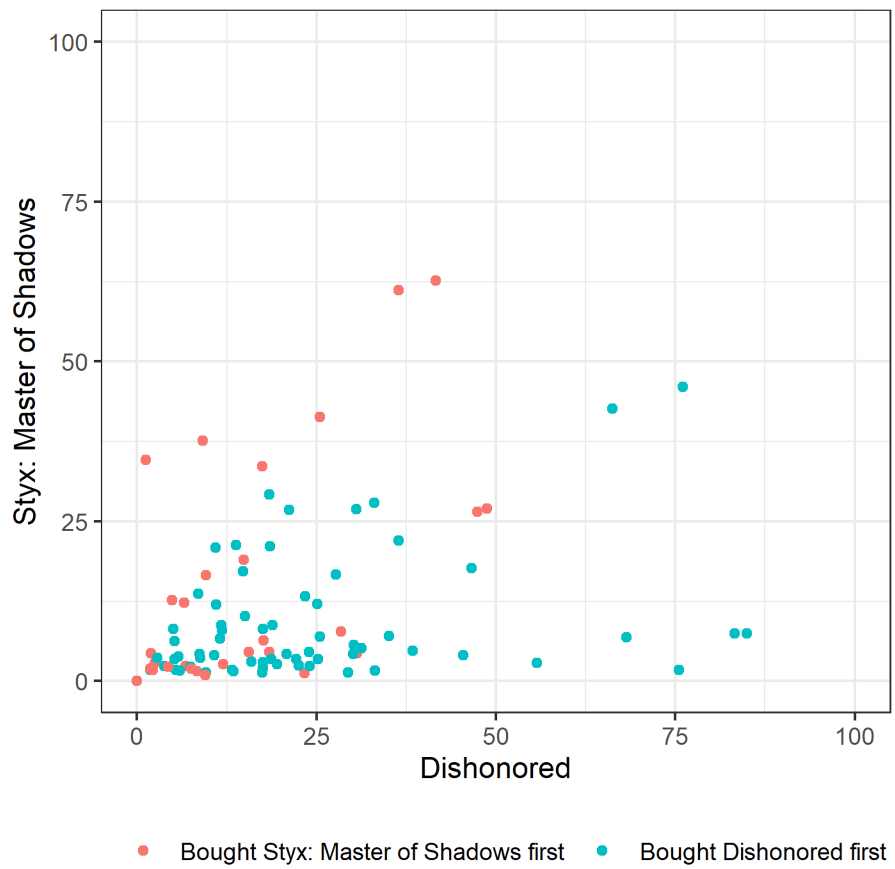

Figure 15: Total playtimes of the two games for individuals owning both