What Makes Players Pay?
An Empirical Investigation of In-Game Lotteries*

Tomomichi Amano[†]

Andrey Simonov[‡]

June 18, 2024

In 2020, gamers spent more than \$15 billion on loot boxes, lotteries of virtual items in video games. Paid loot boxes are contentious. Game producers argue that loot boxes complement the gameplay and expenditures on loot boxes reflect players' enjoyment of the game. Consumer protection groups argue that the mechanics of paid loot boxes closely resemble gambling and can attract spenders who get the direct thrill from paying for uncertain rewards. We use a unique dataset from a prototypical mobile game to estimate and compare the tastes of regular and high-spending players. While regular players enjoy loot boxes primarily for their gameplay complementarity, high-spenders get the vast majority of loot box value from the direct utility. Thus, the estimates confirm that high spenders have fundamentally different tastes for loot boxes. We use the estimates to simulate the outcomes under counterfactual game and loot box designs and evaluate policy actions proposed by consumer protection groups and regulators. Our estimates favor spending caps over a blanket ban on loot boxes.

JEL Codes: D12, D18, D61, D91, L82, L83, M31, M38.

Keywords: Product design, lotteries, gambling, dynamic demand models, video games.

I became a gambling addict over a game where there is no return, no reward, for spending my money. I flushed $16,000 down the toilet over [loot boxes in] a game.

– Anonymous Player [2017] of *Final Fantasy Brave Exvius*

In the opinion of the Administrative Jurisdiction Division, obtaining and opening the [randomized] packs is not an isolated game ... The vast majority of packs are obtained by and used for game participation ... Because the packs are not a stand-alone game, they are not a game of chance and do not require a license.

– Dutch Council of State [2022] on loot boxes in *FIFA22*

1   Introduction

Around three billion people around the world – and two-thirds of Americans – played video games in 2021 [Newzoo, 2021a, ESA, 2022]. Video games constitute the largest category in the entertainment industry, with the FTC estimating that its global revenues exceed movie box office revenues by a factor of five in 2022 [Gameranx, 2022]. Monetization of this $170 billion market heavily relies on loot boxes, an in-game lottery of virtual items, which generated more than $15 billion of revenue for gaming companies in 2020 [Juniper Research, 2021]. The importance of loot boxes for the video game industry is projected to only further increase.

Loot boxes are contentious. On the one hand, the industry argues that loot boxes enhance gameplay, improve players' performance, and in other ways are complementary to gameplay [e.g. ESA, 2019]. Under this view, high spenders are those who are deeply engrossed in the game. As such, loot boxes are similar to other uncertain rewards frequently used in product design to enhance repeated engagement [e.g. Shen et al., 2019]. On the other hand, consumer protection groups argue that loot boxes are similar to unregulated gambling – players pay to get these uncertain rewards. This can elevate other drivers of engagement with loot boxes, including a direct thrill from revealing the uncertainty of loot box outcomes [e.g. Ely et al., 2015] and behavioral biases like self-control problems [e.g. DellaVigna and Malmendier, 2006]. The latter could lead to "overspending" with negative financial and health consequences and is likely since high spenders have gambling problems [e.g. Zendle et al., 2019, Close et al., 2021] and could be minors who are most susceptible to developing problematic habits [e.g. Kristiansen and Severin, 2020]. This view is suggested by the high inequality in in-game spending, with 50%-70% of revenues coming from 1-2% of players [Udonis, 2024]. The controversy around loot boxes is evident in how varied the regulatory response has been across jurisdictions: while some countries ban or regulate them (e.g. Belgium, China), others have decided not to (e.g. New Zealand, Poland) or to continue investigating them (e.g. US, UK).

We use a unique dataset covering all in-game actions and transactions in a prototypical mobile game to estimate and compare the tastes of regular and high-spending players. In the utility model, we separate out tastes driven by the complementarity between loot boxes and gameplay – which aligns with the arguments of gaming companies on the value of loot boxes – and the direct thrill consumers get from opening loot boxes. The comparison of these taste estimates between players allows us to evaluate whether high-spenders spend a lot due to the higher game value or because of their higher direct tastes for loot boxes, reflecting loot boxes

1

as a standalone lottery for these players.

Our empirical context is a prototypical free-to-play mobile game in the most popular category, puzzle games [Newzoo, 2021b]. Users play to complete sequential game "stages" of increasing difficulty. To make progress, they increasingly need more potent virtual characters. Such rarer characters can be collected by opening loot boxes using in-game or real currency. Users acquire the in-game currency by playing the game. Users accumulate inventories of these virtual characters and select the best ones to play game stages. We observe the universe of data from this game, including full records of 2.5 million users' decisions to play, open loot boxes, and spend money, as well as their character inventory, in-game currency stock, and realizations of loot boxes. Loot boxes are used heavily: players open around 20 million paid loot boxes, and 96% of the company's in-game revenues come from loot box openings. As in other games, the expenditures are highly concentrated across users, with 90% of revenue coming from 1.5% of players. Following the industry's convention, we refer to these users as "whales" and label the rest as "non-whales" [e.g. Udonis, 2024].

This rich transaction and activity data provides a unique context to estimate players' tastes for the game, loot boxes, and the complementarity between the two. Since this is a single-player game, most of the loot box-game complementarity comes from the expected gameplay value of the items consumers might receive from loot boxes: items help make progress in the game. We refer to this as the functional value of loot boxes. We leverage the timing of players' loot box openings to identify their functional value; users should be more likely to open loot boxes when loot box items have a higher incremental effect on game win probabilities. Since, by definition, loot box outcomes are random, comparable players end up in different inventory states throughout the game. We use this exogenous variation in states to identify how much functional value players get from loot boxes by comparing changes in play and loot box opening probabilities across states with varying returns to opening loot boxes.

Leveraging the randomness of game and loot box outcomes, we compare the tastes of whales and non-whales in a series of reduced-form regressions. On the one hand, both whales and non-whales have a distaste for playing the game after a loss; players are 40-50% more likely to open a loot box rather than play one more time after losing the game stage. On the other hand, players show drastically different tastes for loot boxes. Non-whales sharply increase their stage play propensity when they receive rare items from loot boxes that will help them succeed in the game, suggesting the functional value of loot boxes for these players. In contrast, the response of whales is small and not distinguishable from zero once they have a relatively strong inventory, suggesting that the functional value of loot boxes does not play such a prominent role for whales.

To measure the tastes of whales and non-whales, we build and estimate an empirical model of gameplay and loot box choices. A consumer chooses between playing the game, opening loot boxes, or leaving the game forever in a series of discrete choices given her current inventory, game stage, and currency stock, among other factors. The consumer gets utility from playing the stage regardless of the outcome and additional utility from "winning," i.e. advancing a stage. A consumer is more likely to win if she holds a better inventory of items. Thus, in

expectation opening a loot box improves future utility flow by adding a good character to her inventory, capturing the functional value of loot boxes. At the same time, the consumer can get direct thrill utility from opening loot boxes; we capture this direct utility as a fixed effect of loot boxes plus a first-order Markov state dependence parameter, to account for potential inertia in consumer choices. To open loot boxes, a consumer needs to spend in-game currency, which can be acquired by playing the game more or paying with real money. Thus, the variation in players' currency stock introduces a dynamic trade-off between opening loot boxes and playing the game, and pins down the price responsiveness of players.

We estimate the dynamic discrete choice model using a two-step procedure [Hotz and Miller, 1993]. While our state space is large – 4.9 million states – the estimation is simplified because of the existence of a terminal action, leaving the game. Following Arcidiacono and Miller [2011], we express the expected value functions as a function of the conditional choice probabilities estimated in the first stage, and estimate utility parameters in the second. We allow for different preferences for whales and non-whales by estimating the model separately. We check the robustness of our estimates by allowing for more heterogeneity in users' tastes, clustering whales and non-whales based on users' play propensities [Bonhomme et al., 2022].

The estimates confirm the fundamental difference in tastes for loot boxes of whales and non-whales. Whales have a higher taste for opening loot boxes regardless of their outcomes, and the parameter governing state dependence in loot box openings is around twice as high for whales than non-whales. Under the current game design whales get around 90 times more utility from loot boxes than non-whales, while they get only 2.7 times more utility from playing the game itself. Quantifying the relative importance of the functional and direct thrill values of loot boxes, we show that non-whales primarily get the functional value of loot boxes: the functional mechanism accounts for almost 90% of loot boxes' value for non-whales. In contrast, whales get 97% of loot box value from the direct mechanism.

To understand how possible changes in the game design could affect the utility players get from the game and loot boxes, we simulate the outcomes of various policy actions proposed by consumer protection groups and regulators. A blanket ban on loot boxes [e.g. like the one discussed by Forbruker Rådet, 2022] without other changes to game design will reduce the utility non-whales get from playing the game by 25.4%, driven by the complementarity between the gameplay and loot boxes for these players. In contrast, whales maintain nearly the same level of utility from the gameplay and only lose utility from opening loot boxes. Simulations show that spending caps recover the vast majority of the functional value of loot boxes while preventing the firm from profiting from the overspenders. This part of our result can be viewed as supportive of proposals of Close and Lloyd [2021] and Leahy [2022], who advocate for actions like spending caps, pre-committed limits, and forced breaks from opening loot boxes.

We further use our taste estimates to evaluate alternative game designs. For this, we simulate gameplay and loot box opening decisions under different levels of game difficulty and loot box outcome variability. We show that by making the game harder, i.e. decreasing win probabilities across stages, the gaming company can extract more revenue from whales by increasing the complementarity between loot box outcomes and in-game usage. Yet in doing

so it substantially loses the engagement of non-whales. Our estimates show that the current design of the game balances these two forces well, highlighting the value of "free" non-whale players and the importance for firms to balance out the revenue and growth objectives [e.g. Gupta et al., 2009, Lee et al., 2017]. In contrast, changing the variability of loot box outcomes has a substantially lower impact on revenues and players' behavior, indicating a weaker role of the variance of the loot box outcomes in game design.

Our paper contributes to several streams of literature. A growing literature examines the drivers of users' engagement with video games and their implications, such as the role of user communities [Albuquerque and Nevskaya, 2022], need for challenge and skills [Huang et al., 2019] as well as risk-seeking and task completion [Zhao et al., 2022], learning of game features and sensitivity to promotions [Runge et al., 2022, Sunada, 2018], and satiation and its effects on re-sales [Ishihara and Ching, 2019]. Couched within the literature of broader media and entertainment markets [e.g. Crawford and Yurukoglu, 2012, Fan, 2013, Sweeting, 2013, Jeziorski, 2014, Martin and Yurukoglu, 2017, Cagé, 2020, Liu et al., 2020, Simonov et al., 2022], our paper demonstrates the key monetization mechanism, in-game purchases, in and of itself interacts with and affects the way in which players engage with the core game product.

Our estimates of players' tastes for loot boxes suggest that these in-game purchases may be associated with purchase and consumption decisions under randomness, and thus related to the empirical literature on gambling [e.g. Jullien and Salanié, 2000, Narayanan and Manchanda, 2012, Park and Manchanda, 2015, Taylor and Bodapati, 2017, Park and Pancras, 2022] and self-control problems in media consumption [e.g. Acland and Chow, 2018, Hoong, 2021, Allcott et al., 2022, Aridor, 2022].

Methodologically, we extend the domain of economic single-agent dynamic models with a finite dependence property [e.g. Arcidiacono and Miller, 2011, Scott, 2014, Kalouptsidi et al., 2021] to the realm of in-game decision making.

Finally, we contribute to the growing body of work on loot boxes in social sciences and engineering that inform public policy decisions on how to understand and regulate these purchase mechanisms [e.g. King and Delfabbro, 2018, Griffiths, 2018, Drummond and Sauer, 2018, Zendle et al., 2020, Chen et al., 2020, Xiao, 2021]. While most work relies on interviews, surveys, and correlational analysis to understand the value of loot boxes for players [e.g. Zendle et al., 2019, Close and Lloyd, 2021, Spicer et al., 2022], we are the first to provide a revealed preference measure of players' tastes. We also bring in unique field data and evaluate alternative product designs. We thus contribute to an emerging literature that studies the implications of product (self-)regulation in video games and their implications for product usage [Nevskaya and Albuquerque, 2019, Jo et al., 2020].

In Section 2, we describe the institutional context surrounding loot boxes. Section 3 describes the focal video game and our data. In Section 4, we build a structural model of gameplay and loot box openings. Section 5 presents the reduced-form evidence. Section 6 discusses the estimation procedure, and Section 7 presents the resulting estimates. Leveraging counterfactual simulations, in Section 8, we evaluate the implications of policy, game and loot box design for the firm's revenue and game engagement. We conclude in Section 9.

4

2 Institutional Context

2.1 Video Game Companies' View

The video game industry has rapidly expanded in recent years. In 2020, the industry was worth an estimated \$159.3 billion – a 126% increase compared to the \$70.6 billion value in 2012 [Statista, 2021], outgrowing the movie and music industries combined [Investopedia, 2021]. With three billion video game players choosing from more than one million games to play, competition is fierce [remarkablecoder.com, 2019]. This competition, along with advances in technology, has led to an exponential increase in the costs of producing high-quality video games.[1] However, due to competitive pressure, the list prices of major video games have remained at the same level for the last 15 years [extremetech.com, 2020], leading video game companies to search for other monetization methods to cover costs and to fund future products.

One type of such new monetization methods are in-game purchases, commonly referred to as microtransactions. Microtransactions allow users to purchase virtual items within the game, enhancing the gaming experience, and are omnipresent in both free-to-play (or "freemium") games and major titles. Driven by microtransactions, in 2018 the market for free-to-play games grew to \$88 billion worldwide [techcrunch.com, 2019].

One of the most common types of microtransactions is a "loot box." Players purchase a "black box" from which they obtain a randomized selection of virtual items to be used in subsequent gameplay. Apart from subsidizing the list price of the game in a form of "razor and blades" pricing, loot boxes allow companies to exercise price discrimination in other ways, such as bundling (bundling different items together in a loot box) and volume-based pricing (providing a discount for opening multiple loot boxes at the same time).

Loot boxes have potential benefits for players. They allow players to get in-game items that make the game more enjoyable, either by enhancing in-game skills, or by making the game more visually appealing and personalized. The mechanic of uncertain outcomes is similar to other uncertain rewards frequently used in product design to enhance repeated engagement [e.g. Shen et al., 2019]. The uncertainty over the quality of the items players receive in loot boxes can have a functional value since it might be optimal for players to pay for a chance to obtain a high-quality item that will allow them to advance faster.[2] A statement by the Entertainment Software Association, the trade association for the U.S. video game industry, captures this argument well: "Loot boxes are a voluntary feature in certain video games that provide players with another way to obtain virtual items that can be used to enhance their in-game experiences" [gameinformer.com, 2018].

2.2 Regulators' View

While loot boxes have become instrumental for video game companies to generate revenues, they have also received attention from regulators, who often have a less favorable view of

---

[1]Major titles can cost hundreds of millions of dollars to produce, e.g. see gamedeveloper.com [2018].

[2]This argument is akin to the indivisibility of expenditures argument in explaining the demand for lotteries, especially in low-income areas [Kwang, 1965, Rosen, 1997].

the practice. Loot boxes are paid lotteries and share many features with gambling, which is regulated in most countries. When opening loot boxes, players receive a randomized selection of items, suggesting that players potentially get a similar direct utility from uncertainty as they would when gambling, for instance, in casinos [Zendle et al., 2019].[3] This similarity between loot boxes and other gambling contexts suggests that issues associated with problem gambling, such as addiction, may also be associated with loot boxes.[4] Rare loot box outcomes give players similar psychological arousal and rewards as slot machines [Larche et al., 2021]. These concerns are particularly pronounced since many players are minors [e.g. Kristiansen and Severin, 2020], who are prohibited from gambling in many jurisdictions.[5]

The controversy around loot boxes is evident in how varied the regulatory response has been across jurisdictions. Some countries have chosen to classify loot boxes as gambling and to regulate or ban them. For instance, Belgium has banned all loot boxes, arguing that these are games of chance [screenrant.com, 2018, Belgium Gaming Commission, 2018]. China has requested game producers to disclose the probabilities of items received from loot boxes [gamedeveloper.com, 2016]. In contrast, New Zealand and Poland have declared that loot boxes do not constitute gambling [e.g. gamedeveloper.com, 2017]. A number of countries are still investigating the nature of loot boxes, including the United Kingdom [gov.uk, 2022] and the United States [Federal Trade Commission, 2020].

A core element of this debate is whether paid loot boxes allow users to enjoy the game more or provide value as a stand-alone product. If loot boxes were a stand-alone product, they would be considered a "game of chance" (in which skill plays a limited role) rather than a product feature that constitutes part of a video game that is an overall "game of skill" (in which higher-skilled players are more likely to win), allowing to classify loot boxes as gambling.[6] However, if players pay for loot boxes because they value them as part of the overall game, a similar argument does not apply. This logic was used in a 2022 ruling by the Dutch Administrative Jurisdiction Division that overturned a lower court ruling that loot boxes in a popular video game constituted gambling [Dutch Council of State, 2022]. Instead, the Division stated that loot boxes played a functional role in an overall game of skill, and the loot boxes could not be evaluated as an isolated product.

---

[3]Some evidence suggests that players themselves liken the activity to gambling [pcgamer.com, 2017].

[4]Zendle et al. [2019, 2020] show there is a correlation between opening loot boxes and future problem gambling behavior, and the United Kingdom's National Health Service argued in 2020 that loot boxes lead to gambling addiction among minors [NHS, 2020]. A study by the Netherlands' gaming authority has argued that loot boxes "have integral elements that are similar to slot machines," including a "near miss" effect of almost winning something and visual cues, among other things [pcgamer.com, 2018].

[5]Most jurisdictions that allow gambling require people to be 18 years old or older to be allowed to gamble [casino.org, 2020].

[6]Another important element of the regulatory discussion is whether the items received from loot boxes are items of "value," monetary or otherwise. For a succinct discussion of the legality of loot boxes, see jdsupra.com [2019].

3 Empirical Context

3.1 Game Description

We focus on a Japanese free-to-play mobile puzzle game that was available from April 2015 till July 2019. There are 173 stages. In each stage, a grid of colored gemstones is presented. A virtual (computer) enemy and the player take turns "attacking" their opponent, with the player doing it by connecting gemstones of the same color. The player's "attack" is enhanced by the quality of the characters that the player has picked from her inventory to accompany her in a game stage — having better characters enables the player to attack more effectively. The optics of the game is prototypical of popular "connect-the-dots" games, such as Candy Crush saga. Figure 1a shows an example of one stage that is visually similar to the focal game; enemies are displayed in the top section of the screen, and gemstones are in the middle.

The goal of the game is to progress through all 173 stages. The players are guided through the "roadmap" of stages, similar to one depicted in Figure 1b. They can proceed to the next stage only after completing the previous stage.[7] As the stages progress, the game generally becomes more difficult, thereby requiring higher player skills and better inventory.

The characters — the main virtual items that players can collect in their inventories — are called "divers." Before playing a stage, players pick divers from their inventory (up to four divers) that they will use to assist in their attacks. The player starts from a limited inventory of divers and gets more divers as the game progresses. There are two primary mechanisms of receiving new divers. Firstly, players can collect and strengthen their characters through organic gameplay, namely by clearing stages, but this requires a substantial investment of time and effort by players. Secondly, at any time players can open loot boxes.[8] Opening a loot box always results in the player obtaining one diver or other items, but the rarity and color – vertical and horizontal attributes of divers – can vary. For instance, the player may receive a diver that is comparable or inferior to a diver that she already has in her inventory. Figure 1c depicts an example of information that is disclosed on a loot box purchase screen. Players observe the probability distribution of possible loot box outcomes before purchasing a loot box.

Purchases in the game are made with virtual currency we will call "coins." Players can get coins organically, as they make progress in the game – for instance, winning a stage for the first time gives a player one to three coins. Users can also purchase coins with real money. The price of one coin varies from roughly 60 to 120 cents, due to non-linear pricing. For example, 12 coins were frequently available for $8.10.

Players use coins primarily to purchase loot boxes. Loot boxes have a non-linear pricing schedule. One loot box has a regular price of five coins, while a set of 11 loot boxes is commonly

---

[7]Apart from these 173 stages, the players can play special events, which can be weekly or more idiosyncratic. Yet, the 173 stages represent the game core that is shown on the main screen and highlighted as the main objective.

[8]There are two types of loot boxes in this game. "Rare loot boxes" are paid for using in-game currency ("coins"), while "normal loot boxes" are opened with in-game points. The latter tend to provide divers of low rarity and are less relevant for building up the inventory. Throughout the analysis, we focus on rare loot boxes; from now on, we mean "rare loot boxes" when saying "loot boxes" unless we specify otherwise.

50 coins, a five-coin discount. Players may receive other discounts; the most typical discount is when a user opens a loot box for the first time at specific parts of the game; in this case, one loot box costs 3 coins and 11 loot boxes cost 40.

In this game, apart from loot boxes, there are other microtransactions that players can purchase. For instance, the game allows for buy-ins to continue playing a stage that the user is just about to lose. However, these alternative in-game microtransactions are much less popular — loot boxes account for 96% of in-game expenditures.

Figure 1: Example screenshots of a stage game, stages roadmap, and a loot box purchase screen

(a) Example of a stage game   (b) The roadmap of stages   (c) Example of a loot box

Panel (a) depicts a game stage. Enemies are displayed on the top of the screen. The divers that the player has chosen from her inventory to play this round are displayed at the bottom of the screen. The player "attacks" by connecting gemstones of the same color. The more gemstones are connected, the stronger the divers' attacks. Panel (b) presents the screen caption of the stages roadmap; players need to complete stage 1 to go to stage 2, stage 2 to go to stage 3, etc. Panel (c) is an example of information displayed before a user purchases a loot box. The screen shows the probabilities of getting a diver of a particular "rarity" — a measure of quality. These figures have been created by a professional artist for the purposes of illustration and are visually similar to the focal game.

3.2 Data

We now describe our data, provided by the gaming company and extracted directly from the company's production databases. The data contains the universe of observations of all player actions, gameplay and loot box realizations, expenditures, and game feature descriptions.

Play and Loot Box Data. The most substantial part of our data are logs of users' actions. In the gameplay log, the key variables are user ID, stage ID, date and time of the action, and an indicator for whether the stage was successfully completed. In the loot box log, the key variables are user ID, date and time of opening a loot box, and the loot box outcome.

There is a total of 2.52 million users who played the game and are recorded in our dataset, who are responsible for a total of 217.9 million play occasions and 49.6 million loot box openings. Out of these, 96.7 million play occasions are of the main stage, and 19.8 million openings are of rare loot boxes. There is an option of purchasing 11 loot boxes at once, for a small volume-based discount – around 2.6 million rare loot box opening occasions come from 11 loot boxes opened at once. A user is assumed to have left the game after her last recorded action.

Table 1 presents summary statistics of users playing the game and opening loot boxes. An average consumer plays 38.4 main stages and opens 8 rare loot boxes. She reaches stage 18, receives 78 coins – 72.8 coins through the gameplay and 5.2 through purchase with real money – and plays on 21.2 unique sessions and on 11.4 unique days. This distribution is heavily skewed – for instance, the median user opens only 3 rare loot boxes, reaches only stage 4, receives only 18 coins, and plays only one session. The right tail of the distribution is very long, with some players opening more than 6 thousand rare loot boxes, purchasing 50 thousand coins with real money, and playing more than 9 thousand sessions. Figure 7 in Appendix A.1 visualizes this heavy right skew in the users' play and loot box openings occasions by plotting their joint distribution. Even using a log scale, the distributions are skewed right. The number of plays and loot box openings are highly correlated; those that play a lot also open many loot boxes.

In-Game Currency and Loot Box Prices. Our data on currency transactions allows us to distinguish between the acquisition of new currency by completing the main game and by purchasing them. We construct a stock of currency the player has at each in-game action, and record whether this action was made with the currency that the player acquired organically or purchased with hard currency.

The second part of Table 1 presents the resulting summary statistics. Most, 183.2 million out of 196.3 million, or 93.3%, of the in-game currency used is obtained organically. The rest, 13.1 million coins, are purchased using real money. Out of the purchased coins, players spend 99.8%, or 13 million; these paid coins are spent mostly right after the purchase. Players are less careful in using all of the organically obtained coins – only 155 million, or 87.7%, are spent.

Players can spend their coins on loot boxes or other microtransactions. For instance, a user can pay to continue playing a stage that she is about to lose. However, the acquired coins are mostly spent on loot boxes – 91.1% of total spending and 95.7% of their spending with coins purchased with real money are on loot boxes.

Interestingly, the usage of paid coins is much more concentrated than the organic coins –

Table 1: Summary statistics across users.

	Min	Mean	Median	Max	SD	Total
# of actions	1	106.31	10	119,007	540.61	267,521,534
- Played main stage games	0	38.44	4	28,586	115.67	96,719,354
- Played event games	0	48.16	0	83,805	378.10	121,186,908
- Opened rare lootboxes	0	7.88	3	6,204	30.54	19,829,420
- Opened normal lootboxes	0	11.84	2	24,426	58.93	29,785,852
Max main stage achieved	0	18.54	4	173	32.27	–
Win share: main stage games	0	0.94	1	1	0.14	–
Win share: event games	0	0.74	0.84	1	0.29	–
Opened 11 rare lootboxes at once	0	1.02	0	3,256	8.45	2,559,307
In-game currency received	0	78	18	178,698	457.52	196,267,039
- through gameplay	0	72.80	18	128,831	359.30	183,187,916
- through a purchase	0	5.20	0	49,867	119.89	13,079,123
In-game currency spent	0	66.76	9	178,606	442.07	168,001,147
- got through gameplay	0	61.58	9	128,739	340.72	154,962,617
- got through a purchase	0	5.18	0	49,867	119.75	13,038,530
Sessions	1	21.21	1	9,102	103.80	53,368,821
Unique days played	1	11.42	1	1,548	45.14	28,744,507
Length of play (in calendar days)	0	38.19	0	1,553	137.29	–

Actions correspond to playing the game or opening lootboxes. A session is defined as a sequence of actions that are no more than 1 hour apart.

while 90% of all organic coins are driven by 31.5% of the players, 90% of paid coins are driven only by 1.5% of the players. Following the industry's convention, we label these high-spending 1.5% of players as "whales." There is a total of 37,294 players in this group responsible for $10,132,617/95,617,205 \approx 10.6\%$ of all in-game actions. We label the rest of players as "non-whales."

We also observe loot box prices. In most observations (89%), the price of a single loot box is 3 or 5 coins, where 3 is a discounted price for the first time a loot box type is opened. The players face a discounted price of 3 coins on 55.8% of occasions. Similarly, in 88.8% of the observations, the price of an eleven-pack bundle is 40 (discounted) or 50 (regular price). The next two most common price levels of 11-packs of loot boxes, 18 and 25, jointly account for another 6.6% of observations.

Divers, Loot Boxes, and Diver Inventory. We have data on the characteristics of divers, the probabilities of getting different divers from loot boxes, as well as actual loot box

realizations. We construct current diver inventory from this data.

There are 2,009 unique divers. Players get divers by making progress in the game, and by opening loot boxes. The two features of the divers that determine their usefulness for the stage game are rarity and color. Rarity is a vertical attribute that determines the overall strength of the diver, varying from 0 to 6 (seven levels). Five colors represent a horizontal attribute, determining the ability of the diver to remove gems of the same color.

We have data on the probability distribution of loot boxes, as well as loot box realizations, from which we construct diver inventory. There are large differences in the divers received from normal and rare loot boxes, stored in the inventory, and used during playing the game (Figure 8 in Appendix A.1). Normal loot boxes are opened with in-game points and are very likely to give lower-quality divers of rarity smaller than three. They are therefore not very relevant for players focused on the functional value of the divers, as most players are likely to have some of the comparable quality in their inventory. In contrast, rare loot boxes almost never provide divers of rarity zero or one, and relatively large probabilities of getting divers of rarity greater than three. Players tend to keep these divers in inventory. Players are more likely to use rare divers in actual play, with divers of rarity four, five, and six played more than the probabilities of receiving them or just having them in inventory. We present a similar picture for diver colors in Figure 9 in Appendix A.1; there are no similar systemic patterns.

3.3   Descriptive Statistics

Game Progress.   Throughout our analysis, we focus on the main stage play – the 173 stages that players need to clear sequentially to "complete" the game. Given the sequential nature, early stages see more unique players and plays than latter stages (Figure 10 in Appendix A.1). While 2.5 million players play and proceed from stage 1, the number of unique users and plays drops rapidly. The number of unique players is only around 1 million by stage 10. Starting from stage 12, there are occasional spikes in the number of plays per stage. This is because some stages are more difficult than others; players may lose and need to re-play to make progress.

Win Probabilities.   In general, higher stages are more difficult. Figure 2 plots win rates by stage, as well as probabilities of opening loot boxes and spending coins. Win rates are calculated by dividing the number of times a stage has been won, by the times the stage was attempted. They are around 95-96% in the first few stages but drop to 81.6% in stage 12, and 65.4% in stage 16. The win rate roughly drops every 4 stages, which are designed to be difficult as players need to defeat a particularly strong character (a "boss"). Win rates decrease as stages progress, with the final stages of the game having win probabilities only of 47.7-48%.

The lower two lines in Figure 2 present the probabilities of opening at least one rare loot box and of spending at least some coins while at each game stage. Around 12% of players open at least one loot box and 18.8% of players spend at least some coins at an average stage. There is a detectable spike in loot box opening and coins spending after winning stage 3, since that is the end of the game tutorial when players are introduced to rare loot boxes and suggested to open one. Otherwise, spikes in loot box openings and coins spendings occur every four levels,

11

Figure 2: Win, Loot Box Opening, and Coins Spending Probabilities for Each Stage
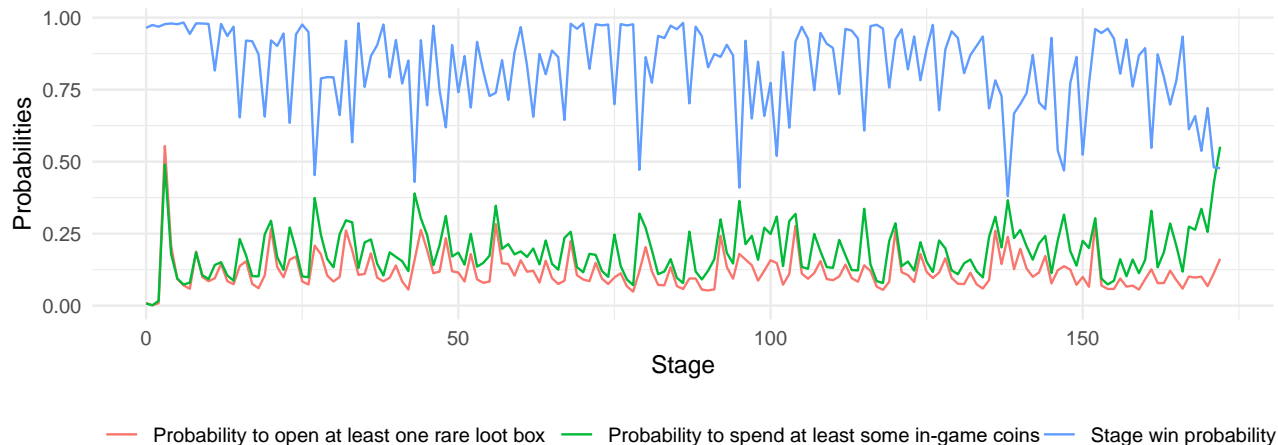
Win rates are calculated by dividing the number of times a stage has been won, by the times the stage was attempted, using only sequential occasions of main stage plays. Probabilities of opening at least one rare loot box and spending at least some in-game coins (in-game currency) are calculated by taking an average across players who are at a given stage.

aligned with the harder "boss" stages. They suggest that players tend to open more loot boxes when they are struggling to make progress in the game.

The probabilities in Figure 2 are nearly identical if we replicate the description using only the 38 thousand players who have reached stage 173 in their gameplay; we present these results in Figure 11 in Appendix A.1. This similarity highlights the lack of selection on skill among players who "survive" till latter stages of the game, as shown in Appendix XXX.

We test the relationship between the complexity of stages and loot box openings by regressing loot box openings and coin spending probabilities on stage win probabilities. Results are presented in Table 7 in Appendix A.1. All variables are in logs, which gives coefficients an interpretation of elasticity estimates based on across-stage correlations. If the win rates of stages are 1 percent lower, the probabilities to open rare loot boxes are 0.77% higher, and probabilities to open a rare loot box for real money are 1% higher. Similarly, on stages with a 1 percent lower win rate players tend to spend 1.7% more coins and 1.4% more coins than they have paid for with real money.[9] Estimates are similar if we use only the 38 thousand players who reached the final stage. Higher probabilities of opening loot boxes at harder stages suggest their functional value; we investigate this relationship more rigorously in Section 5.

---

[9]Probabilities to open at least one loot box (for any or paid coins) and spend at least some (any or paid) coins are highly correlated across stages, with correlations varying between 55% and 97%. Figure 12 in Appendix A.1 presents all four variables together, with variable averages across stages normalized to one to make visual comparisons simpler.

Inventory quality.   As players make process throughout the game, they accumulate more rare divers in their inventory. Figure 3 presents the quality of players' inventory across stages by summing the rarity level of the top four divers in the inventory; only four top divers are the most relevant for gameplay since players can choose up to four divers to play in a given stage. To make inventories comparable across stages, we only use inventories of players who complete the game (those who win all 173 stages). The average increases from 9 in the early stages up to 21.5 by the end of the game. There is a discrete jump after stage 3 since that is when players open a rare loot box as part of the game's tutorial.

Figure 3: Average top-4 diver rarity in inventory across stages
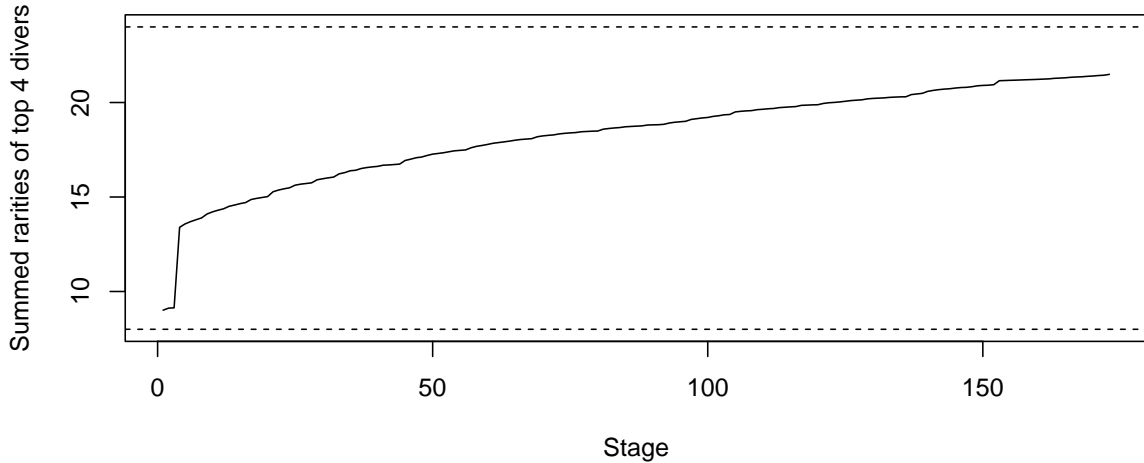
Average inventories are computed by taking the sum of the top four divers' rarity across players, using only the first observation per player per stage (to make sure player weights are equal across all stages). We use only observations of players who reached and won stage 173.

Diver value.   How valuable are divers' rarity for winning in-game stages and making progress? To examine this, we regress a player's likelihood to win a stage on the summed rarity of the top four divers in their inventory, similar to the variable presented in Figure 3. We allow the effect of the summed rarity of the top four divers to be stage-specific by adding stage-rarity interaction terms, and control for stage and user fixed effects to use only within the stage and within user variation.[10] Figure 13 in Appendix A.1 visualizes the estimated stage-specific coefficients of rarity effects on stage win probabilities. The coefficients are positive and statistically significant (at 5% level; clustering done on the stage level) for 155 out of 173 game

---

[10]Appendix A.2 presents the estimates from alternative specifications of regressions, that examine the effect of the rarity of divers in players' inventories on players' win probabilities. The effect of a rarity on win probabilities is consistently strong across different parametric assumptions.

13

stages. On average, one extra rarity point in the sum of rarities across the top four divers increases win probability by 2 percentage points. Out of 18 stages with insignificant estimates of the effects of diver rarities on win probabilities, 10 are in the first 13 stages, presumably because these stages are relatively simple. There is a very slight increase in the magnitude of the effect of diver rarities on win probabilities as the players progress to more advanced game stages, with the fitted effect of inventory rarity on win probabilities increasing from 1.5 percentage points in stage 1 to 2.5 percentage points in stage 173. Overall, these results confirm that diver rarity has strong functional value in the game, and that this functional value of acquiring extra divers is relatively constant throughout the stages.

Returns to loot boxes. Results so far provide suggestive evidence for the existence of the functional value behind loot boxes for players. We have shown that players tend to open more loot boxes at stages they lose, and extra diver rarity they can collect increases their win probabilities. Yet, the combination of evidence presented in Figures 2 and 3 – that the rate of opening loot boxes is relatively flat across stages but that the quality of players' inventory is much higher in latter stages – is hard to rationalize with the functional value of loot box openings. The "return" from loot boxes – expected increases in top divers' rarity – are likely higher in the early than latter stages of the game due to inventory differences.

A piece of descriptive evidence makes this argument more salient. In Figure 4 we visualize the implied "return" from loot boxes in terms of the top divers' rarity. In Panels (a) and (b), we present the probabilities to open a rare loot box per stage of the game – overall (a) and using real money (b) - which are nearly flat across game stages. In the average stage, the user opens a loot box 16.2% of the time and opens it using real money 1.1% of the time. In Panel (c), we visualize the "return" from loot boxes, the expected change in rarity of the top four divers if a player opens one rare loot box. The expected increase in the rarity of top divers in players' inventories is around 0.25-1 in the early stages of the game – meaning that for every loot box opened, a player can expect the rarity of one of the top four divers to go up by 0.25-1 points. It goes down dramatically in the latter stages of the game, all the way to around 0.006-0.01 in the last 50 stages since players have already amassed good inventories. Finally, in Panel (d) of Figure 4 we present the number of loot boxes a player needs to open to increase their expected win probability by 1 percentage point.[11] The functional value of loot boxes dramatically falls across stages; in early stages, 1-2 loot boxes are enough to increase players' win rate by 1 percentage point, while in latter stages a player needs to open 50-60 loot boxes to increase win rate by 1 percentage point. Jointly, Panels (a), (b), and (d) of Figure 4 suggest that loot box openings cannot be explained by the functional value alone– even if we focus only on the immediate loot box returns.[12]

---

[11]We compute this by combining the expected increase in top diver rarity in players' inventory across stages (Panel c) with the effect of one extra rarity point on stage win probability (Figure 13).

[12]Accounting for long-term effects of having high rarity divers in the inventory makes this argument even stronger since the returns from opening loot boxes early on in the game are even higher – early on in the game players may expect to be using divers for more stages.

Figure 4: Returns from opening rare loot boxes by stage
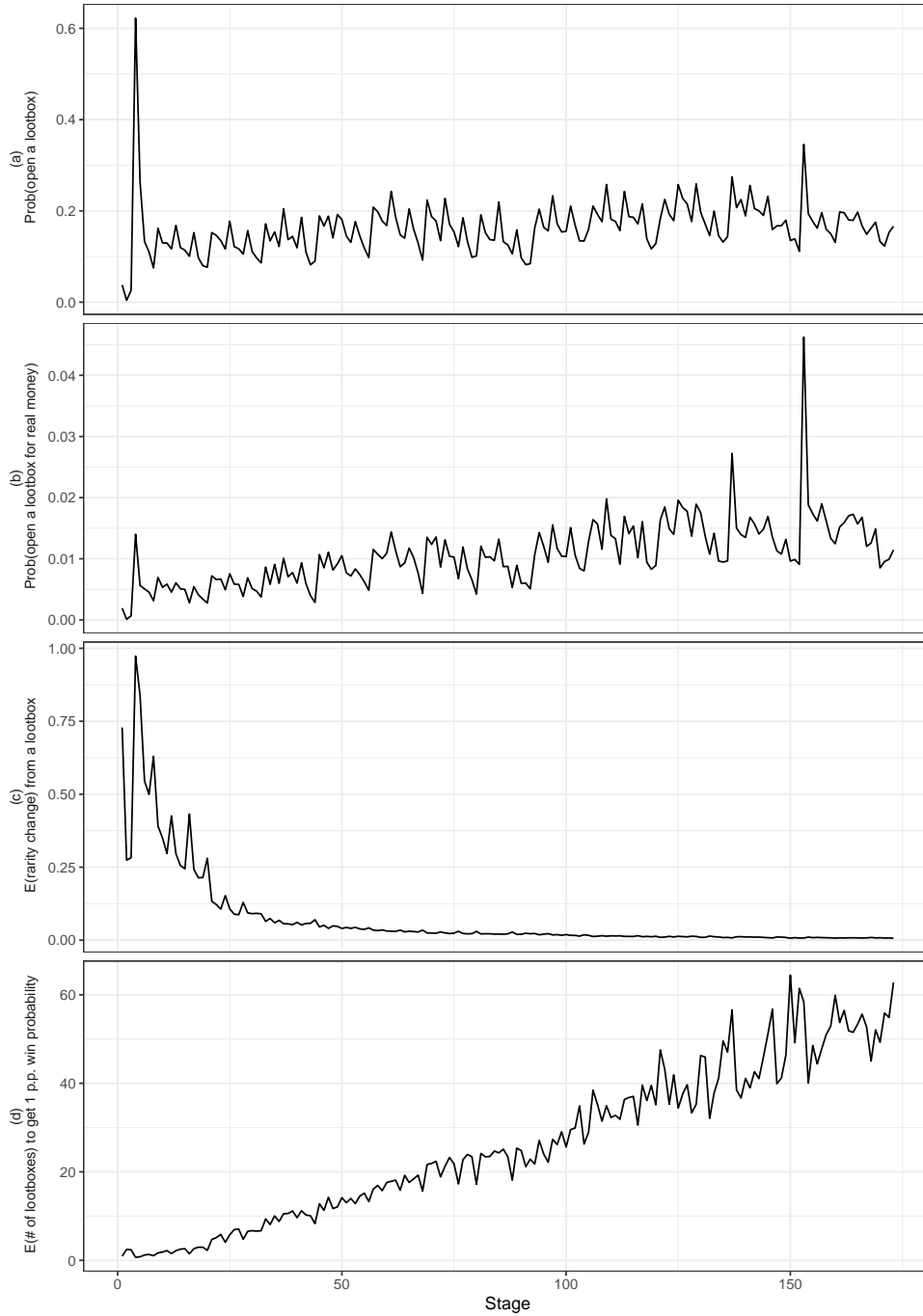
Panels (a) and (b) present probabilities to open a rare loot box per stage of the game, overall and using real money. Panel (c) presents the average change in the rarity of the top four divers in a player's inventory after opening a rare loot box at different stages of the game. Panel (d) presents the number of loot boxes a player needs to open to expect to increase their win probability by 1 percentage point. The expected number of loot boxes is computed by dividing 0.01 by the average change in the rarity of the top four divers in a player's inventory after opening a rare loot box and multiplying by the fitted values from Figure 13, the effect of one rarity point on win probabilities.

15

Currency constraints. Finally, we examine whether currency constraints play a role in players' decisions to open loot boxes. When players want to open a loot box but do not have enough in-game currency to do it, they face a trade-off between playing more and accumulating enough in-game currency, or making a purchase of coins with real money. To understand whether such constraints are binding, we plot the probability of opening a loot box by the accumulated amount of the in-game currency in Figure 14a in Appendix A.1. There are four clear spikes at 3, 5, 40, and 50 coins, corresponding to the prices of 1 and 11 loot boxes (3 is a discounted price sometimes offered instead of 5, and 40 instead of 50). This result shows evidence that constraints are binding. The players may be forward-looking in waiting for more coins to open more loot boxes for a discounted price. To confirm that this effect is not driven by other factors, Figure 14b in Appendix A.1 plots the probability to open a loot box using paid coins, which can be purchased at any point. Results are strikingly different; there are no spikes at the stocks of 3 and 5 coins. There are only small spikes at the stock of 40 and 50 coins, potentially corresponding to players using the coins they have purchased earlier.

4 Model

4.1 A Toy Example

We first present a simple model of gaming with loot boxes to build intuition about the functional and direct value of loot boxes.

Consider a static problem of a consumer who makes two discrete choices: whether or not to play the game, $Y_G = \{0, 1\}$, and whether or not to open a loot box in the game, $Y_L = \{0, 1\}$. As is common in games, users extract value from making progress in the game (e.g. by "winning" a stage or "clearing" a puzzle). If a consumer chooses to play ($Y_G = 1$) she gets

$$u(Y_G = 1, Y_L) = \alpha_G + \beta Pr(\text{win}|Y_L), \tag{1}$$

where α_G is a preference for playing the game and β is a preference for winning in the game.

Before playing, the consumer does not know whether she will win and takes an expectation over her odds of winning. It is a function of Y_L, the decision to open a loot box. Opening a loot box ($Y_L = 1$) can weakly increase the win probability, $Pr(\text{win}|1) - Pr(\text{win}|0) \geq 0$ – i.e. because opening a loot box gives the player a chance to get items that will help her win.

Apart from the gaming utility, a consumer that opens a loot box ($Y_L = 1$) gets utility

$$u(Y_L = 1) = \alpha_L - \gamma p, \tag{2}$$

where α_L is the direct utility from the loot box. This may be a "true" normatively respectable preference over the uncertainty or a taste that is driven by misperceptions from behavioral biases. We denote the price of the loot box as p, and γ is the marginal utility of currency. The consumer gets zero utility if she does not play the game and does not open the loot box.

4.2 An Empirical Model

We now build a more formal model of gaming with loot boxes that is tailored to our data.

A player i makes progress sequentially through 173 stages. After the player completes all stages, she can continue replaying the stages. On choice occasion t, player i decides one of four actions a_{it} to take: play the game ($a_{it} = 1$), open a single loot box L_{rt} ($a_{it} = 2$), open an eleven pack of loot boxes ($a_{it} = 3$), or leave the game forever ($a_{it} = 0$). The player makes these decisions given the inventory of the divers that she holds, D_{it}; the accumulated stock of the in-game currency, c_{it}; the current stage she is playing s_{it}; an indicator whether the player has lost round s_{it} before, q_{it}; a variable d_{it} that captures state dependence in loot box openings in a primitive fashion, as we define below; and prices of opening 1 and 11 loot boxes, p_{it}^1 and p_{it}^{11}.

Playing the stage game ($a_{it} = 1$). The non-random contemporaneous utility of choosing $a_{it} = 1$ is

$$u\left(a_{it} = 1\right) = \alpha_{G,s_{it}} - \beta q_{it} \tag{3}$$

where $\alpha_{G,s_{it}} = \alpha_{G,s}$ is the stage-specific utility of playing the stage game, and q_{it} is an indicator variable that takes the value of one if the player has lost the current stage s before. This captures the disutility of the player of losing and having to reply the same stage. In the empirical context, players make progress through a sequential series of stages. The $-\beta$ captures the disutility of having lost the current stage and having to replay it – anticipating this, players may seek better characters to make progress in the game. The parameter is assumed to be common across players and stages.

Opening a single loot box ($a_{it} = 2$). Instead of playing the stage, a consumer can open loot boxes. Consider the case of $a_{it} = 2$, in which the player opens one loot box. We assume that the loot box a player can open is determined by the current stage, $L_s = L_{s_{it}}$. With probability $\Pr_s$ she gets a diver d from a loot box L_s, $d \in D_{L_s}$, updating her inventory of divers to $D_{i,t+1} = \{D_{it}, d\}$.[13]

Opening a loot box comes at a cost of p_{it}^1 coins, subtracted from the stock of in-game currency, $c_{i,t+1} = c_{it} - p_{it}^1$.[14] Following the expenditure patterns of in-game currency, we assume that depletion from the stock itself is not utility decreasing. The stock is likely to consist mainly of coins obtained for free, which consumers do not fully use. On the other hand, if $p_{it}^1 > c_i$, the consumer spends hard currency to purchase $p_{it}^1 - c_{it}$ drops.

A consumer obtains (constant) direct thrill utility from opening a loot box, $\alpha_{L,1}$. On top of this, if a consumer opens a loot box immediately after another loot box, she gets an extra thrill value η, which captures inertia in openings of loot boxes. We capture this behavior with an indicator variable d that captures if the play's previous action was also to open a loot box,

$$d_{it} = \mathbb{1}\left(a_{i,t-1} \in \{2,3\}\right). \tag{4}$$

---

[13]Inventory has a capacity constraint. If the constraint is met, we assume the player keeps the best divers.

[14]The price p_{it}^1 can vary by player and play occasion since price discounts can depend on the timing of play, as well on the previous actions of the player, as we discuss in Section 3.2.

Combining these terms, the non-random component of utility from action $a_{it} = 2$ is

$$u\left(a_{it} = 2\right) = \alpha_{L,1} - \gamma \mathbb{1}\left(p_{it}^1 > c_{it}\right) \times \left(p_{it}^1 - c_{it}\right) + \eta d_{it} \tag{5}$$

where γ is the marginal (dis)utility of purchasing coins.

Opening an eleven-pack of loot boxes $\left(a_{it} = 3\right)$. As we described previously, distinguishing between purchases of one and eleven-packs of loot boxes is important given the non-linear pricing. We allow for a separate option for consumers to open eleven loot boxes, introducing action-specific coefficients on direct utility $(\alpha_{L,11})$, as well as a purchase (γ) and state-dependence (η) coefficients. The non-random component of utility from choosing $a_{it} = 3$ is

$$u\left(a_{it} = 3\right) = \alpha_{L,11} - \gamma \mathbb{1}\left(p_{it}^{11} > c_{it}\right) \times \left(p_{it}^{11} - c_{it}\right) + \eta d_{it}. \tag{6}$$

Choosing to leave the game $\left(a_{it} = 0\right)$. A player can choose to leave the game forever, $a_{ts\tau} = 0$. The terminal utility is normalized to zero,

$$u\left(a_{it} = 0\right) = 0. \tag{7}$$

Utility maximization. Given these choice-specific utilities, the player i decides on an action a_{ij} that maximizes the present-discounted value of the future stream of utilities,

$$max_{\{a_{it} \; \forall t\}} E \sum_{t=1}^{\infty} \beta^{t-1} \left[u_{it}(a_{it}, O_{it}; \theta) + \varepsilon_{iat}\right] \tag{8}$$

where θ are the model parameters to be estimated, $O_{it} = \left\{R_{it}, c_{it}, s_{it}, q_{it}, d_{it}, p_{it}^1, p_{it}^{11}\right.$ are state variables, and ε_{iat} are player, choice, time specific idiosyncratic shocks.

Using Bellman's equation, we define the value function as

$$V\left(O_{it}, \varepsilon_{iat}\right) = \max_{a_{it} \in \{0,1,2,3\}} u\left(a_{it}\right) + \varepsilon_{iat} + \beta E_{O', \varepsilon' | O_{it}, a_{it}} V\left(O', \varepsilon'\right) \tag{9}$$

Assuming that the idiosyncratic shock ε_{iat} is distributed Type-1 extreme value, we can express the conditional choice probabilities of choosing action a_{it} as

$$CCP\left(a_{it} | O_{it}\right) = \frac{exp\left(u\left(a_{it}\right) + \beta E_{O', \varepsilon' | O_{it}, a_{it}} V\left(O', \varepsilon'\right)\right)}{\sum_{\tilde{a}_{it} \in \{0,1,2,3\}} exp\left(u\left(\tilde{a}_{it}\right) + \beta E_{O', \varepsilon' | O_{it}, \tilde{a}_{it}} V\left(O', \varepsilon'\right)\right)} \tag{10}$$

State Transitions. Here we summarize the transitions of the state variables O_{it}. We discuss the empirical estimates of transition probabilities in Section 6 and Appendix A.3.

There are two sources of updates to diver inventory, D_{ij} – the player receives divers when making progress in the game and when opening loot boxes.

The currency stock, c_{ij}, is also updated either via organic in-game progress as she makes progress in the stage game, or purchases of the currency with real money. Currency decreases by as much as p_{it}^1 or p_{it}^{11} when the player chooses to open a loot box.

The current stage of the game, s, is updated when a player wins the stage. The win probability is affected by which stage the player is on (stages differ in complexity), what is the rarity of divers in the player's inventory, and whether the player has lost this stage before (they play it the second time), $\Pr(\text{win}|s_{it}, q_{it}, D_{it})$. To make the estimation of the effect of the player's inventory on win probability tractable, we approximate the player's inventory as a function of the rarity of divers she has, as rarity is the primary vertical dimension by which divers are differentiated functionally. For this, we track the top four rarity divers, as players can use at most four divers from inventory for any given stage of gameplay. We then create a single index, $R_{it} = \sum_{l=1}^{4} \text{rarity}_{itl}$, a summed rarity of the top four divers in the inventory.[15]

The probability to win is

$$\Pr(\text{win}|s_{it}, q_{it}, D_{it}) = \zeta_{1,s,q} + \zeta_{2,s,q} * R_{it} + \zeta_{3,R_{it}} \tag{11}$$

where R_{it} summarizes the rarity of divers in player i's inventory at time t. Coefficients $\zeta_{1,s,q}$, $\zeta_{2,s,q}$ and $\zeta_{3,R_{it}}$ allow for stage-and-loss-specific effects of inventory rarity on the win probability.[16] When the player loses a stage, s_{it+1} is not updated ($s_{it+1} = s_{it}$) and q_{it+1} is updated to one ($q_{it+1} = 1$). If she wins a stage, s_{it+1} is updated to $s_{it} + 1$ and q_{it+1} is set to zero.

The state-dependence variable d_{it} is updated using the formula in Equation 4.

The transition of the loot box prices mainly depends on the previous actions of the player. The player faces a discounted price the first time she opens a single or eleven-bundle of loot boxes. We allow price states to transition across states ({-3,-5} for one loot box and {-18, -25, -40, -50} for eleven-packs) given the empirical distribution of these transition probabilities conditional on stages and past actions of players to account for other occasional discounts.

5 The Reduced-Form Evidence

Before estimating the full model, we present the reduced-form evidence on the tastes of whales and non-whales. While we need a structural model to account for players' forward-looking behavior and expectations about state transitions, short-term reactions of players to exogenous shocks in their game performance and loot box outcomes allow us to characterize the directions of taste differences between whales and non-whales without imposing any restrictive assumptions.

5.1 Effect of Winning the Stage

The first shock we leverage is whether the player has won or lost the stage game, controlling for the player's skill, inventory, and progress so far. Whether the player wins a stage depends on realizations of stage complexity, driven by how close different colors of gemstones are (randomly) allocated throughout the stage. We leverage this randomness in players' stage performances to examine the effect of winning or losing on immediate play and loot box decisions.

---

[15]Since divers of rarities 0-2 do not vary much in their characteristics, we pool them together and assign them rarity 2. Thus, there are 17 possible values of the sum of the top 4 diver rarities, from 8 to 24.

[16]Appendix A.3 presents the estimates of the win probability function.

Building on equation 3, we estimate the effect of stage loss on the probability of opening a loot box (instead of playing again)

$$\mathbb{1}(a_{it} \in \{2,3\}|a_{i,t-1} = 1) = \kappa_i + \kappa'_{s,R_{i,t-1}} + c\,I(\text{win}_{i,t-1}) + \xi_{it}, \tag{12}$$

where κ parameters correspond to user and stage by accumulated inventory rarity (period $t-1$) fixed effects and $I(\text{win}_{i,t-1})$ is an indicator of whether the player has just won while playing the game in the previous choice occasion. User fixed effects control for any differences in skills across players, and stage by accumulated inventory fixed effects control for the inventory and game progress. All observations used are conditional on playing the game in period $t-1$. We use the main stage play and rare loot box openings as two types of actions available. We cluster the standard errors two-way, on the user and stage level.

Table 2: The effect of having won the stage game in period $t-1$ on the likelihood to open a loot box in period t

	Dependent variable: $I(a_{it} \in \{2,3\}|a_{i,t-1} = 1)$				
	All	Stage < 87	Stage ≥ 87	Non-whales	Whales
	(1)	(2)	(3)	(4)	(5)
$I(\text{win}_{i,t-1})$	-0.0384***	-0.0442***	-0.0262***	-0.0393***	-0.0296***
	(0.0044)	(0.0043)	(0.0054)	(0.0041)	(0.0065)
Number of Fixed Effects:					
Stage x Inventory Rarity	3,243	2,094	1,149	3,197	3,120
User	2,436,145	2,435,963	129,998	2,398,851	37,294
Average $I(a_{it} \in \{2,3\}|a_{i,t-1} = 1)$	0.0768	0.0809	0.0676	0.0766	0.0779
R^2	0.1577	0.1891	0.0544	0.1708	0.0525
Number of observations	95,617,362	73,617,737	21,999,625	85,484,745	10,132,617

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

All observations are conditional on the user playing a main stage in period $t-1$. All specifications include stage by accumulated inventory rarity and user fixed effects. All standard errors are clustered two-way, on the user and stage level.

Table 2 presents the estimates across stages and player types. In Column (1), we present the effect of winning the stage on the probability of opening a loot box based on all relevant observations in our data. If a player just won the game, she is 3.8 percentage points less likely to open a loot box – a 50% decrease in loot box play probability compared to the baseline average of loot box openings (after playing the stage) of 7.7 percentage points. The probability of playing another stage respectively increases by 3.8 percentage points.

In Columns (2) and (3), we split the sample by game stages, estimating the effect separately for players in the first and second half of the game (below and above stage 87, respectively). Point estimates of the effect of a stage win on the rate of opening loot boxes are slightly stronger early on in the game – a win decreases the loot box opening probability by 4.4 and

20

2.7 percentage points before and after stage 87, respectively – but given the standard errors this difference is only marginally significant.

In Columns (4) and (5), we now split the sample into whales and non-whales. We find that the effect of a stage win on the rate of opening loot boxes is slightly stronger for non-whales than for whales – a win decreases the loot box opening probability by 3.9 and 3 percentage points for non-whales and whales, respectively. However, this difference is driven entirely by the difference in the composition of stages played by whales and non-whales; proportionally, whales play more latter stages of the game. Overall, the estimates in Table 2 confirm that losing a stage decreases the value of play for whales and non-whales, consistent with the preference for winning in the game for both groups.

5.2   Effect of the Loot Box Outcomes

Above, we have shown that players dislike losing and repeating the same game stage. If players open loot boxes for their functional value, a more positive outcome (a rarer diver) received from a loot box should increase the gameplay utility, since it improves the players' inventory and increases their probability of winning. To examine the effect of players' inventory rarity and loot box outcomes, we estimate

$$I(a_{it} \in \{2,3\}|a_{i,t-1} \in \{2,3\}) = \kappa_i + \kappa'_{s,R_{i,t-1}} + bR_{it} + \xi_{it}, \tag{13}$$

where κ parameters correspond to user and stage by accumulated inventory rarity (period $t-1$) fixed effects and R_{it} is the inventory rarity in period t.[17] All observations used are conditional on opening a loot box in period $t-1$. As above, we use only playing the main stage and opening rare loot boxes as two types of actions available. We cluster the standard errors two-ways at the user and stage level.

Since players can have different rates of loot box openings – and players who opened more loot boxes in the past are both more likely to have accumulated better inventory and to open more loot boxes in the future – we cannot regress $I(a_{it} \in \{2,3\}|a_{i,t-1} \in \{2,3\})$ on R_{it} directly. Instead, we instrument for R_{it} with rarity$_{L_{it-1}}$, the realized diver rarity of the loot box just opened in period $t-1$.[18] By construction, this rarity realization is random and has an impact on a player's inventory before it is built up to a perfect level, making it a valid instrument.

The first part of Table 3 presents the estimates of b from the instrumental variable regression 13, as well as first stage results, using all players in our data. Column (1) presents the estimates for the entire sample. On average, one extra rarity point of the diver received from a loot box by players, rarity$_{L_{it-1}}$, increases the rarity of the top four divers in a player's inventory on average by 0.123 points. This increase, in turn, decreases the probability that a player opens another loot box by 0.151 * 0.123 * 100 = 1.86 percentage points. Put differently, a full point increase in R_{it} on average corresponds to a 100 * 0.151/67.7 ≈ 22.3% decrease in the baseline probability to keep opening loot boxes.

---

[17]We abuse notation by using the same notation for the nuisance parameters as in regression 12.

[18]If a player opens a pack of loot boxes at once, we count the highest rarity among divers drawn as rarity$_{L_{it-1}}$.

21

Table 3: The effect of the loot box outcome in $t-1$ on the likelihood to open a loot box at time t

	Dependent variable: $I(a_{it} \in \{2,3\}|a_{i,t-1} \in \{2,3\})$		
	All	$R_{it-1} < 16$	$R_{it-1} \geq 16$
	(1)	(2)	(3)
I. All players:			
$\hat{R}_{it}$ (IV: rarity$_{L_{it-1}}$)	-0.1508***	-0.0837***	-0.6358***
	(0.0480)	(0.0069)	(0.1788)
First stage ($R_{it} \sim$ rarity$_{L_{it-1}}$)	0.1231***	0.4447***	0.0262***
	(0.0376)	(0.0386)	(0.0077)
Average $I(a_{it} \in \{2,3\}|a_{i,t-1} \in \{2,3\})$	0.6771	0.5476	0.7472
R^2	0.3092	0.4047	0.2348
Number of observations	18,419,425	6,466,010	11,953,415
II. Non-whales:			
$\hat{R}_{it}$ (IV: rarity$_{L_{it-1}}$)	-0.1725***	-0.0880***	-0.8094***
	(0.0513)	(0.0075)	(0.2168)
First stage ($R_{it} \sim$ rarity$_{L_{it-1}}$)	0.1562***	0.4458***	0.0320***
	(0.0383)	(0.0380)	(0.0075)
Average $I(a_{it} \in \{2,3\}|a_{i,t-1} \in \{2,3\})$	0.6439	0.5471	0.7141
R^2	0.3236	0.4155	0.2552
Number of observations	14,563,179	6,116,584	8,446,595
III. Whales:			
$\hat{R}_{it}$ (IV: rarity$_{L_{it-1}}$)	-0.0080	-0.0313***	-0.0212
	(0.0363)	(0.0040)	(0.0952)
First stage ($R_{it} \sim$ rarity$_{L_{it-1}}$)	0.0400**	0.4238***	0.0152**
	(0.0171)	(0.0435)	(0.0061)
Average $I(a_{it} \in \{2,3\}|a_{i,t-1} \in \{2,3\})$	0.8025	0.5571	0.8269
R^2	0.1537	0.2229	0.1238
Number of observations	3,856,246	349,426	3,506,820

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

All observations are conditional on the user opening a rare loot box in period $t-1$. All specifications include stage-by-rarity inventory in period $t-1$ and user fixed effects. All standard errors are clustered two-way, on the user and stage level.

An average effect based on the entire sample masks important differences in random realizations of loot boxes. As we have discussed above, drawing more rare divers should be more important for players with a less developed inventory. To examine the difference in the effects of rarity$_{L_{it-1}}$ on gameplay, Columns (2) and (3) of the first part of Table 3 break down the sample to players with inventory in $t - 1$ below and above 16 points. As expected, the effect of rarity$_{L_{it-1}}$ realization on diver inventory R_{it} is much more pronounced for players with less developed inventories. A one-point increase in the rarity of a received diver improves the top four diver rarity in the inventory by the expected 0.445 points if the inventory rarity was below 16, but only by 0.026 points if the inventory rarity was above 16. The implied effects of one extra rarity$_{L_{it-1}}$ point on the probabilities to open another loot box are 0.084 * 0.445 * 100 = 3.74 and 0.646 * 0.026 * 100 = 1.68 percentage points decrease, respectively. The latter effect is smaller particularly since one rarity$_{L_{it-1}}$ point is less useful for players with stronger inventories. Overall, both effects confirm that players are more likely to switch to playing the game if the quality of their inventory improves.

The second and third parts of Table 3 further splits the sample by non-whales and whales. For non-whales (part II) we find that results are similar to the full sample – all columns show that getting a better diver from a loot box increases the players' probability to switch from opening loot boxes to playing the game, and the effects are stronger for players with less developed inventories. The magnitudes of the estimates are similar to the magnitudes in part I based on the full sample of players.

For whales (part III), the results are drastically different. While the first stage effects are the same as before – the effect of one extra rarity$_{L_{it-1}}$ point on diver inventory R_{it} is 0.424 for the subset with $R_{it-1} < 16$ (Column 2) and 0.015 for the subset of observations with $R_{it-1} \geq 16$ (Column 3) – the implied effects of incremental R_{it} on gameplay or loot box opening decisions are much smaller. First, for whales with weaker inventory, $R_{it-1} < 16$, improvement of the inventory (R_{it}) by one extra rarity point decreases their probability to open another loot box only by 3.1 percentage points – in contrast to 8.8 percentage points for non-whales. This implies that one extra rarity point in rarity$_{L_{it-1}}$ realization decreases the probability that a player opens another loot box by 0.031 * 0.424 * 100 = 1.31 percentage points (3.9 for non-whales). For whales with stronger inventory, $R_{it-1} \geq 16$, there is no detectable effect of the improvement of the inventory (R_{it}) on loot box opening decisions. The effect of one extra rarity point in rarity$_{L_{it-1}}$ on a player's probability to loot again is 0.0212 * 0.0152 * 100 = 0.03 percentage points, an effect that is estimated very precisely (standard error of 0.133, projecting $I(a_{it} \in \{2,3\}|a_{i,t-1} \in \{2,3\})$ directly on rarity$_{L_{it-1}}$).[19]

The difference in the estimates for non-whales and whales suggests the different sources of value that loot boxes provide to these player groups. In particular, the lack of reaction of whales to the diver rarity realizations suggests that direct value might be driving their loot

---

[19]We confirm that the difference in estimates for non-whales and whales is not driven by the overall engagement of players with the game: we get statistically similar point estimates for non-whales if we condition on the 1.5% of players by the number of actions. If anything, the magnitudes of the IV estimates for this group are larger in magnitude; the IV estimates are -0.809 (s.e. of 0.36) for all rarity levels, -0.114 (0.01) for rarity levels < 16, and -3.282 (s.e. 1.25) for rarity levels above or equal to 16.

box consumption. We now proceed to estimate the structural utility model to quantify the relative importance of the functional and direct values of loot boxes.

6 Estimation

To estimate users' tastes, we use a two-step procedure as in Hotz and Miller [1993]. Below we describe how we operationalize the state space and estimate empirical conditional choice probabilities (CCPs). We then describe how the terminal action of users – leaving the game – simplifies the estimation procedure, following Arcidiacono and Miller [2011].

States. In our model, the state variables are $O_{it} = \{R_{it}, c_{it}, s_{it}, q_{it}, d_{it}, p_{it}^1, p_{it}^{11}\}$. There are 17 states of R_{it}, the sum of the rarity for the top four divers. We allow for currency stock to lie between integers of 0 and 51, as this captures the vast majority of observations and the full range of observed loot box prices. The user's current stage s_{it} ranges from 0 to 173. The indicator for whether the player has lost the stage, q_{it}, and the state-dependence state, d_{it}, both vary between zero and one — or 2 levels. Finally, the price of a single loot box opening takes one of two states, $\{3, 5\}$, while the price of a bundle of eleven loot boxes takes four price levels, $\{18, 25, 40, 50\}$. Thus, there is a total of 4.9 million possible states.

We estimate the state transition probability matrices from the data, using equation 11 for the win probability estimates and frequency estimators for the rest of the state variables. Appendix A.3 provides more details on the estimation procedure and estimates.

Terminal Action Conditional Choice Probabilities (CCPs) Estimation.
CCPs are the probability that an agent optimally chooses an action a given her state O.

We estimate the model parameters using empirical estimates of the CCP for the terminal action $a_{it} = 0$. To see why we only need the CCP of leaving the game for estimating the utility parameters, note that we can express the integrated value function, $\int_{\epsilon_{it}} V(O_{it}, \varepsilon_{iat}) dF(\epsilon_{iat})$, as a function (denoted by $\psi(\cdot)$) of the CCPs and the choice-specific value function $v_a(O_{it}) = u(a|O_{it}) + E_{O', \varepsilon'|O_{it}, a_{it}} V(O', \varepsilon')$,

$$\int_{\epsilon_{it}} V(O_{it}, \varepsilon_{it}) dF(\epsilon_{iat}) = \psi_a[CCP(a|O_{it})] + v_a(O_{it}). \tag{14}$$

as shown by Arcidiacono and Miller [2011]. Using this expression for the terminal action, $a = 0$, and the properties of the Type-1 extreme value distribution of ϵ_{it}, we get that

$$\int_{\epsilon_{it}} V(O_{it}, \varepsilon_{it}) dF(\epsilon_{iat}) = -\log(CCP(a_{it} = 0|O_{it})) + (\text{Euler constant}), \tag{15}$$

where $CCP(a_{it} = 0|O_{it})$ is the conditional choice probability of the terminal action.

We can directly estimate the empirical analog of the CCP for the terminal action $a_{it} = 0$ from the data. We fit

$$\mathbb{1}(a_{it} = 0|R_{it}, c_{it}, s_{it}, q_{it}, d_{it}) = a_{sq}^1 + a_{sd}^2 + a_{sq}^3 R_{it} + a_s^4 c_{it} + a_s^5 c_{it}^2 + a_s^6 c_{it}^3 + \xi_{it}, \tag{16}$$

where a are the coefficients of interest. To keep the relationship between the states and players' actions flexible, we allow for stage-by-lose indicator and stage-by-state-dependence parameter fixed effects (a^1_{sq} and a^2_{sd}), the stage-by-loss-specific slopes on the rarity of the inventory (a^3_{sq}), and the stage-specific third order polynomials of the currency stock.[20]

Utility Parameters' Estimation. Given the empirical estimates of the terminal CCPs, we estimate the parameters of the players' utilities using a simple weighted least squares regression. This requires us to obtain estimates for $\beta E_{O',\varepsilon'|O_{it},a_{it}} V(O',\varepsilon')$ for each state O_{it} and action a_{it} combination.

Using Equation 15, we obtain estimates of $\int_{\epsilon_{it}} V(O_{it},\varepsilon_{it}) dF(\epsilon_{iat})$ from empirical terminal CCPs. Then, using transition probabilities of the state variables estimated in the first step, we compute the expected value function, $\beta E_{O',\varepsilon'|O_{it},a_{it}} V(O',\varepsilon')$, by integrating over the expected transitions of the states and the chosen action. We set $\beta = 0.99$.[21]

These estimates of $\beta E_{O',\varepsilon'|O_{it},a_{it}} V(O',\varepsilon')$ allow us to express the current period utilities of consumers' actions using the inversion in Berry [1994],

$$log(sh^a_{it}) - log(sh^0_{it}) - \beta E_{O',\varepsilon'|O_{it},a_{it}} V(O',\varepsilon') = u(a_{it}) + \xi_{it}, \tag{17}$$

where sh^a_{it} are the empirical probabilities of choosing this action in this state observed in the data, and ξ_{it} is the idiosyncratic shock to the probabilities of choosing this stage. With the estimates of $\beta E_{O',\varepsilon'|O_{it},a_{it}} V(O',\varepsilon')$ from the first stage, we can compute the left-hand side of Equation 17 directly from the data, and the right-hand side of this equation is linear in the parameters defined in Equations 3-6. Observations are weighted by the number of realized observations of each state. Confidence intervals are computed using Bayesian bootstrap [Rubin, 1981], with clustering (draws of observation weights) done at the stage level.

From Equations 3-6, there are 179 utility parameters of interest – baseline utility from playing stages (174 parameters), tastes to open 1 or 11 loot boxes (2), the state dependence parameter (1), the player's preference not to be in the losing state (1), and the player's disutility from spending on loot boxes (1).

7 Results

We estimate the structural model for non-whales and whales separately. We visualize the estimates of consumer tastes for the stage play in lieu of a table with extra 174 parameters.

The utility from playing (Figure 5) increases over the first several stages – especially for non-whales – and fluctuates subsequently. This is aligned with the game design: the first 10-15

---

[20]Our parameter estimates are robust to alternative approximations of $CCP(a_{it} = 0|O_{it})$, such as additionally including rarity fixed effects, removing stage-by-loss interactions, and allowing for different slopes for when a player does and does not have enough currency to open one loot box for a regular price of 5 coins. The probability of choosing to leave the game does not depend on loot box prices; this is because we do not observe the price of a loot box if the player has chosen the terminal action. We have experimented with the parametric assumptions in the CCP_0 estimation; the functional form is flexible enough to account for non-linearities in the CCPs.

[21]Preference estimates are robust to using alternative values of 0.9 and 0.9999.

Figure 5: Estimates of Preference for Stage Play, $\alpha_{G,s}$

(a) Non-whales
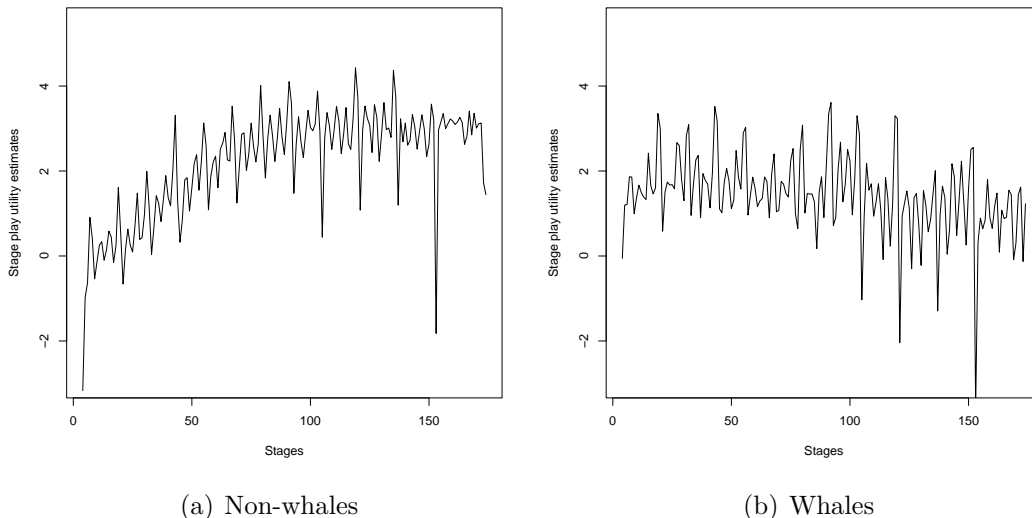
(b) Whales

stages are relatively simple to complete, whereas the latter stages are more difficult. For non-whales, stage utility is increasing in the first part of the game, indicating that they put more weight on making progress in the game and reaching latter stages. Further, both non-whales and whales derive more utility from playing harder stages. The cyclical game design, i.e. a boss stage every 4 stages (Section 3.3), is also consistent with the fluctuations in stage utilities.

Table 4: Estimates of Preference for Loot Boxes and Winning in the Game

	Non-Whales		Whales	
	Estimate	S.e.	Estimate	S.e.
One loot box ($\alpha_{L,1}$)	-1.7180	(0.4514)	0.1125	(0.1844)
State dependence (η)	1.1262	(0.2751)	2.0187	(0.2671)
Eleven-pack loot box ($\alpha_{L,11} - \alpha_{L,1}$)	0.0312	(0.4311)	-0.4776	(0.1913)
Payment (γ)	-0.1954	(0.0098)	-0.1545	(0.0039)
Lose the game (-β)	-0.4293	(0.1136)	-1.4488	(0.0846)

Standard errors are computed using Bayesian bootstrap, with clustering (draws of observation weights) done at the stage level.

Table 4 presents the rest of the estimated utility parameters. Both types prefer playing the game to opening loot boxes – the average stage fixed effect is around 1.5-2 utils for both types of players, while the average preference for opening one or loot boxes is -1.72 for non-whales and 0.11 for whales. However, both non-whales and whales exhibit state dependence in loot box consumption, with the magnitude of the state dependence more pronounced for whales than non-whales (2.02 versus 1.13 extra utils if another loot box has just been opened).

We note though that the state dependence coefficient captures both the potential structural and spurious state dependence in our context and should be interpreted as a reduced-form parameter. Players do not get much of a direct value from an eleven-pack of loot boxes over one loot box. Demand for loot boxes is negatively sloping for both non-whales and whales, with a slightly steeper slope for non-whales (-0.2 versus -0.15). Finally, both non-whales and whales dislike losing in the game and getting stuck on one level. The magnitude of the effect is stronger for whales.

Preference Heterogeneity. In our main specification, we allow for heterogeneity along players' monetary importance for the gaming company – whales, the 1.5% of players who are responsible for 90% of revenues, and non-whales, the rest of the players. In Appendix A.4, we extend our analysis in two ways to test whether parameter estimates vary when we allow for more heterogeneity, and to show that the division of players into whales and non-whales constitute the main dimension of heterogeneity. First, we split the sample of players along a different dimension – how many events the players have participated in throughout the game. The taste estimates are consistent across the two groups – they share the same level of state dependence preferences, preferences for eleven-pack loot box, payment, and past loss of the played stage. The only significant difference in the preferences of low- and high-engagement players of event games is their preference for loot boxes over the gameplay – high-engagement stage players also have a higher preference for opening loot boxes – which is expected since their players are more invested in the game.

Second, we separate out whales and non-whales into additional clusters based on the player-level propensities to play main-stage games over opening loot boxes. Table 10 in Appendix A.4 presents the estimates. As expected, in both groups players with a high preference for gameplay have lower estimates of their preferences for loot boxes. The state dependence tastes of players are consistent across high- and low-preference for gameplay clusters for whales. Non-whales are more price elastic compared to whales in both high- and low-preference for gameplay groups. The high-preference for gameplay group (cluster 2) does not care as much about losing the game, but overall losses are still more costly for whales compared to non-whales. We confirm that our key takeaways are robust to allowing for heterogeneity within the groups of whales and non-whales when we decompose the mechanisms behind the loot box value for players.

8 Loot Box Value Decomposition and Counterfactuals

We use our estimates of players' tastes to compare the relative importance of the functional and direct value of loot boxes for whales and non-whales, simulate the effects of loot box regulation, and evaluate alternative game designs.

8.1 Loot Box Value Decomposition

We start by quantifying the relative weight of the two roles of loot boxes in players' tastes. For this, we compare the value of the expected present discount utility flow under two scenarios. First, under the current game design, loot boxes increase current utility by providing users with an option to open a loot box today. This captures the direct utility of loot boxes. Second, they increase the future flow of utility by potentially affecting the inventory of divers. This increases the future likelihood of receiving win utility, as well as increases the likelihood that the user will stay in the game rather than leaving. This captures the functional value of loot boxes, which is realized in future periods.

Given the logit form of our choice model, the baseline future expected utility (up to a constant) takes the form of the "log-sum" of utilities attributable to each of the four actions that a player can take,

$$\tilde{V}_{\text{baseline}}(O) = \ln \left(\underbrace{\exp\left(u(a=1) + \beta E_{R',\hat{O}',\varepsilon'|O,a=1} V(O',\epsilon')\right)}_{\text{Playing stage game}} \right. \tag{18}$$

$$\left. + \underbrace{\sum_{\tilde{a}\in\{2,3\}} \exp\left(u(a=\tilde{a}) + \beta E_{R',\hat{O}',\varepsilon'|O,a=\tilde{a}} V(O',\epsilon')\right)}_{\text{Opening 1 or 11 loot boxes}} + \underbrace{\exp\left(u(a=0)\right)}_{\text{Exit game}} \right) \tag{19}$$

where $u(a)$ correspond to the action-specific current period utilities outlined in Section 4. For the purposes of the decomposition exercise, we abstract away from the role of prices and currency by assuming loot boxes are free for the immediate period. $V(O,\varepsilon)$ is the true value function that is consistent with the full empirical model as defined in Equation 9, and the expectation is taken over the distribution of the next period states $O' = \{R', \hat{O}'\}$ conditional on the current state O and action a, where we write out R' apart from the rest of the states $\hat{O}' = \{c, s, q, d, p^1, p^{11}$ since we adjust its transition probability in the utility decomposition.

We contrast this to the future expected utility of a scenario in which users do not have the option to open a loot box in the current period,

$$\tilde{V}_{\text{n.l.}}(O) = \ln \left(\underbrace{\exp\left(u(a_0=1) + \beta E_{R',\hat{O}',\varepsilon'|O,a_0=1} V(O',\epsilon')\right)}_{\text{Playing stage game}} + \underbrace{\exp\left(u(a=0)\right)}_{\text{Exit game}} \right) \tag{20}$$

where the expected utility associated with having the option to open loot boxes has been removed. The difference between these two values $\tilde{V}_{\text{baseline}} - \tilde{V}_{\text{n.l.}}$ represents the value that users get from an option to open one or eleven loot boxes in the current time period.

We now decompose the value of loot boxes into direct and functional mechanisms by

shutting down only the functional value of loot boxes. For this, we shut down the expected change in inventory quality ($R' = R$) for the immediate period, which in turn affects future utility flow by changing the state transitions, and the players' expectation of future expected rarity levels,

$$\tilde{V}_{\text{n.f.}}(O) = \ln \left(\underbrace{\exp(\cdots)}_{\text{Playing stage game}} + \underbrace{\sum_{\tilde{a}\in\{2,3\}} \exp\left(u(a=\tilde{a}) + \beta E_{\boxed{R},\hat{O}',\varepsilon'|O,a=\tilde{a}} V(O',\epsilon')\right)}_{\text{Opening 1 or 11 loot boxes}} + \underbrace{\exp(\cdots)}_{\text{Exit game}} \right) \quad (21)$$

The role of the direct utility of loot boxes in driving loot box openings can then be obtained as a ratio of the expected utility not explained by the functional component to the overall value of loot boxes,

$$\frac{\tilde{V}_{\text{n.f.}} - \tilde{V}_{\text{n.l.}}}{\tilde{V}_{\text{baseline}} - \tilde{V}_{\text{n.l.}}}. \quad (22)$$

We weigh the expected values by the empirical frequency in which we observe players at the given states to present our results.

Table 5 shows the resulting decomposition for whales and non-whales. The first two columns ("overall") show the decomposition weighted by the users' empirical distribution of states. Whales and non-whales have fundamentally different tastes for loot boxes, confirming our conclusions from the reduced-form estimates in Table 3. For whales, only 3% of loot boxes' value comes from the functional mechanism, while for non-whales the functional mechanism accounts for almost 90% of loot boxes' value. Part of this difference in the sources of loot boxes' utility is explained by whales being more likely to play in the latter stages of the game, where the functional value is lower because the player already has a better inventory and is closer to finishing the game. For instance, at one of the first stages, stage 5, whales get 22% of loot box value from the functional mechanism, but only 4% at stage 50. However, even conditional on stages, the functional value of loot boxes for whales is 2-4 times lower than for non-whales.

Table 5: Decomposition of the Loot Box Value

	Overall		Stage 5		Stage 50		Final stage	
	Non-Whales	Whales	Non-Whales	Whales	Non-Whales	Whales	Non-Whales	Whales
Functional value	89.51	3.04	93.73	21.92	6.72	3.91	0.24	0.08
	(2.67)	(0.72)	(1.76)	(3.58)	(4.57)	(0.95)	(0.13)	(0)
Direct value	10.49	96.96	6.27	78.08	93.28	96.09	99.76	99.92
	(2.67)	(0.72)	(1.76)	(3.58)	(4.57)	(0.95)	(0.13)	(0)
State Dependence	2.50	33.38	2.04	37.17	30.89	42.26	24.56	29.71
	(0.71)	(3.08)	(0.58)	(6.23)	(7.46)	(4.22)	(7.31)	(2.86)
Option to loot	7.98	63.58	4.23	40.90	62.40	53.84	75.20	70.21
	(2.43)	(2.52)	(1.53)	(3.56)	(9.8)	(3.46)	(7.4)	(2.86)

Standard errors are computed using Bayesian bootstrap, with clustering (draws of observation weights) done at the stage level.

We further decompose the direct utility players get from loot boxes into a part explained

by the state dependence in loot box consumption and a remainder part associated with the action of opening a loot box. For this, we extend the decomposition done in equation 22 and shut down the current period utility associated with the state dependence parameter. We find that state dependence accounts for a substantial share of the direct utility that players get from loot boxes, around 1/3 for whales and slightly less, around 23%, for non-whales.[22]

To confirm that our loot box value decompositions are not driven by the assumption of homogeneity of preferences within the groups of whales and non-whales, we redo the loot box value decomposition using the preference estimates with heterogeneity, presented in Table 10. Results are presented in Table 11 in Appendix A.5. We confirm that our main takeaways hold. The share of functional value in loot boxes is much smaller for whales (7.3%) than non-whales (66%), and state dependence accounts for 27%-42% of the direct value for players with a significant coefficient on the state dependence variable (clusters 2, 3, and 4).[23]

Based on a snapshot – by turning off various drivers of value attributable to loot boxes for one immediate period – we conclude that direct thrill utility is an important source of utility for loot box for whales. This decomposition is a "partial effect" in that it does not reflect the effect on overall gameplay if consumer preferences or the gaming environment vary. Changing the environment likely affects in-game behavior and consequently the likelihood of players arriving in each game state. To better understand what consequences loot boxes have for the overall gameplay experience of consumers — and to understand what it means for policymakers and game designers — we turn to our counterfactual simulations.

8.2   Counterfactuals and Product Design Implications

We study the managerial and policy implications of our estimates. Policymakers may be interested in regulating specific features of loot boxes and microtransactions more generally. Managers may seek to enhance engagement with the game by altering the difficulty of game stages or the variability of the loot box. These changes affect the composition of active users, and how users engage with loot boxes and the game — ultimately affecting firm revenues as well as consumer surplus. These are questions of product design, and require an assessment of how these policy decisions affect the overall gameplay. Therefore, through a series of simulations, we alter the difficulty of the game and the design of loot boxes, as well as evaluate various policy actions proposed by consumer protection groups and regulators.

Setup.   For any given counterfactual scenario and player type (e.g. non-whales), we draw simulated players from the empirical distribution of consumer types at the beginning of the game. For each scenario, we create a corresponding transition matrix and utility function. For instance, to simulate the outcomes in a more difficult version of the game, we decrease the probability that a player wins a stage and transitions to the next stage compared to the baseline transition matrix. Based on the transition matrix and utilities, we simulate the implied

---

[22]Once again, we highlight that the state dependence coefficient should be interpreted as a reduced-form parameter.

[23]For cluster 1, the state dependence estimate in Table 10 is negative and noisy.

expected future utility function by iterating over the Bellman equation as in Rust [1987].[24] We then compute the implied conditional choice probabilities for any given counterfactual scenario, which we use to simulate how the players would play the game and open loot boxes. We start the simulation from stage 5, as the first few stages of the games were "tutorials" in which players were taught how to play the game, and are characterized by high exit rates which may not be attributable to the design of the stage game and of the loot box.

Policies: Loot Box Ban and Spending Limits. We first evaluate sources of consumer utility under the current game design and under various counterfactual policy actions proposed by consumer protection groups and regulators.

Figure 6: Revenue and Consumer Surplus under Loot Box Bans and Spending Caps

| (a) Overall | (b) Non-Whales | (c) Whales |
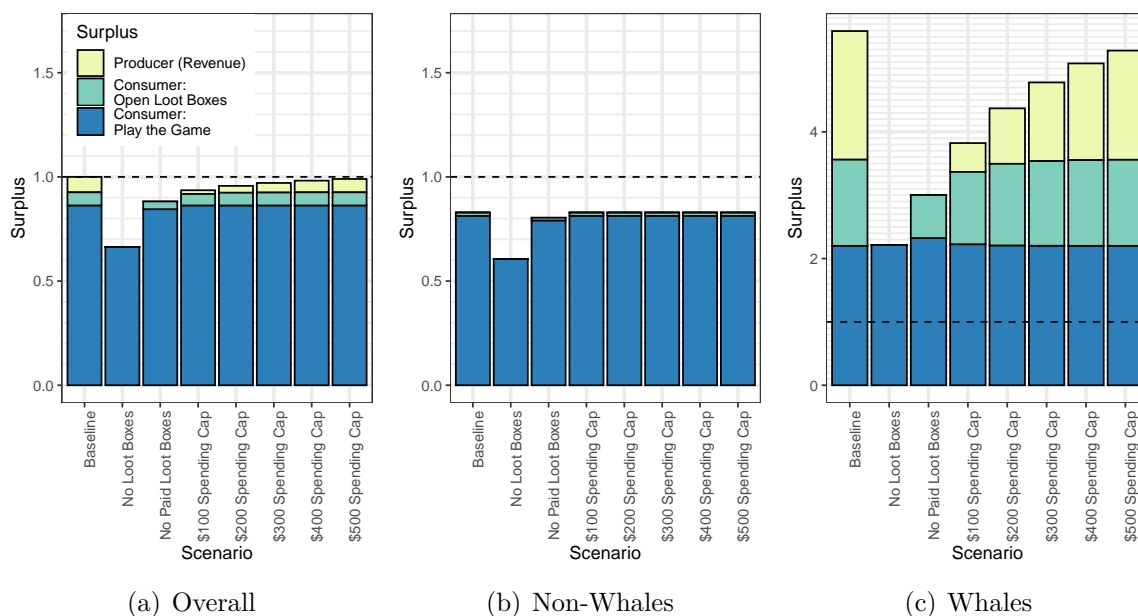
Estimates of producer and consumer surplus by scenario and types of players. The overall surplus per (average) player under the current scenario is normalized to one (dashed line). We use an ex-post measure of consumer surplus – based on players' realized path of actions – to decompose the surplus from playing the stage game and from opening loot boxes.

Figure 6 presents the resulting revenue and consumer surplus.[25] The first bar in all sub-figures corresponds to the current scenario, where the total surplus for an average player is

---

[24]The sequential nature of the model makes solving this problem easier because it can be broken out into pieces. Namely, we use the simulated expected utility function from stage $s + 1$ as input into finding the expected utility function for states in stage s.

[25]We note that we count full utility estimates as "consumer surplus." We acknowledge that while the utility from loot boxes includes normatively respectable preferences – like the functional value or the entertainment value from the resolution of uncertainty [e.g. Ely et al., 2015] – it also includes the utility arising from behavioral biases like self-control problems [Lockwood et al., 2021]. Thus, we caution the reader from interpreting these estimates as welfare measures.

normalized to one.[26] Producers get 7.4% of the total surplus in the form of revenue (abstracting away from fixed costs), with players getting 6.3% from opening loot boxes and the rest 86.3% from playing the game.[27] Splitting these average estimates for non-whales and whales, we see that non-whales (subfigure b in Figure 6) get the vast majority, 97.8%, of the total surplus from playing the game, with only 1.8% coming from opening loot boxes. In contrast, whales (subfigure c in Figure 6) get only 39.3% of the total surplus from playing the game, but another 24.3% from opening loot boxes. Overall, whales get around 90 times more utility from loot boxes than non-whales, while they get only 2.7 times more utility from playing the game itself. This difference emphasizes that whales enjoy loot boxes over the gameplay disproportionally more compared to non-whales.

In the second set of bars ("No Loot Boxes"), we evaluate the effect of a complete loot box ban in our video game [e.g. like the one proposed by Forbruker Rådet, 2022]. We simulate this scenario by removing the option to open loot boxes (actions 2 and 3). Compared to the baseline scenario, the total surplus based on the tastes of an average player drops by 33.7%. This includes a drop to zero of revenue and consumer surplus from loot boxes – by construction, since loot boxes are now banned – but also a 23.2% drop in the surplus players get from playing the game itself. The decrease in surplus associated with gameplay comes only from non-whales (their utility from gameplay drops by 25.4%), while whales are virtually unaffected. This result highlights a strong complementarity of the game and loot boxes that non-whale players experience and, once again, highlights the drastic difference in the nature of loot box tastes for whales and non-whales.

Next, in the third set of bars in Figure 6 ("No Paid Loot Boxes"), we evaluate a scenario where loot boxes are still in the game, but players cannot open any loot boxes by spending real money. We simulate this scenario by setting the price coefficient to negative infinity – so players can never open loot boxes if they do not have enough coins that were acquired through the organic gameplay. Compared to the baseline scenario, consumer surplus from playing the game drops by 2.1%, and consumer surplus from opening loot boxes drops by 41.3%. A drop in consumer surplus from playing the game is driven entirely by non-whales, while for whales, consumer surplus from playing the game slightly *increases* (by 5.5%) under the ban of paid loot boxes – because such ban forces these players to replay hard stages more and get more play utility. In contrast, whales are responsible for the drop in consumer surplus from opening loot boxes – without an option to open paid loot boxes, they lose 50% of loot box openings surplus. By construction, the firm's revenue (producer surplus) is zero.

In the last five bars across subfigures of Figure 6, we evaluate the effect of imposing spending caps of different levels [similar to the proposals of Drummond et al., 2019, Close and Lloyd, 2021, Leahy, 2022]. We vary the spending caps from $100 to $500 per player over their

---

[26]Figure 21 in Appendix A.7 presents the number of plays and loot box openings for an average player, with the total number of actions similarly normalized to one.

[27]We use an ex-post measure of consumer surplus – based on players' realized path of actions – to decompose the surplus from playing the stage game and from opening loot boxes. Specifically, we track and take the (discounted) sum of utilities and realization of error draws to obtain the consumer surplus from the game.

lifetime. We simulate this scenario in a stylized way – players stay playing the game without any constraints, but once they hit the spending cap they cannot open more loot boxes using real money (but can keep using coins).[28] In these scenarios, the actions and surplus of non-whale players are not affected, since none of these players cross the lowest threshold of spending $100 in our simulations. The consumer surplus of whales from playing the game is also very close to the baseline level. However, spending caps introduce large differences in how much surplus whales get from loot boxes, and how much revenue the firm collects. Under a $100 spending cap, whales get 84% of their consumer surplus from loot box openings, while the firm gets only 22.4% of the baseline revenue (24.3% of total baseline revenue). As the cap increases to $300, whales recover 98.3% of their baseline surplus from opening loot boxes, and the company recovers 61.1% of the baseline revenue from whales (62.2% of total baseline revenue). Finally, under a $500 spending cap, whales recover almost all, 99.9%, of their baseline surplus from opening loot boxes, and the company recovers 85% of the baseline revenue from whales (86.5% of total baseline revenue). This implies that the firm extracts almost all incremental surplus from loot boxes that are opened after the first $500 of expenditures of the player.

Overall, counterfactual simulations show that a blanket ban on loot boxes is likely too stringent as a policy action. Apart from removing all surplus that players get from opening loot boxes, it significantly reduces consumer surplus from playing the game itself, due to the complementarity between loot boxes and gameplay that non-whales – the vast majority of players – experience. Banning only paid loot boxes could be a better middle-ground solution, recovering most of the consumer surplus from gameplay and around 50% of the consumer surplus from opening loot boxes. However, conditional on the development of the game, no producer surplus is generated. This suggests that the company will not be able to recover the fixed costs of game production. To address this, policymakers can implement spending caps. Simulations show that spending caps allow for retaining most of the surplus that players get from loot box openings (compared to the baseline) and still generate revenue for the game producer. A high spending cap (e.g. $300 or more) has close to no effect on the baseline consumer surplus but restricts the firm from profiting off of high-rollers, who give almost all their surplus to the firm in exchange for playing the loot box lottery.

Game design: changing the difficulty of the game. We now evaluate a scenario where the gaming company changes game difficulty. Game difficulty is one of the primary levers of game design that the company has since the puzzle-solve nature of the game is the core mechanic in the game and its main stated attraction. The design of the stage game is likely to interact with the players' preferences for loot box opening, which further showcases the industry's argument that the role of loot boxes can only be assessed within the overall context of the game.

More specifically, we simulate how users would behave if the win probabilities of game

---

[28]Thus, the players are myopic regarding their "budget restriction" – they do not anticipate that they will hit a spending limit. This model reflects a scenario where players are not well informed about the spending limit and approximates a more complex model where players track how far they are from this budget constraint as an additional state variable.

stages were modified from the original game. We implement this by uniformly increasing or decreasing win probabilities of game stages by 10 percent increments, to make the game easier or more difficult. We place a lower bound of win probability at 10% so that the game is not prohibitively difficult such that it prevents players from making any progress.

Table 6 presents the resulting revenues and user engagement metrics. The first three columns show the expected revenue generated from opening loot boxes for real money – overall and split by non-whales and whales. We normalize the current revenue, under the baseline difficulty of the game, to unity. As the game becomes more difficult, the expected revenue the company can harness from the existing user base increases dramatically – from 100% to 223% if the win probability is decreased by 10 percent, and all the way to 2349% if it is halved. This increase in revenue is driven entirely by the expenditures of whales who dramatically increase the number of rare loot boxes they open to still make progress in the game. In contrast, revenues from non-whales fall by around 30% when stages' win probabilities are decreased by 50 percent points – since fewer non-whales engage with the game and keep playing.

Table 6: Simulations under varying game stage win probabilities

	Revenue			# of Stage Games Played			Share of Consumers At Stage 20		
	Overall	Non-whales	Whales	Overall	Non-whales	Whales	Overall	Non-whales	Whales
Harder									
Win prob -50%	23.49	0.05	659.77	0.28	0.18	2.89	0.21	0.15	1.93
Win prob -40%	15.92	0.05	446.70	0.37	0.29	2.52	0.33	0.27	1.95
Win prob -30%	7.24	0.06	202.17	0.49	0.42	2.3	0.47	0.41	1.97
Win prob -20%	3.33	0.07	91.86	0.64	0.58	2.14	0.63	0.58	1.98
Win prob -10%	2.23	0.07	61.02	0.81	0.77	2.04	0.82	0.77	1.98
Current win prob	**1.00**	0.07	26.21	**1.00**	0.96	1.96	**1.00**	0.96	1.98
Win prob +10%	0.73	0.07	18.26	1.05	1.02	1.91	1.05	1.02	1.98
Win prob +20%	0.62	0.07	15.52	1.06	1.03	1.88	1.07	1.04	1.98
Win prob +30%	0.58	0.06	14.58	1.06	1.03	1.87	1.08	1.05	1.98
Win prob +40%	0.55	0.06	13.85	1.06	1.03	1.86	1.1	1.06	1.98
Win prob +50%	0.54	0.06	13.58	1.06	1.03	1.85	1.1	1.07	1.98
Easier									

We simulate how players engage with the game when we uniformly increase or decrease win probabilities of game stages by 10 percent increments. Revenue is the average sum of payments made from opening rare loot boxes over the lifetime of a player. The number of times the player played the stage game includes the attempts the player lost. The weighted average value of each metric across non-whales and whales under the current design of the game is normalized to unity (presented in bold).

Given such a large revenue upside from making the game more difficult, a natural question arises: Why does not the company implement this game design change? This design would target the preferences of whales and extract more value from them. However, by targeting the preferences of whales, the company will be providing an inferior product to all non-whales – who constitute the vast majority of the company's players – and will decrease their adoption and engagement with the game. While non-whales bring only a tiny share of the game's revenue, they generate direct network effects that increase the adoption of the game by players, including whales – e.g. because more popular games are higher in games' rankings in the App Store and Google Play, generate more word-of-mouth, or create other positive peer effects. This reflects

the value of "free" customers and the importance for firms to target them and balance out the revenue and growth objectives [e.g. Gupta et al., 2009, Lee et al., 2017].

To examine the effect of changes in the difficulty of the game on user engagement, the last six columns in Table 6 report the number of times an average user plays the stage game and the share of consumers reaching stage 20 – a benchmark for strong user engagement with the game, having completed the initial tutorial stages and becoming familiar with the game. Metrics for the baseline difficulty case are once again normalized to unity. The users' engagement with the game drops prominently as the game becomes more difficult. Stage plays go down to 81% if the game's complexity is increased by 10 percent, and to 28% if increased by 50 percent. In contrast, making the game easier increases the number of stage plays, although by a much smaller magnitude – making the game 50 percent easier increases the number of play occasions only by 6%. The metric of the shares of consumers at stage 20 – the last three columns – follows a very similar pattern, with the number of play occasions dropping by 79% if the game becomes 50 percent more difficult and increasing by 10% if the game becomes 50 percent easier.

Changes in user engagement are driven primarily by non-whales' responses. If the win probability of the game decreases by 50 percent, non-whales decrease their engagement by 81% in terms of the number of play occasions and by 84% in terms of the share of players reaching stage 20. In contrast, the effect for whales is much smaller or even reversed. As the game becomes more complex, the number of times whales play the game on average *increases*, by 4% if win probabilities increase by 10 percent and by 47% if by 50 percent. Such a positive relationship between the game's difficulty and engagement for whales shows the complementarity between loot box openings and the gameplay. As the game is harder to complete, whales open more loot boxes and get better items. Concurrently they open more loot boxes due to positive state dependence and ultimately play more. The last column shows that while whales make more play attempts, 2.5% fewer whales reach stage 20 if the win probability is halved.

Looking at Table 6 overall, we conclude that the current game design balances out the revenue and engagement objectives of the firm – a 10 percentage point decrease in win probability leads to 123% higher revenues but also to 19% less stage play. A further increase in win probability by 10 percentage points decreases revenue by 27% but increases engagement only by 5%. Whales and non-whales play different roles in this trade-off – whales bring most of the revenue change, while non-whales are responsible for most of the engagement change.

Loot boxes design: changing the variance of outcomes. Finally, we simulate outcomes under alternative loot box designs. One of the main controversial features of loot boxes is the *randomness* of loot box outcomes. While we cannot shut down the randomness of loot box outcomes entirely – unbundling loot boxes and offering a menu of items for purchase would require estimates of the direct values of items – we simulate counterfactual behavior of users under higher and lower variance of loot box outcomes.

The effect of altering the variability of loot box outcomes is informative about the role of "chance" in loot boxes for the users' gameplay. One observation, which we discuss concretely below, is that the expected return from opening loot boxes conditional on a player's current inventory levels is increasing in variance. Players only benefit if they receive divers that are of

a higher rarity than what they have in their inventory. In other words, the role of "chance" in loot boxes is inherently associated with functional value as well.

To implement these simulations, we obtain new transition matrices corresponding to alternative loot box designs. In these alternative scenarios, we keep constant the expected rarity outcomes but either increase or decrease the standard deviation of rarity outcomes by 30 and 50 percentage points. Figures 19-20 in Appendix A.6 visualize the resulting state transition matrices.

Table 12 in Appendix A.7 presents the simulated revenues and player actions under different levels of the variability of loot box outcomes. The first three columns correspond to players under the current game and loot box preferences. The first thing that jumps out is that compared to changes in the game design presented in Table 6, changes in probabilities of loot box outcomes have a much lower impact on revenues and players' behavior. For instance, the overall revenue decreases at most by 6% and increases at most by 30%, closely tracking the number of loot box openings. Interestingly, game revenue is *higher* under the smaller variability of loot box outcomes – e.g. if the standard deviation of loot box outcomes is 50% lower, the expected revenue is 30% higher and the number of rare loot boxes opened is 27% higher. This is despite a lower expected benefit of loot boxes under lower standard deviations of outcomes. The reason for this counterintuitive response is due to whales: they open loot boxes not because of their functional value but due to their direct taste. Without the functional value, it is harder to play the stage game, so whales open more loot boxes. In contrast, non-whales do not change their loot box opening behavior much, as they open very few loot boxes to begin with.

Columns 4-6 further highlight the importance of the state-dependent behavior in loot box consumption for whales. In these columns, we re-run the simulations setting the state dependence parameter (η) of players to zero. Results are drastically different compared to the baseline case. First, revenues drop significantly, by 91%, and no longer prominently change with loot box variability. This decrease in revenue is primarily driven by the decrease in the number of loot boxes whales open, which drops by more than 90%. The effect for non-whales is similar in direction but much smaller in magnitude – they start opening loot boxes 23% less – highlighting the higher importance of the direct value of loot boxes for whales. Interestingly, the effect of shutting down the state dependence parameter has different effects on non-whales and whales in terms of the volume of gameplay. Non-whales start to play the game a bit (3.1%) less, driven by the complementarity of the functional value of loot boxes, and gameplay. In contrast, whales play the game a bit more (2%), since now opening loot boxes is much less attractive to them and they did not get much functional value from them in the first place.

Overall, results suggest that the company will be better off if it reduces the variability of loot box outcomes – both revenue and players' engagement will be slightly higher. This suggests that incentives of the company and regulators – who often argue against the randomness of loot box outcomes – are to some degree aligned. However, it is difficult to extrapolate our results to the case of directly selling deterministic products, as we cannot account for changes in direct value that are associated with changes in randomness.

9 Conclusion

In this paper, we provide evidence of fundamentally different tastes for loot boxes, lotteries that are built into video games, of regular and high-spending players ("whales"). We separate out two sets of tastes for loot boxes, the functional value of receiving virtual items that complement gameplay and the thrill utility of opening loot boxes. In the context of a free-to-play puzzle game, we show that regular players open loot boxes primarily for their complementarity with the game, a functional value of items that players receive to make in-game progress. In contrast, whales open loot boxes primarily due to their direct value.

This drastic difference in consumer tastes gives credibility to concerns of the consumer protection groups and regulators who worry that the mechanics of paid loot boxes closely resemble gambling and can attract spenders who get the direct thrill from paying for uncertain rewards. We use the taste estimates to evaluate various proposed policy actions and argue that spending caps are more appropriate than a blanket ban on loot boxes. We also showcase how our taste estimates can be used to evaluate the effect of counterfactual game and loot box designs on the firm's revenue and players' engagement with the game.

Our estimates and analysis have several important limitations. First, while we can separate out the functional value from the direct taste for loot boxes, we do not have the right variation in the data to decompose the direct tastes into normatively respectable preferences – like the entertainment value from the resolution of uncertainty [e.g. Ely et al., 2015] – and behavioral biases like self-control problems and addiction [Lockwood et al., 2021]. Future work is required to separate out these important mechanisms, e.g. using the experimental design like in Allcott et al. [2022]. Second, our dataset lacks player characteristics; separating out taste estimates by income groups, age, and exposure to problem gambling is an important step in evaluating the degree of exposure of minors and individuals with gambling problems to loot boxes. Finally, we study a single-player game with close to no social interactions, ruling out some forms of the value of loot boxes that can be important in other contexts, like the cosmetic value of the items.

References

Dan Acland and Vinci Chow. Self-control and demand for commitment in online game playing: evidence from a field experiment. *Journal of the Economic Science Association*, 4(1):46–62, 2018.

Paulo Albuquerque and Yulia Nevskaya. The impact of new content and user community membership on usage of online games. 9(1):1–24, 2022. ISSN 2196-291X, 2196-2928. doi: 10.1007/s40547-022-00127-2. URL https://link.springer.com/10.1007/s40547-022-00127-2.

Hunt Allcott, Matthew Gentzkow, and Lena Song. Digital addiction. *American Economic Review*, 112(7):2424–63, 2022.

Anonymous Player. A whale of a tale. https://www.reddit.com/r/FFBraveExvius/comments/7jmezv/a_whale_of_a_tale/ [Acccessed: January 31, 2023], 2017.

Peter Arcidiacono and Robert A Miller. Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica*, 79(6):1823–1867, 2011.

Guy Aridor. Drivers of digital attention: Evidence from a social media experiment. *Available at SSRN 4069567*, 2022.

Belgium Gaming Commission. Report: Loot boxes. `https://www.gamingcommission.be/nl/publicaties/rapporten` [Acccessed: January 10, 2023], 2018.

Steven T Berry. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262, 1994.

Stéphane Bonhomme, Thibaut Lamadon, and Elena Manresa. Discretizing unobserved heterogeneity. *Econometrica*, 90(2):625–643, 2022.

Julia Cagé. Media competition, information provision and political participation: Evidence from french local newspapers and elections, 1944–2014. *Journal of Public Economics*, 185: 104077, 2020.

casino.org. Gambling ages around the world. `https://www.casino.org/features/gambling-ages-worldwide/` [Acccessed: January 10, 2023], 2020.

Ningyuan Chen, Adam N Elmachtoub, Michael L Hamilton, and Xiao Lei. Loot box pricing and design. *Management Science*, 2020.

James Close and Joanne Lloyd. Lifting the lid on loot-boxes: Chance-based purchases in video games and the convergence of gaming and gambling. *Gamble Aware*, 2021.

James Close, Stuart Gordon Spicer, Laura Louise Nicklin, Maria Uther, Joanne Lloyd, and Helen Lloyd. Secondary analysis of loot box data: Are high-spending "whales" wealthy gamers or problem gamblers? *Addictive Behaviors*, 117:106851, 2021.

Gregory S Crawford and Ali Yurukoglu. The welfare effects of bundling in multichannel television markets. *American Economic Review*, 102(2):643–85, 2012.

Stefano DellaVigna and Ulrike Malmendier. Paying not to go to the gym. *american economic Review*, 96(3):694–719, 2006.

Aaron Drummond and James D Sauer. Video game loot boxes are psychologically akin to gambling. *Nature human behaviour*, 2(8):530–532, 2018.

Aaron Drummond, James D Sauer, and Lauren C Hall. Loot box limit-setting: A potential policy to protect video game users with gambling problems? 2019.

Jean-Pierre Dubé, Günter J Hitsch, and Peter E Rossi. State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics*, 41(3):417–445, 2010.

Dutch Council of State. Penalty imposed unjustly: "loot boxes" in computer game fifa22 are not a game of chance. `https://www.raadvanstate.nl/actueel/nieuws/@130206/dwangsom-onterecht-opgelegd-loot-boxes/` [In Dutch; Acccessed: January 10, 2023], 2022.

Jeffrey Ely, Alexander Frankel, and Emir Kamenica. Suspense and surprise. *Journal of Political Economy*, 123(1):215–260, 2015.

ESA. Esa leadership desk: Unpacking the loot box conversation. `https://www.theesa.com/news/esa-leadership-desk-unpacking-the-loot-box-conversation/` [Acccessed: February 3, 2023], 2019.

ESA. 2022 essential facts about the video game industry. `https://www.theesa.com/wp-content/uploads/2022/06/2022-Essential-Facts-About-the-Video-Game-Industry.pdf` [Acccessed: February 3, 2023], 2022.

extremetech.com. After 15 years, retail video game prices may be rising. `https://www.extremetech.com/gaming/312375-after-15-years-retail-video-game-prices-may-be-rising` [Acccessed: January 10, 2023], 2020.

Ying Fan. Ownership consolidation and product characteristics: A study of the us daily newspaper market. *American Economic Review*, 103(5):1598–1628, 2013.

Federal Trade Commission. Loot boxes: What's in play? `https://www.ftc.gov/business-guidance/blog/2020/08/loot-boxes-whats-play` [Acccessed: January 10, 2023], 2020.

Forbruker Rådet. Report on loot boxes: Insert coin. `https://www.forbrukerradet.no/`

`report-on-loot-boxes-insert-coin/` [Acccessed: January 10, 2023], 2022.

gamedeveloper.com. Online games will be required to disclose random loot box odds in china. `https://www.gamedeveloper.com/disciplines/online-games-will-be-required-to-disclose-random-loot-box-odds-in-china` [Acccessed: January 10, 2023], 2016.

gamedeveloper.com. New zealand says lootboxes 'do not meet the legal definition for gambling'. `https://www.gamedeveloper.com/business/new-zealand-says-lootboxes-do-not-meet-the-legal-definition-for-gambling-` [Acccessed: January 10, 2023], 2017.

gamedeveloper.com. The cost of games. `https://www.gamedeveloper.com/business/the-cost-of-games` [Acccessed: January 10, 2023], 2018.

gameinformer.com. Esa backs up loot boxes as ftc begins investigations. `https://www.gameinformer.com/2018/11/28/esa-backs-up-loot-boxes-as-ftc-begins-investigations` [Acccessed: January 10, 2023], 2018.

Gameranx. Gaming is five times bigger than movies now. `https://gameranx.com/updates/id/416500/article/gaming-is-five-times-bigger-than-movies-now/` [Acccessed: May 17, 2023], 2022.

gov.uk. Government response to the call for evidence on loot boxes in video games. `https://www.gov.uk/government/consultations/loot-boxes-in-video-games-call-for-evidence/outcome/government-response-to-the-call-for-evidence-on-loot-boxes-in-video-games` [Acccessed: January 10, 2023], 2022.

Mark D Griffiths. Is the buying of loot boxes in video games a form of gambling or gaming? *Gaming Law Review*, 22(1):52–54, 2018.

Sunil Gupta, Carl F Mela, and Jose M Vidal-Sanz. The value of a" free" customer. 2009.

Ruru Hoong. Self control and smartphone use: An experimental study of soft commitment devices. *European Economic Review*, 140:103924, 2021.

V Joseph Hotz and Robert A Miller. Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, 60(3):497–529, 1993.

Yan Huang, Stefanus Jasin, and Puneet Manchanda. "level up": Leveraging skill and engagement to maximize player game-play in online video games. 30(3):927–947, 2019. ISSN 1047-7047, 1526-5536. doi: 10.1287/isre.2019.0839. URL `https://pubsonline.informs.org/doi/10.1287/isre.2019.0839`.

Investopedia. How the video game industry is changing. `https://www.investopedia.com/articles/investing/053115/how-video-game-industry-changing.asp` [Acccessed: January 10, 2023], 2021.

Masakazu Ishihara and Andrew T. Ching. Dynamic demand for new and used durable goods without physical depreciation: The case of japanese video games. 38(3):392–416, 2019. ISSN 0732-2399, 1526-548X. doi: 10.1287/mksc.2018.1142. URL `https://pubsonline.informs.org/doi/10.1287/mksc.2018.1142`.

jdsupra.com. Video gaming: Is my loot box legal? `https://www.jdsupra.com/legalnews/video-gaming-is-my-loot-box-legal-57172/` [Acccessed: January 10, 2023], 2019.

Przemysław Jeziorski. Effects of mergers in two-sided markets: The us radio industry. *American Economic Journal: Microeconomics*, 6(4):35–73, 2014.

Wooyong Jo, Sarang Sunder, Jeonghye Choi, and Minakshi Trivedi. Protecting consumers from themselves: Assessing consequences of usage restriction laws on online game usage and spending. *Marketing Science*, 39(1):117–133, 2020.

Bruno Jullien and Bernard Salanié. Estimating preferences under risk: The case of racetrack bettors. *Journal of Political Economy*, 108(3):503–530, 2000.

Juniper Research. Video game loot boxes to generate over $20 billion in revenue by 2025,

but tightening legislation will slow growth. `https://www.juniperresearch.com/press/video-game-loot-boxes-to-generate-over-$20-billion` [Acccessed: May 18, 2023], 2021.

Myrto Kalouptsidi, Paul T Scott, and Eduardo Souza-Rodrigues. Linear iv regression estimators for structural dynamic discrete choice models. *Journal of Econometrics*, 222(1): 778–804, 2021.

Daniel L King and Paul H Delfabbro. Predatory monetization schemes in video games (eg 'loot boxes') and internet gaming disorder. 2018.

Søren Kristiansen and Majbritt Christine Severin. Loot box engagement and problem gambling among adolescent gamers: Findings from a national survey. *Addictive Behaviors*, 103:106254, 2020.

Ng Yew Kwang. Why do people buy lottery tickets? choices involving risk and the indivisibility of expenditure. *Journal of Political Economy*, 73(5):530–535, 1965.

Chanel J Larche, Katrina Chini, Christopher Lee, Mike J Dixon, and Myra Fernandes. Rare loot box rewards trigger larger arousal and reward responses, and greater urge to open more loot boxes. *Journal of gambling studies*, 37(1):141–163, 2021.

D Leahy. Rocking the boat: Loot boxes in online digital games, the regulatory challenge, and the eu's unfair commercial practices directive. *Journal of Consumer Policy*, 45(3):561–592, 2022.

Clarence Lee, Vineet Kumar, and Sunil Gupta. Designing freemium: Balancing growth and monetization strategies. Technical report, Working paper, 2017.

Xiao Liu, Matthew Shum, and Kosuke Uetake. Passive and active attention to baseball telecasts: Implications for content (re-) design. *NYU Stern School of Business Forthcoming*, 2020.

Benjamin Lockwood, Hunt Allcott, Dmitry Taubinsky, and Afras Y Sial. What drives demand for state-run lotteries? evidence and welfare implications. Technical report, National Bureau of Economic Research, 2021.

Gregory J Martin and Ali Yurukoglu. Bias in cable news: Persuasion and polarization. *American Economic Review*, 107(9):2565–99, 2017.

Sridhar Narayanan and Puneet Manchanda. An empirical analysis of individual level casino gambling behavior. *Quantitative Marketing and Economics*, 10(1):27–62, 2012.

Yulia Nevskaya and Paulo Albuquerque. How should firms manage excessive product use? a continuous-time demand model to test reward schedules, notifications, and time limits. 56(3):379–400, 2019. ISSN 0022-2437, 1547-7193. doi: 10.1177/0022243718821698. URL `http://journals.sagepub.com/doi/10.1177/0022243718821698`.

Newzoo. Newzoo global games market report 2021. `https://newzoo.com/insights/trend-reports/newzoo-global-games-market-report-2021-free-version` [Acccessed: February 3, 2023], 2021a.

Newzoo. Mobile puzzle games: Popularity, revenues and opportunities across east and west. `https://newzoo.com/resources/blog/puzzle-game-mobile-data-revenues-player-behavior-consumer-insights-us-japan-china-korea` [Acccessed: May 17, 2023], 2021b.

NHS. Country's top mental health nurse warns video games pushing young people into 'under the radar' gambling. `https://www.england.nhs.uk/2020/01/countrys-top-mental-health-nurse-warns-video-games-pushing-young-people-into-under-the-radar-g` [Acccessed: January 10, 2023], 2020.

Hee Mok Park and Puneet Manchanda. When harry bet with sally: An empirical analysis of multiple peer effects in casino gambling behavior. *Marketing Science*, 34(2):179–194, 2015.

Hee Mok Park and Joseph Pancras. Social and spatiotemporal impacts of casino jackpot events. *Marketing Science*, 41(3):575–592, 2022.

pcgamer.com. Behind the addictive psychology and se-

ductive art of loot boxes. `https://www.pcgamer.com/behind-the-addictive-psychology-and-seductive-art-of-loot-boxes/` [Acccessed: January 10, 2023], 2017.

pcgamer.com. Netherlands gaming authority cracks down on loot boxes in some games. `https://www.pcgamer.com/netherlands-gaming-authority-cracks-down-on-loot-boxes-in-some-games/` [Acccessed: January 10, 2023], 2018.

remarkablecoder.com. How many video games exist? `https://remarkablecoder.com/how-many-video-games-exist/` [Acccessed: January 10, 2023], 2019.

Sherwin Rosen. Manufactured inequality. *Journal of Labor Economics*, 15(2):189–196, 1997.

Donald B Rubin. The bayesian bootstrap. *The annals of statistics*, pages 130–134, 1981.

Julian Runge, Jonathan Levav, and Harikesh S. Nair. Price promotions and "freemium" app monetization. 20(2):101–139, 2022. ISSN 1570-7156, 1573-711X. doi: 10.1007/s11129-022-09248-3. URL `https://link.springer.com/10.1007/s11129-022-09248-3`.

John Rust. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society*, pages 999–1033, 1987.

Paul Scott. Dynamic discrete choice estimation of agricultural land use. 2014.

screenrant.com. Belgium & the netherlands classify loot boxes as illegal gambling. `https://screenrant.com/loot-boxes-illegal-belgium-netherlands/` [Acccessed: January 10, 2023], 2018.

Luxi Shen, Christopher K Hsee, and Joachim H Talloen. The fun and function of uncertainty: Uncertain incentives reinforce repetition decisions. *Journal of Consumer Research*, 46(1): 69–81, 2019.

Andrey Simonov, Jean-Pierre Dubé, Günter Hitsch, and Peter Rossi. State-dependent demand estimation with initial conditions correction. *Journal of Marketing Research*, 57(5):789–809, 2020.

Andrey Simonov, Raluca Ursu, and Carolina Zheng. Suspense and surprise in media product design: Evidence from twitch. tv. 2022.

Stuart Gordon Spicer, Laura Louise Nicklin, Maria Uther, Joanne Lloyd, Helen Lloyd, and James Close. Loot boxes, problem gambling and problem video gaming: A systematic review and meta-synthesis. *New Media & Society*, 24(4):1001–1022, 2022.

Statista. Global video game market value from 2020 to 2025. `https://www.statista.com/statistics/292056/video-game-market-value-worldwide/` [Acccessed: January 10, 2023], 2021.

Takeaki Sunada. Customer learning and revenue-maximizing trial design. Technical report, Technical report, Working Paper, 2018.

Andrew Sweeting. Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry. *Econometrica*, 81 (5):1763–1803, 2013.

Wayne Taylor and Anand V Bodapati. Incorporating experience quality data into crm models: The impact of gambler outcomes on casino return times. *SMU Cox school of business research paper*, (19-16), 2017.

techcrunch.com. Free to play games rule the entertainment world with $88 billion in revenue. `https://techcrunch.com/2019/01/18/free-to-play-games-rule-the-entertainment-world-with-88-billion-in-revenue/` [Acccessed: January 10, 2023], 2019.

Udonis. What are mobile game whales and how to catch them? `https://www.blog.udonis.co/mobile-marketing/mobile-games/mobile-games-whales` [Acccessed: June 14, 2024], 2024.

Leon Y Xiao. Regulating loot boxes as gambling? towards a combined legal and self-regulatory consumer protection approach. *Interactive Entertainment Law Review*, 4(1):27–47, 2021.

David Zendle, Rachel Meyer, and Harriet Over. Adolescents and loot boxes: Links with problem gambling and motivations for purchase. *Royal Society Open Science*, 6(6):190049, 2019.

David Zendle, Paul Cairns, Herbie Barnett, and Cade McCall. Paying for loot boxes is linked to problem gambling, regardless of specific features like cash-out and pay-to-win. *Computers in Human Behavior*, 102:181–191, 2020.

Yi Zhao, Sha Yang, Matthew Shum, and Shantanu Dutta. A dynamic model of player level-progression decisions in online gaming. 68(11):8062–8082, 2022. ISSN 0025-1909, 1526-5501. doi: 10.1287/mnsc.2021.4255. URL `https://pubsonline.informs.org/doi/10.1287/mnsc.2021.4255`.

FOR ONLINE PUBLICATION

A   Appendices

A.1   Additional Data Descriptives

Figure 7: A joint distribution of play and loot box openings across users
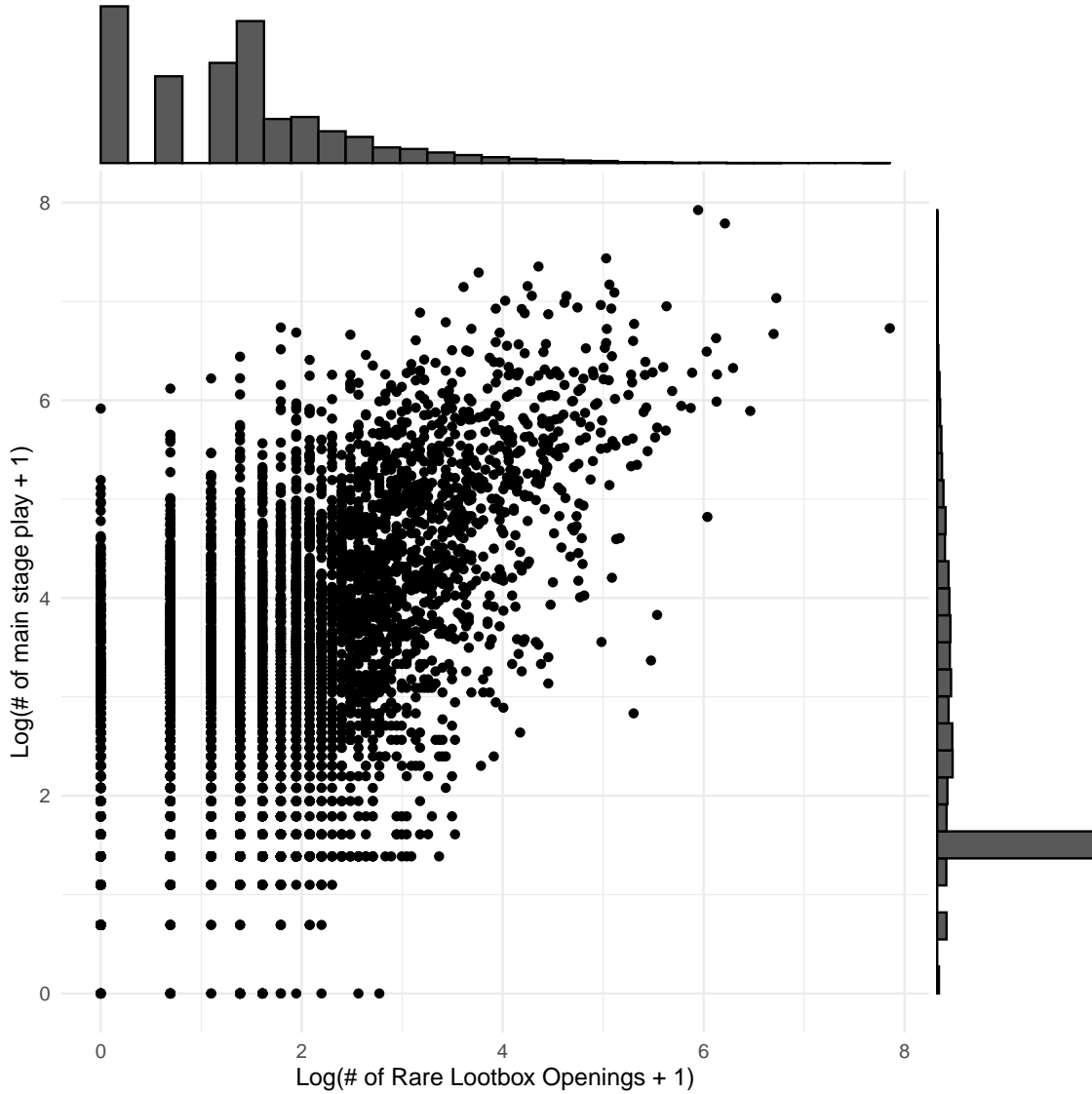
Based on a random sample of 10,000 users.

Figure 8: Distribution of Rarity of Divers.
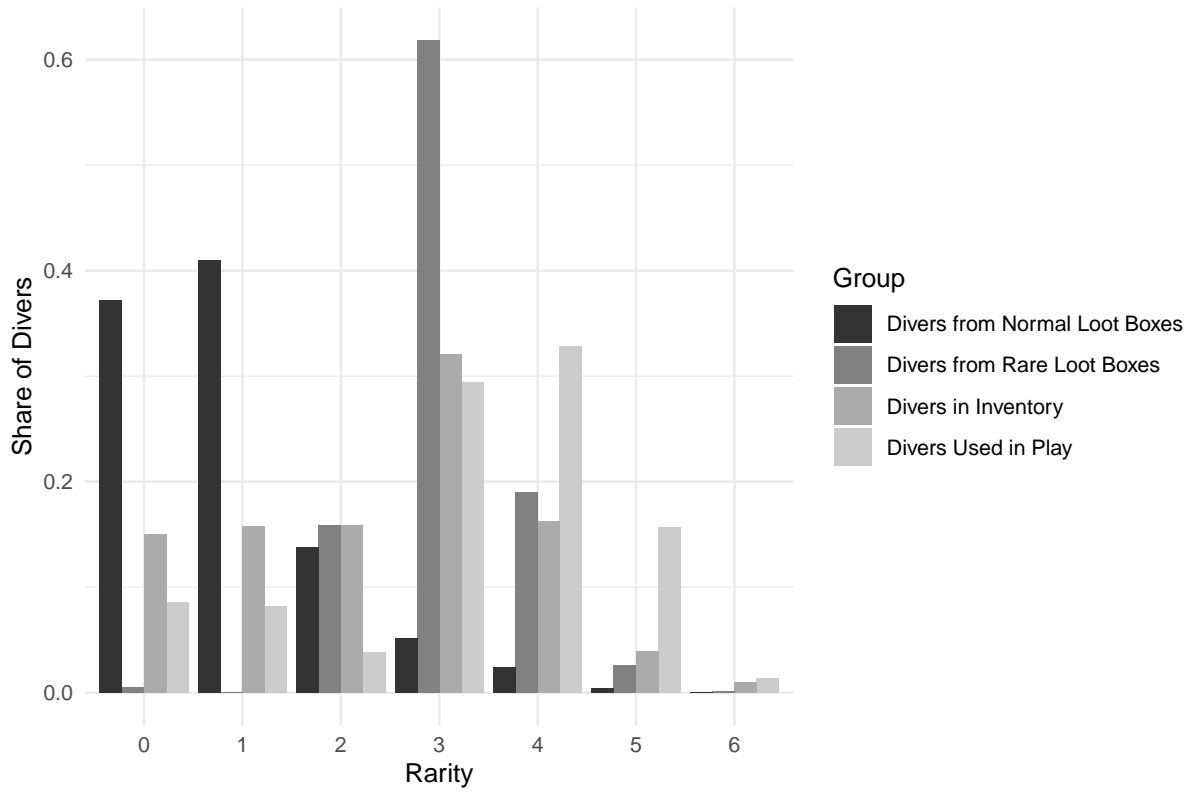
Bars correspond to the shares of rarity observed among the corresponding sets of divers.

Figure 9: Distribution of Color of Divers.
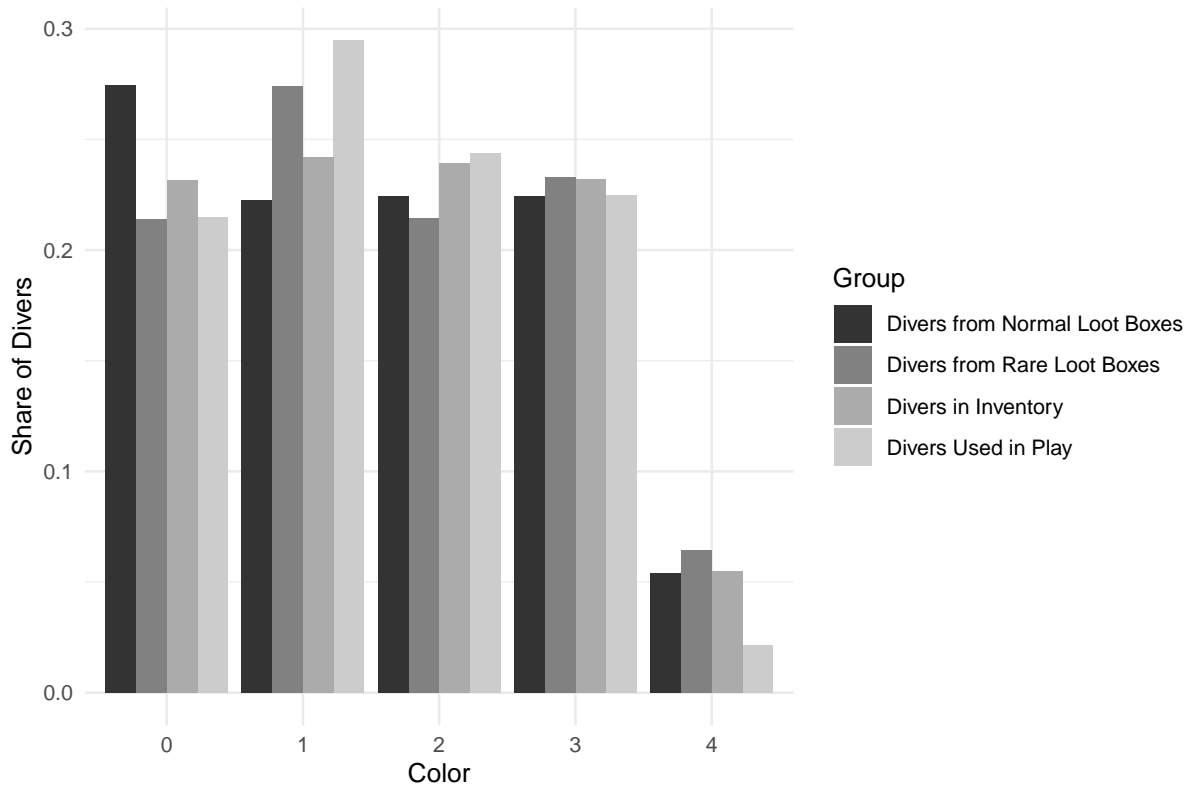
Bars correspond to the shares of color observed among the corresponding sets of divers.

Figure 10: The Number of Plays and Unique Players Present at Each Stage
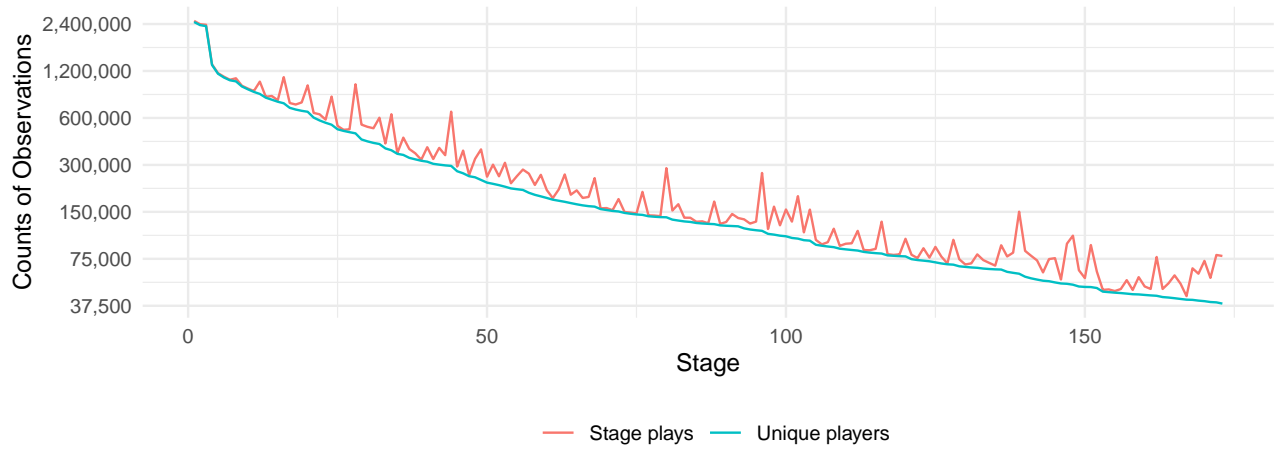
Main stage plays and unique players are counted if a player plays a stage that is one higher than the maximum stage won so far; that is, we do not count occasions of players going back and replaying stages they already won before.

Figure 11: Win, Loot Box Opening, and Coin Spending Probabilities for Each Stage, only for Players who Reached Stage 173
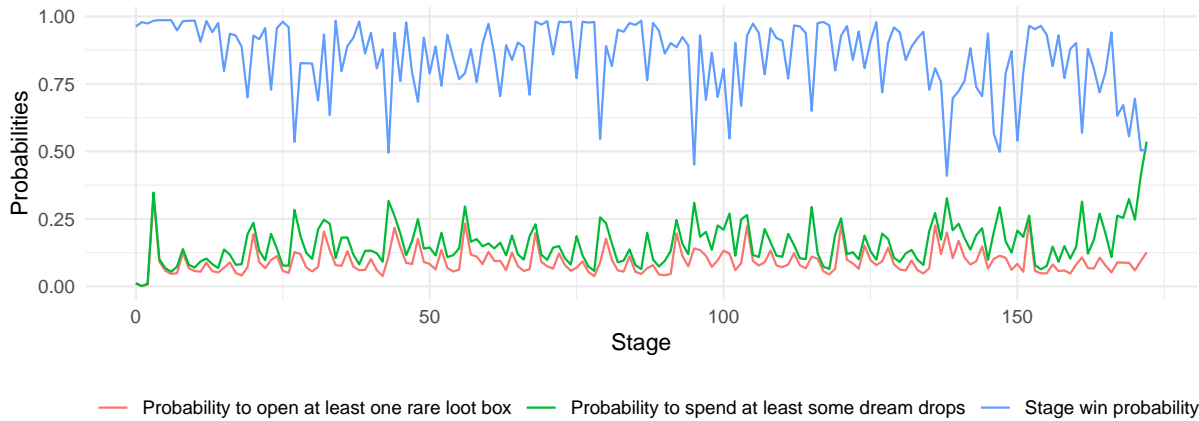
Win rates are calculated by dividing the number of times a stage has been won, by the times the stage was attempted, using only sequential occasions of main stage plays. We use only observations of players who reached and won stage 173. Probabilities of opening at least one rare loot box and spending at least some coins are calculated by taking an average across players who are at a given stage.

A4

Figure 12: Normalized Loot Box Opening and Coin Spending Probabilities for Each Stage
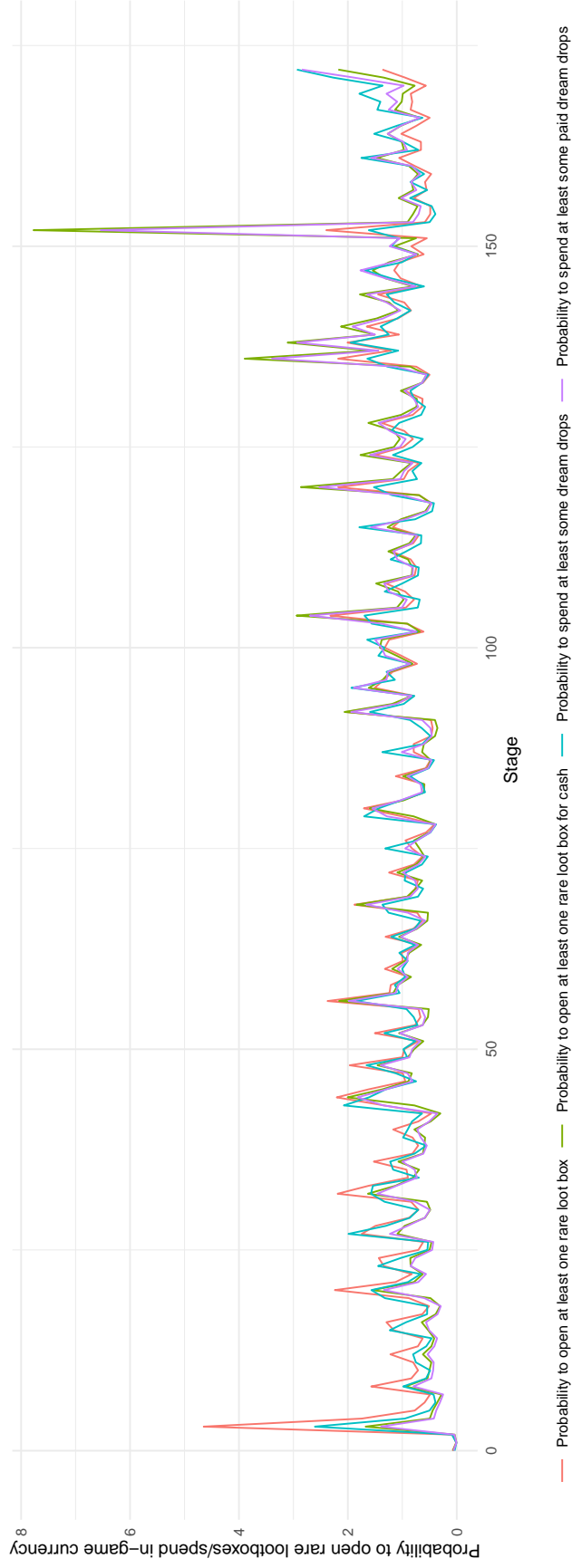
Probabilities of opening at least one rare loot box and spending at least some coins are calculated by taking an average across players who are at a given stage. We then normalize variable averages across stages to one to make visual comparisons simpler.

Table 7: Relationship Between Loot Box Openings, Coins Spendings, and Stage Complexity

| | Probability to open 1+ lootbox (log) | | Probability to spend (log) | |
| | in-game currency | real money | in-game currency | real money |
	(1)	(2)	(3)	(4)
Constant	−2.423***	−5.025***	−2.169***	−4.894***
	(0.063)	(0.073)	(0.058)	(0.069)
Stage win probability (log)	−0.765***	−1.000***	−1.729***	−1.418***
	(0.216)	(0.248)	(0.199)	(0.238)
Observations	173	173	173	173
R^2	0.068	0.087	0.306	0.172

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

The variables are constructed similar to Figure 2; all variables are in logs, meaning that coefficients should be interpreted as an effect of 1% change in stage win probability on percent change in the loot box opening and coins spendings probabilities.

A.2 Effect of the Inventory Rarity on Win Probability (ζ)

Here we use a more flexible regression specification to confirm that divers received from loot boxes do help improve the win probability of game levels. The vertical attribute of divers is rarity. We start by examining how the rarity of divers that players hold in their inventory affects the win probability, by estimating

$$\mathbb{1}\left(win_{it} = 1 | D_{it}, a_{it} = 1\right) = \kappa_i + \kappa'_s + \sum_{r \in 1:6} \zeta'_r \mathbb{1}\left[\max\left(rarity_d \in D_{it}\right) = r\right] + \eta X_{its} + \xi_{its} \quad (23)$$

where $\max\left(rarity_d \in D_{it}\right)$ represents the highest diver rarity that is in the player's i inventory at time t, and $\eta X_{its\tau}$ are additional controls, such as the number of times a player has played the game or opened loot boxes. ζ'_r captures the effect of having a rare diver in the inventory at an average stage for an average player. Standard errors are clustered two-way, on the user and stage level.

Table 8 presents the estimates of ζ'_r. In Column (1) are estimates from the baseline specification, with the user and stage-level fixed effects but without additional controls ηX_{its}. Plays with a maximum diver rarity of one in the inventory, $\max\left(rarity_d \in D_{it}\right) = 1$, are taken as a baseline. When the maximum rarity is 2 or 3, there is almost no detectable change from the baseline – win probability decreases by 0.6-1.2 percentage points. However, when the maximum diver rarity is 4, 5, or 6, the average win probability increases by 6.4, 12.6, and 16.8 percentage points, respectively.

Figure 13: Effect of extra rarity point on stage win probability
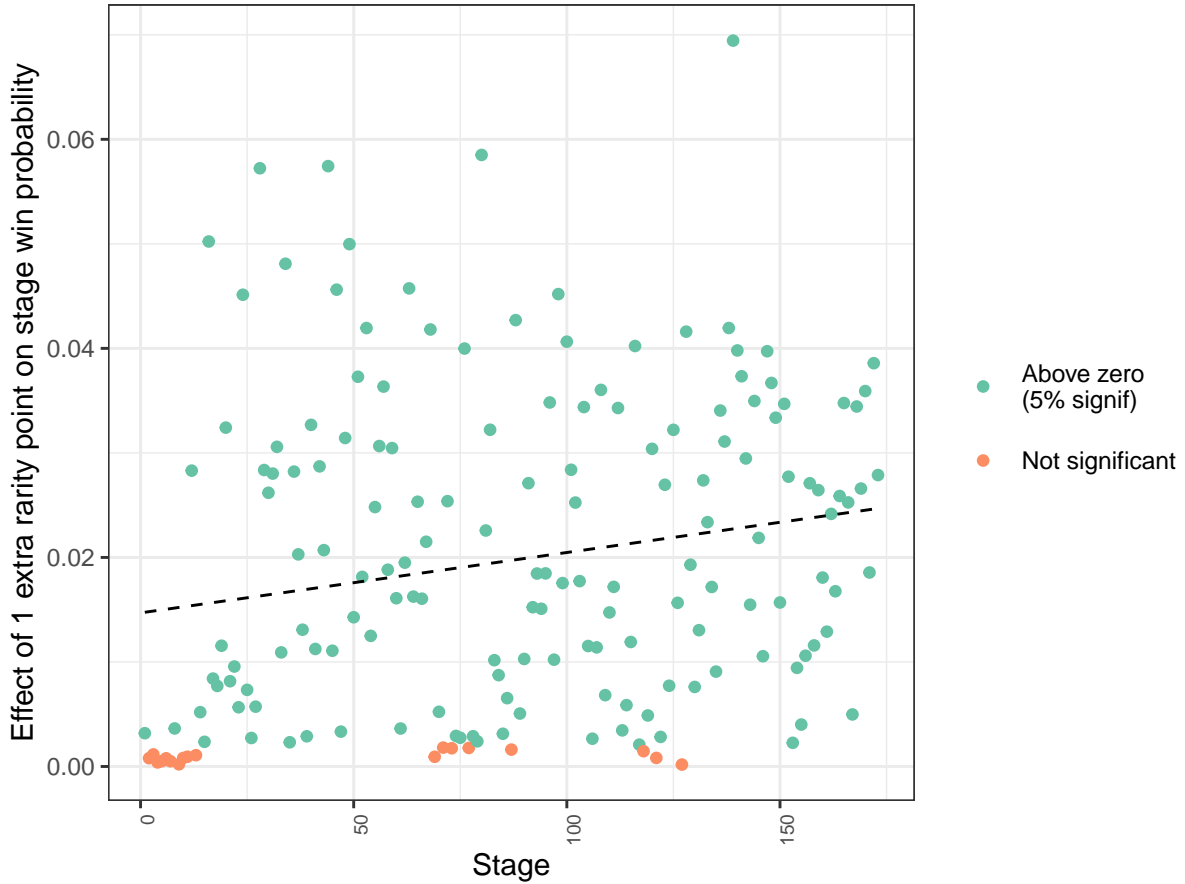
Each point represents a stage-specific estimate of the effect of the rarity of top four divers in players' inventories on stage win probabilities of these players. To compute these effects, we regress a player's likelihood to win a stage on the summed rarity of the top four divers in their inventory, similar to the variable constructed in Figure 3. We allow the effect of the summed rarity of the top four divers to be stage-specific by adding stage-rarity interaction terms, and control for stage and user fixed effects to use only within the stage and within user variation. Standard errors are clustered at the stage level.

Figure 14: Probability of Opening Rare Loot Box Given the Current Stock of Coins.
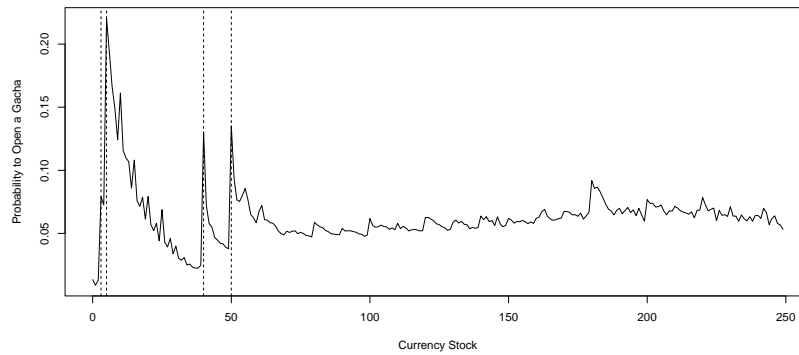
(a) Opened Using Any Coins
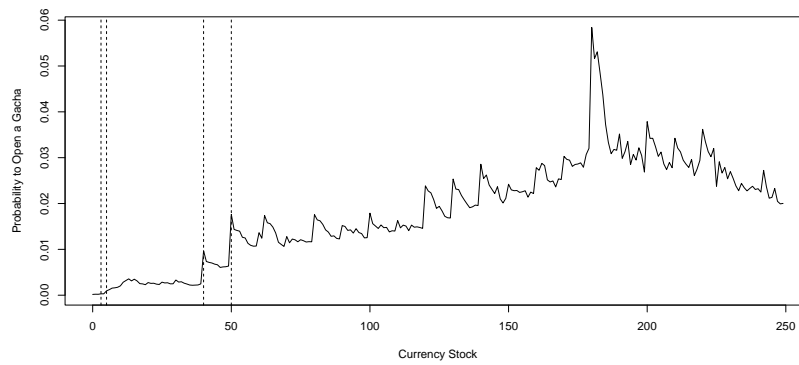
(b) Opened Using Paid Coins

The currency stock is computed using the observations of users' currency transactions.

Table 8: Relationship between Diver Rarity in Inventory and Win Probability

	Dependent variable: $\mathbb{1}\left(win_{it} = 1|D_{it}, a_{it} = 1\right)$			
	(1)	(2)	(3)	(4)
$\max(\text{rarity}_{it}) = 2$	-0.0121	-0.0107	-0.0118	-0.0104
	(0.0009)	(0.001)	(0.0009)	(0.0009)
$\max(\text{rarity}_{it}) = 3$	-0.0066	-0.0074	-0.0186	-0.017
	(0.0004)	(0.0004)	(0.0004)	(0.0004)
$\max(\text{rarity}_{it}) = 4$	0.0643	0.0544	0.0256	0.0257
	(0.0007)	(0.0007)	(0.0007)	(0.0007)
$\max(\text{rarity}_{it}) = 5$	0.1264	0.0972	0.0691	0.0625
	(0.0008)	(0.0009)	(0.0009)	(0.0009)
$\max(\text{rarity}_{it}) = 6$	0.1681	0.1276	0.1147	0.1011
	(0.0009)	(0.0009)	(0.0009)	(0.0009)
Fixed Effects:				
User	Yes	Yes	Yes	Yes
Stage	Yes	Yes	Yes	Yes
Number of loot box openings	No	Yes	No	Yes
Number of times played the game	No	No	Yes	Yes
R^2	0.299	0.302	0.304	0.305
Number of observations		52,842,816		

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

Columns (2)-(4) of Table 8 present the results after including the additional controls. The goal of these controls is to capture any evolution in players' skill or inventory that is not captured by the user and stage fixed effects. The estimates show that our results are robust to these additional controls. Even after including both the count of loot boxes opened and plays by the user (Column 4), the probability of winning is low when the maximum diver rarity in the inventory is 1, 2, or 3. In contrast, the probability of winning is 2.5, 6.3, and 10 percentage points higher when the maximum diver rarity in the inventory is 4, 5, and 6, respectively.

A.3 Estimation of Transition Probability Matrices

We estimate the state transition probability matrices from the data, using a combination of frequency estimators and local approximations by linear models.

A.3.1 Rarity

For the rarity state R_{it+1}, we estimate transition probabilities using the frequency estimator conditional on the current action (play, open 1 loot box, open a pack of 11 loot boxes), whether the player lost last time (q_{it}), the current rarity (R_{it}), and stage (s_{it}). There are 17,748 unique state-action combinations; if the combination has less than 100 observations to estimate the transition probability (40% of unique states, accounting for 0.17% of observations in the data), we replace the estimate with the frequency estimator based on action-q_{it}-R_{it}-stage groups combinations, where stage groups are defined as stage group$_{it} = 1 * I(s_{it} \leq 10) + 2 * I(s_{it} \in \{10, 40])+3*I(s_{it} \in \{40, 80])+4*I(s_{it} > 80)$. Results are robust to using alternative thresholds since very few observations used in the estimation are affected by this approximation.

Figure 15 visualizes marginal distributions of the transition probabilities of rarity states, conditional on different actions. If users choose to play the game (action 1), the rarity of divers in their inventory changes very infrequently. In contrast, if users open one and especially eleven loot boxes there is a substantial transition to higher rarity states. The probability of an increase in the rarity of inventory is higher if the current rarity state of the inventory is lower.

A.3.2 Currency Stock

For the currency state c_{it+1}, we estimate transition probabilities using the frequency estimator conditional on the current action (play, open 1 loot box, open a pack of 11 loot boxes), prices of 1 and 11 loot boxes (p_{it}^1 and p_{it}^{11}), the current currency state (c_{it}), and stage (s_{it}). Since the price of a loot box only affects the currency stock when the corresponding loot box is opened, we do not condition on loot box prices when the player chooses action 1 (resulting in 8,996 unique states), condition on the price of one loot box when the player chooses action 2 (17,992 unique states), and condition on the price of eleven loot boxes when the player chooses action 11 (35,984 unique states).

Figure 16 visualizes marginal distributions of the transition probabilities of currency states, conditional on different actions. If users choose to play the game (action 1), their in-game currency almost never goes down, and either stays the same or increases by 1-2 coins. If users open one loot box, their currency decreases by 3 or 5 coins, depending on the price of the loot box. If the user's currency is below the price of the loot box, with a high probability the currency stock goes to zero, reflecting the idea that the player pays with all the coins they have in the game and then contributes the difference by purchasing loot boxes with real money. Similarly, if the user opens a pack of eleven loot boxes, their currency stock very likely goes to zero.

Figure 15: Average Transition Probabilities of Rarity States, Across Actions
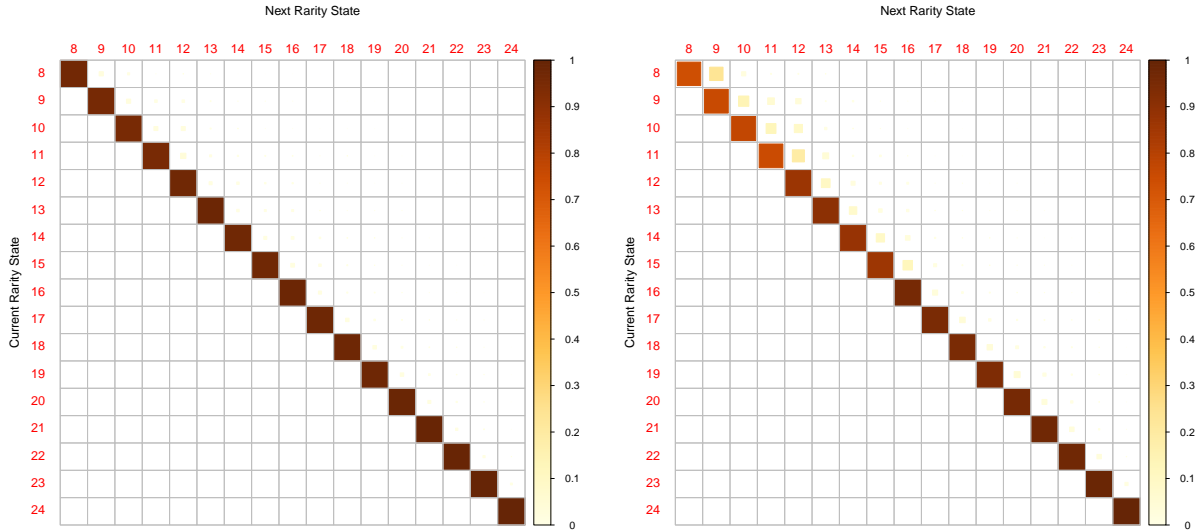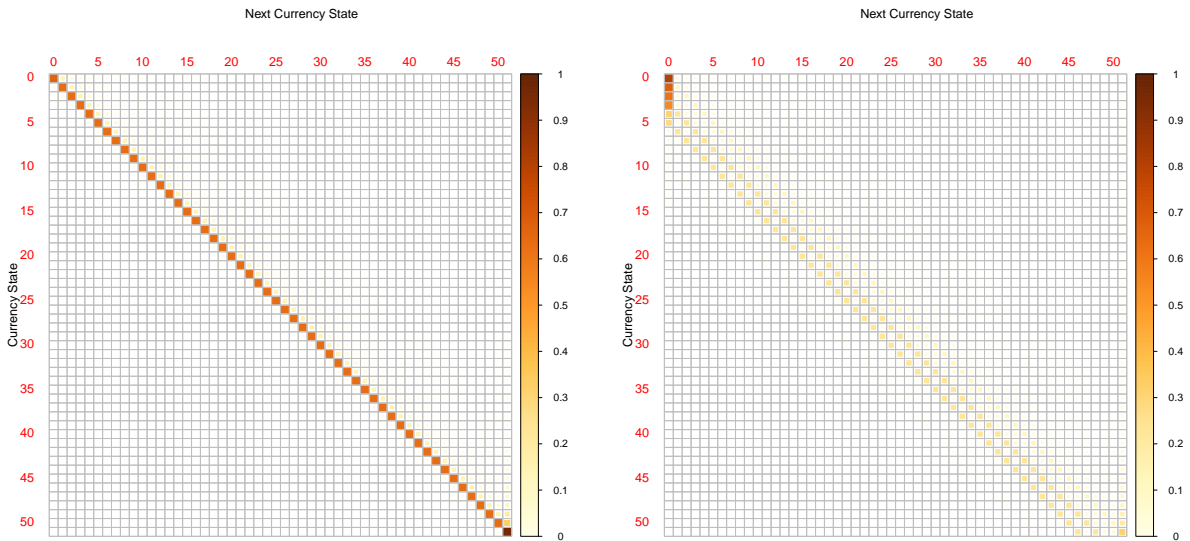

(a) Conditional on Action 1 (Play)

(b) Conditional on Action 2 (One Loot Box)
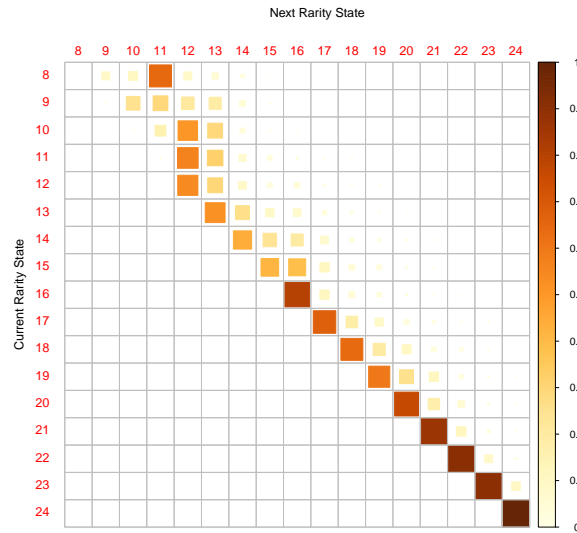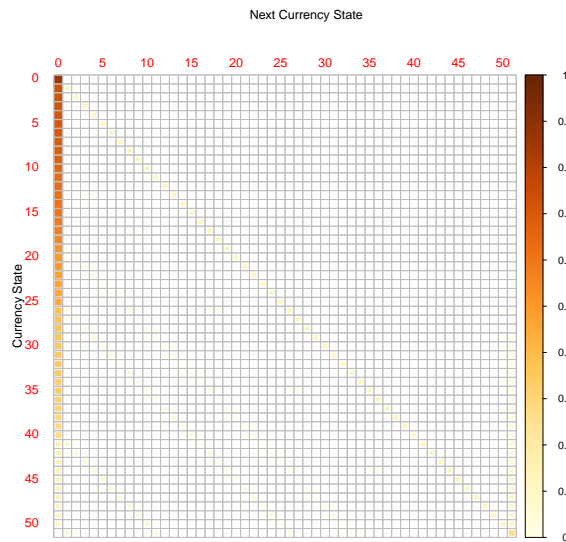
(c) Conditional on Action 3 (Eleven Loot Box)

Transition probabilities are averaged over stages s_{it} and whether the player lost last time q_{it}.

Figure 16: Average Transition Probabilities of Currency States, Across Actions

(a) Conditional on Action 1 (Play)

(b) Conditional on Action 2 (One Loot Box)

(c) Conditional on Action 3 (Eleven Loot Box)

Transition probabilities are averaged over stages s_{it} and loot box prices (for actions 2 and 3).

A.3.3 State Dependence

Transition probabilities for the state dependence variable, d_{it}, are trivial. State dependence gets assigned the value of one any time a player chooses to open a loot box in the previous period, as described by equation 4.

A.3.4 Loot Box Prices

We estimate transition probabilities for prices of 1 (11) loot boxes using the frequency estimator conditional on the current action (play, open 1 loot box, open a pack of 11 loot boxes), prices of 1 (11) loot boxes, and stage (s_{it}). Prices change only after the player opens the corresponding loot box – implying that prices for one loot box do not change when a user plays actions 1 and 3, and prices for eleven loot boxes do not change after actions 1 and 2. This means that we can estimate price transition probabilities only conditional on the current stage and the current price, leading to 348 and 696 possible states for prices of 1 and 11 loot boxes, respectively.

Figure 17 presents the estimates of probabilities of facing baseline prices for one and eleven loot boxes, $p_{it}^1 = 5$ and $p_{it}^{11} = 50$. As the game progresses, players are getting more discounts on these prices.

A.3.5 Stage and Loss Indicator

Transition probabilities for the stage and loss indicator are determined by whether or not a player wins the stage. The win probability function is defined by equation 11 – we allow for flexible stage-and-loss-specific fixed effects and stage-and-loss-specific effects of rarity. We estimate equation 11 separately for whales and non-whales, and the fitted values of the regression are the estimates of win probabilities for each state. We set the small % of probabilities that fall slightly outside of the $[0, 1]$ interval (1.8% for non-whales and 2.2% for whales) to the bounds values.

Figure 17 presents the estimates of probabilities of winning the stage across various stages and rarity states. Figure 17a plots the probabilities against stages of the game; as the game progresses, the win probability declines, and every four stages there are discrete drops in the win probabilities corresponding to the "boss" levels. Win probabilities are slightly higher for whales than non-whales.

Figure 17b presents average win probabilities by the rarity stage of players' inventories, by groups of stages. Across all groups, higher rarity inventory leads to higher stage win probability. However, the effect of the rarity state on win probability is more pronounced in the latter stages of the game.

A14

Figure 17: Probability to be Facing a Baseline Price of One ($p_{it}^1 = 5$) and Eleven ($p_{it}^{11} = 50$) Loot Boxes, Across Stages
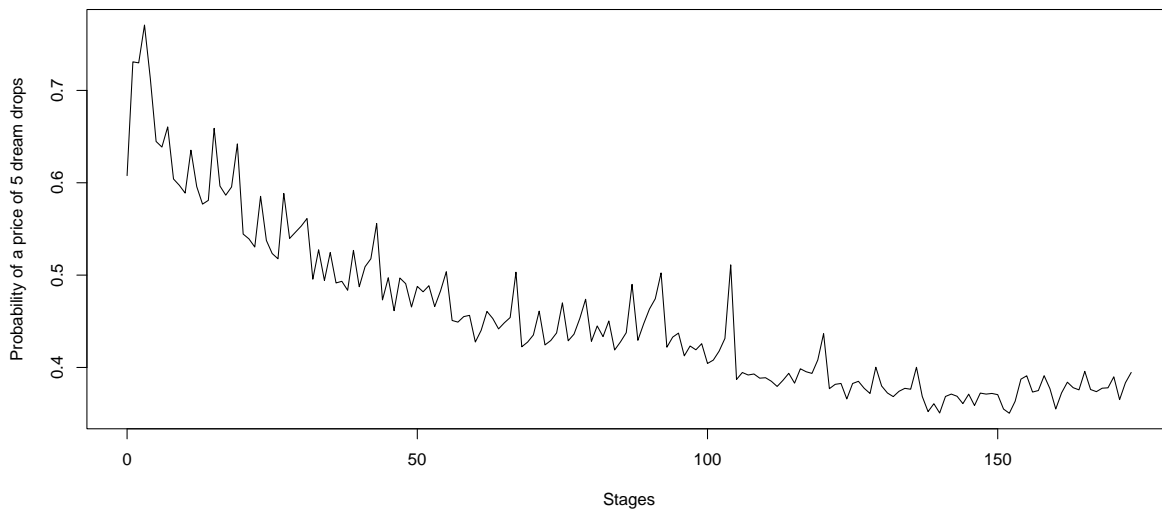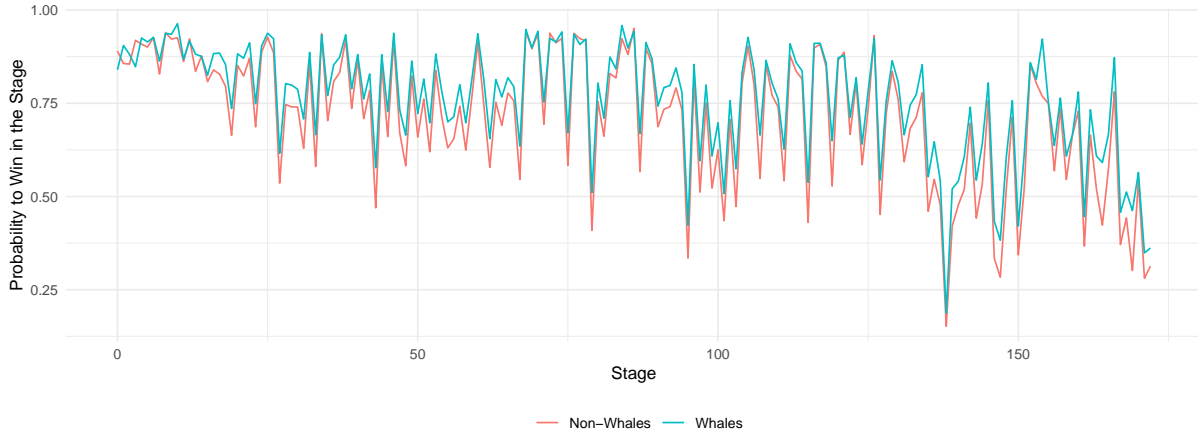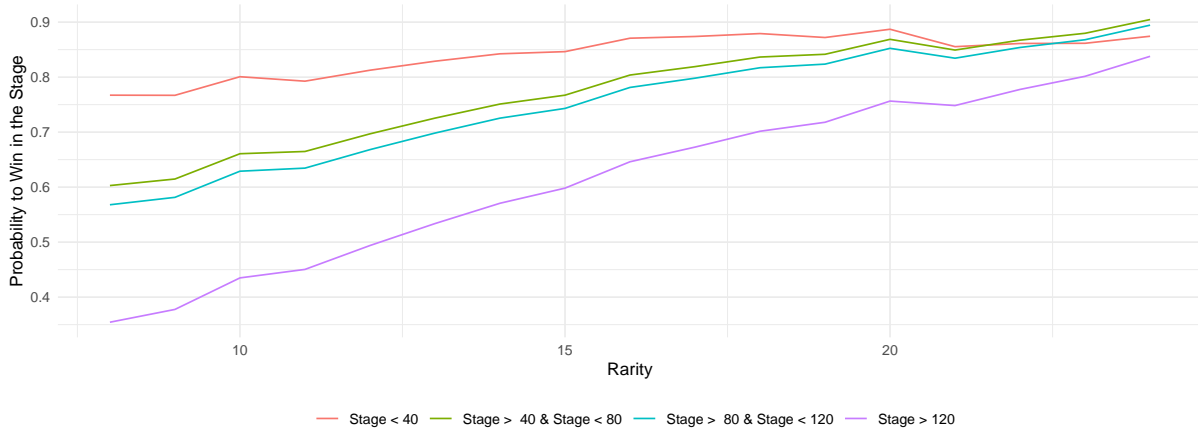

(a) Probability of Facing $p_{it}^1 = 5$
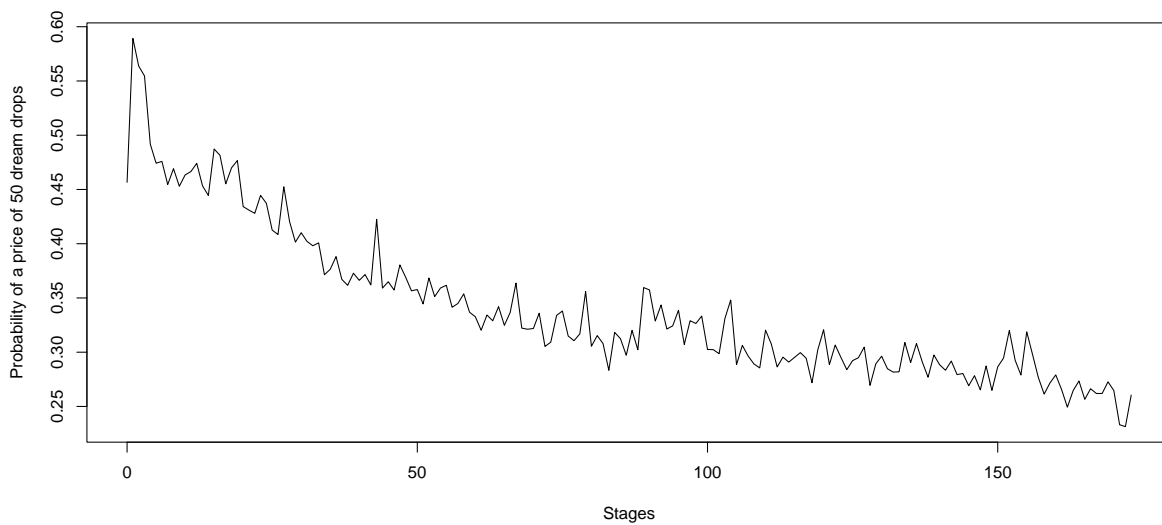
(b) Probability of Facing $p_{it}^{11} = 50$

Figure 18: Estimates of Probabilities to Win the Stage, Across Stages and Rarity States

(a) Across Stages, by Whales/Non-Whales

(b) Across Rarity States, by Stage Groups

Probabilities are the fitted values using the estimates of equation 11 for whales and non-whales, averaged across the states. Subfigure (a) plots the probabilities across stages for whales and non-whales. Subfigure (b) plots probabilities across rarity states by groups of stages, taking an average of estimates for whales and non-whales.

A.4 Heterogeneity of Players' Tastes

In our main specification, we allow for heterogeneity along players' monetary importance for the gaming company – whales, the 1.5% of players who are responsible for 90% of revenues, and non-whales, the rest of the players. Splitting the sample by players' in-game spending is important for understanding a key trade-off: should the company target the design of the game and loot box toward a small minority of paying customers? Or should it target the vast majority of non-paying customers who are still enjoying the game and promote it, either directly through word-of-mouth or indirectly through the apps' rankings in the app store?

However, how much players spend might be not the only critical dimension of heterogeneity for characterizing tastes. This is particularly important since we estimate a reduced-form state dependence parameter, and interpret it as a proxy for the players' impulsive consumption of loot boxes. Past literature has shown that not accounting for heterogeneity can inflate state dependence estimates [e.g. Dubé et al., 2010, Simonov et al., 2020].

We extend our analysis in two ways to test whether parameter estimates, particularly the degree of state dependence in players' choices of loot boxes, vary when we allow for more heterogeneity. First, we split the sample of players along a different dimension – how many events the players have participated in throughout the game. Game events do not count as stage-play in our estimation so clustering does not rely on our main data sample, but the degree to which players participate in events should capture their engagement with the game. We group the top 10% of players by participation in events as high-engagement players, corresponding to those who have participated at least in 50 event games. As expected, high-engagement events players are also heavier players of the main stages of the game and open more loot boxes. As a group, they are responsible for 55.6% of all main stage plays and 58% of rare loot box openings.

Table 9 presents parameter estimates for low- and high-engagement players of event games. Parameter estimates are consistent across the two groups – they share the same level of state dependence preferences, preferences for eleven-pack loot box, payment, and past loss of the played stage. The only significant difference in the preferences of low- and high-engagement players of event games is their preference for loot boxes over the gameplay – high-engagement stage players also have a higher preference for opening loot boxes – which is expected since their players are more invested in the game.

In our second extension of the heterogeneity analysis, we separate out whales and non-whales into additional clusters based on the player-level propensities to play main-stage games over opening loot boxes. For this, we cluster players based on their average preference for playing the main stage – overall and conditional on different levels of currency, rarity, loss, and state dependence states – and allow for two separate clusters (high and low preference for gameplay) for both whales and non-whales.[29] The resulting clusters are approximately equal in

---

[29]To measure players' average preference for playing the main stage, we compute residuals in a model of playing the main stages of the game over opening loot boxes using the linear probability model in Equation 16 (allowing us to control for the player's state). We then use k-means clustering based on averages of the residuals for each player, with residuals unweighted and interacted with the current levels of currency, rarity,

Table 9: Estimates of Preference for Loot Boxes and Winning in the Game, Players Grouped by Activity

	Play ≤ 50 event stages		Play > 50 event stages	
	Estimate	S.e.	Estimate	S.e.
One loot box ($\alpha_{L,1}$)	-3.1063	(0.5607)	0.3576	(0.1014)
State dependence (η)	1.1742	(0.2443)	0.9224	(0.0654)
Eleven-pack loot box ($\alpha_{L,11} - \alpha_{L,1}$)	-0.7636	(0.4206)	-0.9704	(0.3078)
Payment (γ)	-0.1505	(0.0114)	-0.1640	(0.0025)
Lose the game (-β)	-0.4834	(0.2079)	-0.6376	(0.0777)

Players are grouped by the number of special events they participate in. Standard errors are computed using Bayesian bootstrap, with clustering (draws of observation weights) done at the stage level.

size – 52% of non-whales and 45% of whales are allocated in the high-preference for gameplay cluster.

Table 10 presents the estimates for all four groups of players – whales and non-whales with high and low preferences for gameplay. As expected, in both groups players with a high preference for gameplay have lower estimates of their preferences for loot boxes. The state dependence tastes of players are consistent across high- and low-preference for gameplay clusters for whales – for both groups, η estimates are around 2-2.1 – whereas for non-whales low-preference for gameplay cluster (cluster 1) exhibits negative but only marginally significant state dependence in loot boxes opening (η estimate is -0.63 with a standard error of 0.31). Non-whales are more price elastic compared to whales in both high- and low-preference for gameplay groups. Finally, the high-preference for gameplay group (cluster 2) does not care as much about losing the game, but overall losses are still more costly for whales compared to non-whales.

Table 10: Estimates of Preference for Loot Boxes and Winning in the Game, Players Clustered within Groups

	Non-Whales				Whales			
	Cluster 1		Cluster 2		Cluster 1		Cluster 2	
	Estimate	S.e.	Estimate	S.e.	Estimate	S.e.	Estimate	S.e.
One loot box ($\alpha_{L,1}$)	0.3809	(0.2428)	-2.1501	(0.3794)	0.3291	(0.2371)	-0.7989	(0.5314)
State dependence (η)	-0.6333	(0.3082)	1.5379	(0.2199)	2.1345	(0.3013)	1.9681	(0.3854)
Eleven-pack loot box ($\alpha_{L,11} - \alpha_{L,1}$)	-1.4325	(0.5124)	0.6491	(0.3288)	-0.9502	(0.1935)	0.7394	(0.0981)
Payment (γ)	-0.1597	(0.0190)	-0.1995	(0.0070)	-0.1539	(0.0031)	-0.1577	(0.0049)
Lose the game (-β)	-1.8198	(0.2930)	-0.2598	(0.1005)	-2.0050	(0.0852)	-1.3874	(0.0701)
Number of players	1,191,799		1,281,781		20,568		16,807	

Standard errors are computed using Bayesian bootstrap, with clustering (draws of observation weights) done at the stage level.

Overall, we conclude that state dependence of users' choices of loot boxes is prominent even if we allow for more heterogeneity in consumer preferences, in the dimension of players' preferences for play over loot boxes. We confirm that our key takeaways are robust to al-

loss, and state dependence.

lowing for heterogeneity within the groups of whales and non-whales when we decompose the mechanisms behind the loot box value for players.

A.5 Additional Estimation Results

Table 11: Decomposition of the Loot Box Value, Overall by Cluster

| | Non-whales | | Whales | |
	Cluster 1	Cluster 2	Cluster 1	Cluster 2
Functional value	40.65	90.08	0.34	15.82
	(7.89)	(2.05)	(0.08)	(2.55)
Direct value	59.35	9.92	99.66	84.18
	(7.89)	(2.05)	(0.08)	(2.55)
State Dependence	-8.43	2.72	26.94	35.99
	(4.44)	(0.55)	(2.46)	(8.33)
Option to loot	67.78	7.2	72.72	48.19
	(8.74)	(1.82)	(2.39)	(9.98)

Standard errors are computed using Bayesian bootstrap, with clustering (draws of observation weights) done at the stage level.

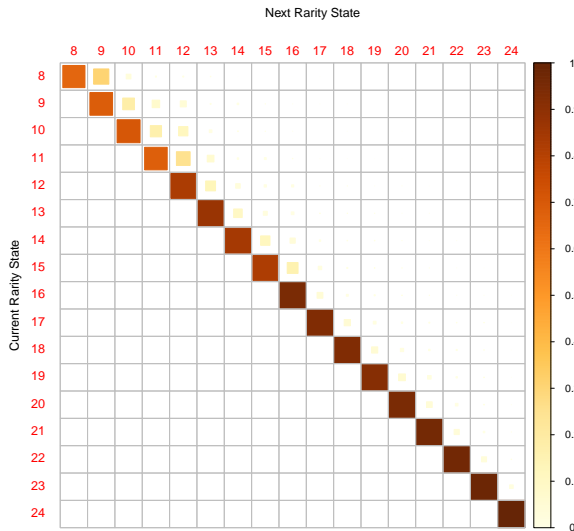## A.6 Counterfactual Rarity Transition Probabilities

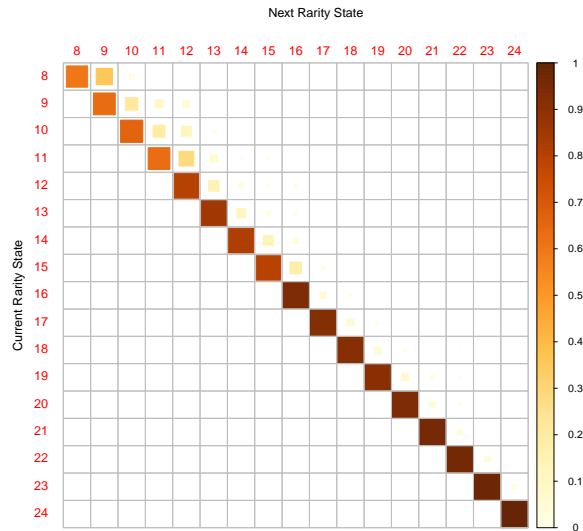Figure 19: Counterfactual Transition Probabilities of Rarity States, Action 2 (One Loot Box Opened)
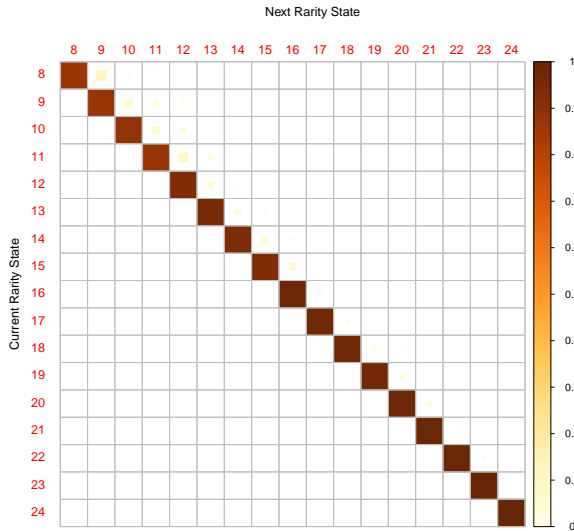
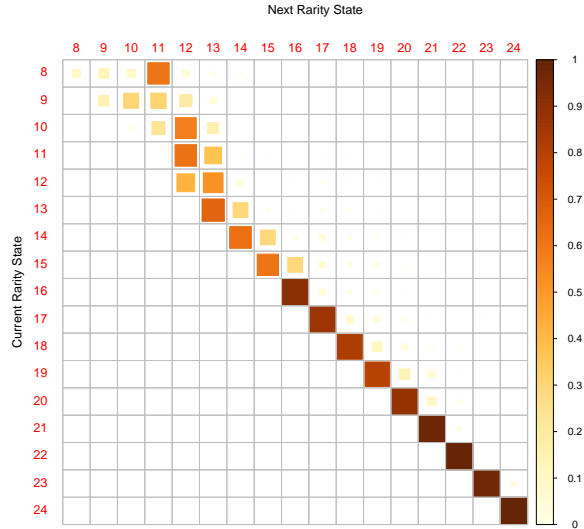(a) St. Dev. -50%
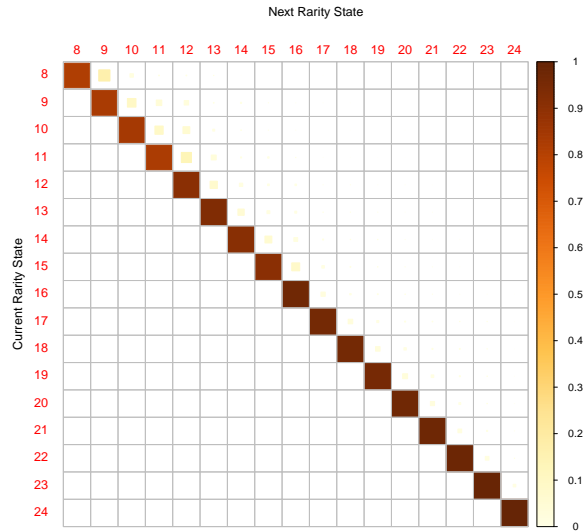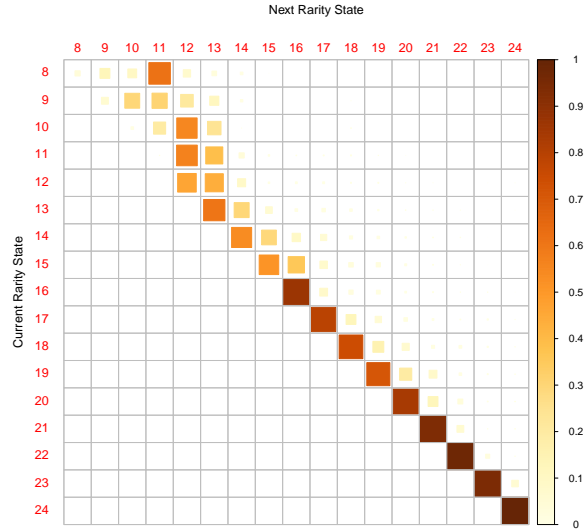

(b) St. Dev. -30%

(c) St. Dev. +30%

(d) St. Dev. +50%

Transition probabilities are averaged over stages s_{it} and whether the player lost last time q_{it}.

Figure 20: Counterfactual Transition Probabilities of Rarity States, Action 3 (Eleven Loot Boxes Opened)
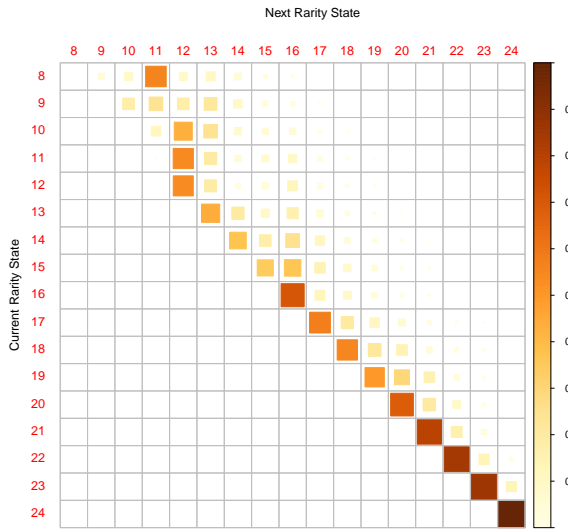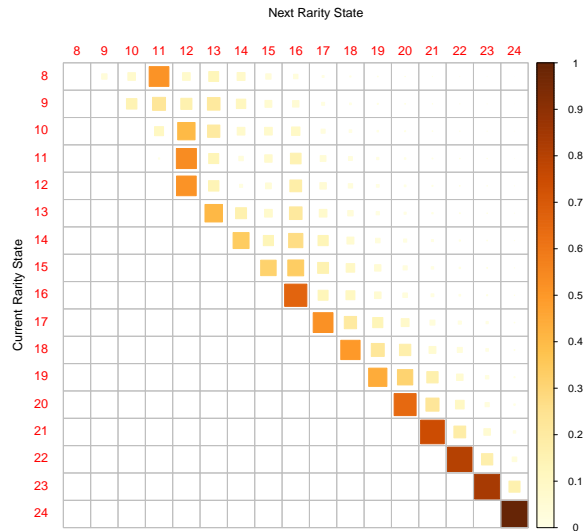

(a) St. Dev. -50%

(b) St. Dev. -30%
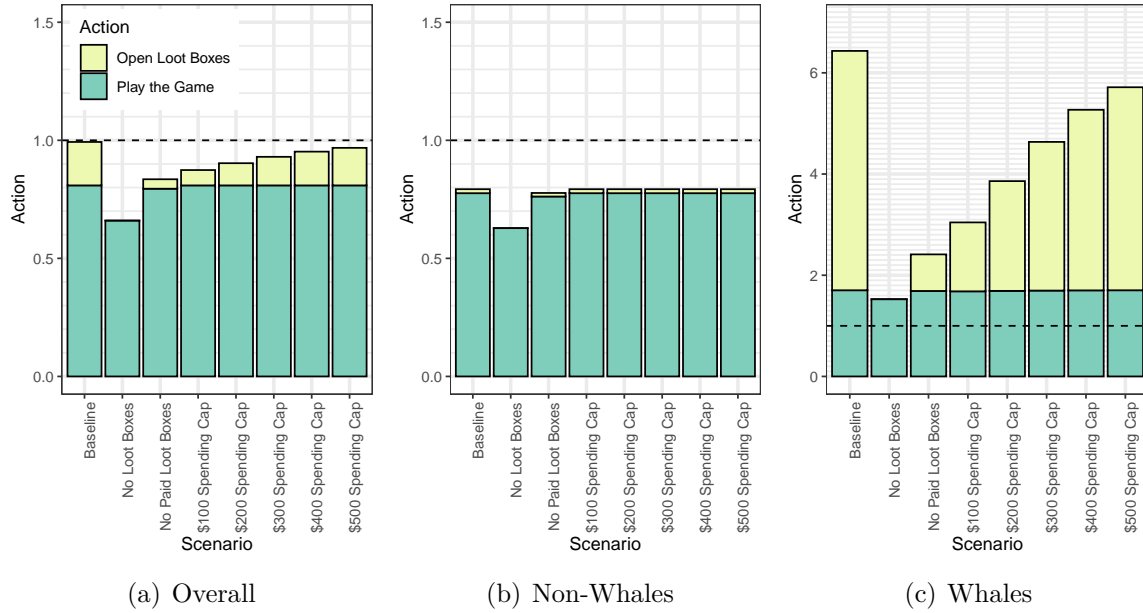
x

(c) St. Dev. +30%

(d) St. Dev. +50%

Transition probabilities are averaged over stages s_{it} and whether the player lost last time q_{it}.

A.7 Additional Counterfactual Results

Figure 21: Players' Actions under Loot Box Bans and Spending Caps

(a) Overall   (b) Non-Whales   (c) Whales

Estimates of the number of consumer actions by scenario and types of players. The overall number of actions per (average) player under the current scenario is normalized to one (dashed line).

Table 12: Simulations under different standard deviations of loot box outcomes

Panel I: Revenue

	Baseline			No State Dependence		
	Overall	Non-whales	Whales	Overall	Non-whales	Whales
Lower st. dev.						
st. dev. -50%	1.3	0.07	34.53	0.09	0.05	1.19
st. dev. -30%	1.05	0.07	27.78	0.09	0.05	1.13
Current st. dev.	**1**	0.07	26.24	0.09	0.05	1.07
st. dev. +30%	0.94	0.07	24.52	0.09	0.05	1.03
st. dev. +50%	0.96	0.07	25.19	0.08	0.05	1.02
Higher st. dev.						

Panel II: # of Loot Boxes Opened

	Baseline			No State Dependence		
	Overall	Non-whales	Whales	Overall	Non-whales	Whales
Lower st. dev.						
st. dev. -50%	1.27	0.13	31.98	0.15	0.1	1.7
st. dev. -30%	1.05	0.13	26.04	0.15	0.1	1.62
Current st. dev.	**1**	0.13	24.64	0.15	0.1	1.55
st. dev. +30%	0.95	0.13	23.27	0.14	0.1	1.48
st. dev. +50%	0.97	0.13	23.94	0.14	0.1	1.44
Higher st. dev.						

Panel III: # of Stage Games Played

	Baseline			No State Dependence		
	Overall	Non-whales	Whales	Overall	Non-whales	Whales
Lower st. dev.						
st. dev. -50%	1.01	0.97	1.99	0.97	0.94	1.99
st. dev. -30%	1	0.96	1.98	0.97	0.93	1.99
Current st. dev.	**1**	0.96	1.96	0.97	0.93	1.98
st. dev. +30%	0.99	0.95	1.94	0.96	0.92	1.98
st. dev. +50%	0.99	0.95	1.93	0.96	0.93	1.98
Higher st. dev.						

We simulate how players would engage with the game if loot boxes were manipulated to result in higher or lower standard deviation in outcomes, with and without state dependence in loot box openings. Revenue is the average sum of payments made from opening rare loot boxes over the lifetime of a player. The number of times the player played the stage game includes the attempts the player lost. The weighted average value of each metric across non-whales and whales under the current design of the game is normalized to unity (presented in bold).