# Managing Congestion in Two-Sided Platforms: The Case of Online Rentals*

Caterina Calsamiglia†    Laura Doval‡    Alejandro Robinson-Cortés§    Matthew Shum¶

## Abstract

Thick two-sided matching platforms, such as the room-rental market, face the challenge of showing relevant objects to users to reduce search costs. Many platforms use ranking algorithms to determine the order in which alternatives are shown to users. Ranking algorithms may depend on simple criteria, such as how long a listing has been on the platform, or incorporate more sophisticated aspects, such as personalized inferences about users' preferences. Using rich data on a room rental platform, we show how ranking algorithms can be a source of unnecessary congestion, especially when the ranking is invariant across users. Invariant rankings induce users to view, click, and request the same rooms in the platform we study, greatly limiting the number of matches it creates. We estimate preferences and simulate counterfactuals under different ranking algorithms varying the degree of user personalization and variation across users. In our case, increased personalization raises both user match utility and congestion, which leads to a trade-off. We find that the current outcome is inefficient as it lies below the possibility frontier, and propose alternatives that improve upon it.

# 1 Introduction

Two-sided matching markets have experienced a significant surge in relevance in recent years, reshaping various industries and fundamentally altering the dynamics of supply and demand. The advent of digital platforms and the proliferation of online interactions have propelled the prominence of two-sided markets to new heights. Examples include ride-sharing and transportation, dating and social networking or, as our present study, accommodation and hospitality. Successful platforms require careful analysis of factors that may affect their capacity to successfully bring together and match both sides of a market. Various design aspects that can have an impact on platforms' effectiveness need to be continuously revisited. The present paper does exactly that; using a specific application, we present a problem and a set of potential solutions applicable to a wide set of platforms.

Given the limited amount of attention users have to process the increasing number of alternatives in online marketplaces, platforms typically employ computational algorithms that steer users towards options considered good matches for them. In this paper, we use rich data on a leading room rental platform to show how currently used algorithms may induce an unnecessary amount of congestion and, in doing so, limit a platform's capacity to exhaust the set of potential matches in a market.

An unintended consequence of the steering algorithms employed by platforms is that they may create additional unnecessary congestion in the market. As in many other dynamic platforms, in the platform we study an algorithm is implemented to prioritize recent and reliable options for users to be matched with. The algorithm generates a partial order, which ties are then broken randomly to generate a strict ranking of all apartments. Crucially, the ties are broken using the *same* random number across all users. This induces users to view, click and request a relatively small set apartments while ignoring similarly valuable ones. We find that the platform's algorithm leads to unnecessary congestion since many users end up clicking on the same small set of apartments, while they would have been equally, or even better off, exploring alternative apartments that are ranked lower by the platform's algorithm and hence do not receive enough attention and consideration from users.

The problem in the current algorithm comes from two main sources. The first has to do with the fact that the order in which rooms are shown greatly determines what individuals view, click and request, which affects final matches. The second is driven by the fact that different users may have different preferences and hence there is a loss from having all users search in the same order. Hence, there are at least two possible solutions. The first and computationally simpler is to have the complete rank be defined by the partial order where ties are broken using a different random number for different

users. The second involves personalizing the ranks for specific subgroups to ensure that the rank responds more clearly to individual preferences.

In our paper we analyze the trade-offs from personalizing more according to individual preferences to target demand to most desired rooms versus creating more randomness in the way rooms are ranked. While highly personalized algorithms steer users to highly preferred alternatives, they may cause congestion if users have similar preferences. Introducing randomization to ranking algorithms alleviates congestion by steering users to different rooms. The relationship between maximal utility when matched versus congestion is the crux of our analysis: can we reduce congestion without penalising users by showing them different rooms?

We estimate preferences and simulate counterfactuals under different ranking algorithms which vary in the degree of user personalization. We find a trade-off between user match utility and congestion, as increased personalization raises both of them in our setting. In terms of congestion, the best algorithm is the one that uses no information about users' preferences and randomizes the ranking of rooms across users. Algorithms that incorporate users' preferences steer users to more highly preferred rooms but generate more congestion. In particular, we show that the algorithm used by the platform during the period we study is inefficient as its outcome lies below the possibility frontier in the utility-congestion space. Some of the algorithms we design are more efficient compared with the status quo. They reduce congestion without decreasing the expected match utility of the rooms users are steered to, or, conversely, increase utility without reducing congestion.

Inefficiencies arising from congestion, as well as the trade-offs faced by platforms when steering consumers to their most preferred alternatives, have been documented in a handful of online platforms: Fradkin (2017) studies search engine design and rejection rates in Airbnb; Horton (2019) studies the effects that signaling features have on matching rates in online labor markets; Chen, Hsieh, and Lin (2023) study matching patterns in an online dating platform and like us, point out that recommendation algorithms used by tech companies may lead to congestion; and Bai, Chen, Liu, Mu, and Xu (2022) study how congestion and information frictions affect firm dynamics and market efficiency in AliExpress.

From a theoretical perspective, alleviating the inefficiencies caused by congestion in two-sided markets has long been seen as one of the goals of effective market design (Roth and Xing, 1997; Roth, 2008). For example, Arnosti, Johari, and Kanoria (2021) study the welfare losses due to inefficient equilibrium search and screening in congested markets; Immorlica, Lucier, Manshadi, and Wei (2021) study the benefits of directing consumer search by limiting choice, which reduces congestion and the cannibalization of "low type" sellers' demand by "high types" ones; Ashlagi, Krishnaswamy, Makhijani, Saban, and

Shiragur (2022) study the trade-offs between expanding choice and increasing matching when designing recommendation algorithms in two-sided platforms; Mekonnen (2023) shows that there is more congestion when preferences are common and quality is revealed with directed search than in random matching. The main theoretical insight of our paper relates to how the effects of congestion are mediated by the differentiation present in consumer preferences, which is evidence of how optimal market design depends on market-specific features. In our case, the trade-off between utility and congestion is driven by the amount of horizontal and vertical differentiation in the market.

A literature parallel to the papers above studies the effects of algorithms on consumer search in online marketplaces (e.g., Koulayev, 2014; Hodgson and Lewis, 2020). While the focus of our analysis is not on modeling optimal consumer search paths, our empirical findings echo the findings in this literature, especially in relation to ranking algorithms.[1] The importance of rankings in online search have been widely documented (e.g., Ghose, Ipeirotis, and Li, 2014; Chen and Yao, 2017; Dang, Ursu, and Chintagunta, 2022; Farronato, Fradkin, and MacKay, 2023). For example, a common finding in hotel booking platforms is that presenting the best alternatives at the top reduces the search costs associated with finding the right product (De los Santos and Koulayev, 2017; Ursu, 2018). However, the platforms studied in these papers are largely many-to-one markets in which product capacity is large enough for any number of users to be matched to the same product; in these markets, congestion is not a first-order issue, leading to design implications which differ greatly from those explored in this paper.

In the next subsection (Section 1.1), we present descriptive evidence to motivate our analysis. In Section 2, we present our model, estimation strategy, and a description of our empirical specification. In Section 3, we discuss the estimation results, and, in Section 4, we study counterfactual exercises. We conclude in Section 5. In Appendix A, we report additional analyses and tables which we refer to throughout the text.

## 1.1 Congestion: motivating evidence

Our platform is one of the largest peer-to-peer rental markets in the world for medium-term rooms and apartments, active in many of the biggest cities in Europe, the Americas, and Asia. During the period of our analysis, it differentiated itself from other competing platforms such as Airbnb by renting rooms for a longer term, at least a month, targeting exchange students and workers. Our platform operates via a website or mobile app in which users interested in rentals submit searches and are shown available apartments listed by landlords. We observe the sequence of clicks to view rooms and submit requests to landlords done by each user. Our platform simply brings together renters and

---

[1]As we discuss below, we model consumer clicking and requesting choice conditional on search. While this implies making the assumption that consumer search is invariant to the ranking algorithm when conducting counterfactuals, we view this favorably compared to making stronger structural assumptions in relation to consumer search behavior.
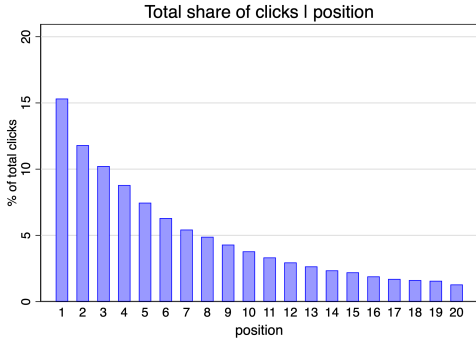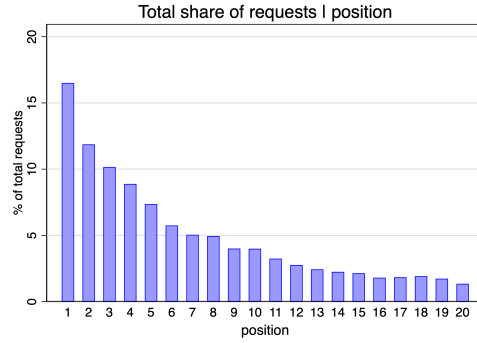
4

Figure (a) Share of clicks                    Figure (b) Share of requests

Figure 1: Concentration of clicks and requests by position

landlords, without always participating in the booking process; in that sense our study platform has more in common with Craigslist than Airbnb, and our dataset contains limited information on what happens after a room is requested by a user. Thus the empirics in this paper focus on users' click and request behavior, for which we have complete data.

Our dataset consists of 100% of all the listed rooms and searches for Barcelona, a major city for business and university studies in Europe. In presenting search results to users, our platform ranks rooms according to an algorithm that uses several features—including the date the room was posted, and whether the room is associated with a "registered" landlord who has linked their bank account on the platform—to determine the rank. A room's rank determines the position in which it is shown in the page of search results. The established criteria only determine a partial order across rooms that is subsequently randomized to generate a complete order. A key feature of this algorithm is that the *same* randomization is used to break ties between rooms across all renters, so that room rankings are invariant across renters.[2] Hence, while renters may be shown different sets of rooms depending on their search criteria, the rooms will be shown *in the same order* to all renters at a given point in time.

Figure 1 shows that clicks (Figure 1a) and room requests (Figure 1b) are concentrated on rooms shown to users at top positions. Specifically, we see that over 15% of all clicks and requests in the data are made to rooms shown to users at the first position. This percentage falls to around 4% for the tenth position, and asymptotes to around 2% for the 20-th position. This evidence suggests that rankings matter to renters' behavior. Indeed, the top two graphs of Figure 2 show that rooms shown in higher positions are more likely to be clicked on (Figure 2a) and requested (Figure 2b).

---

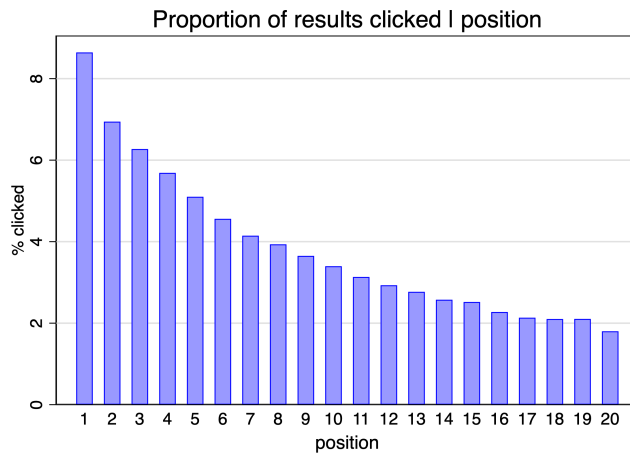[2]We use user, renter, and tenant, interchangeably.

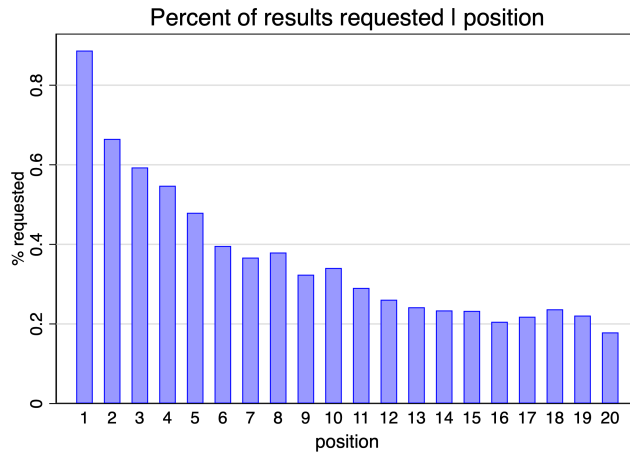Figure (a) $\mathbb{P}(\text{Click})$



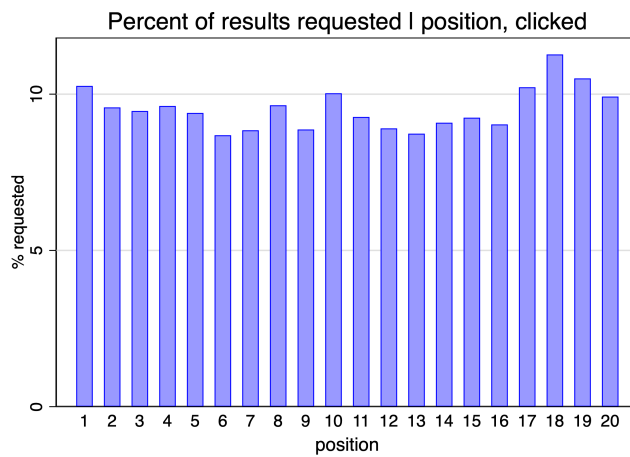Figure (b) $\mathbb{P}(\text{Request})$



Figure (c) $\mathbb{P}(\text{Request} \mid \text{Click=1})$

Figure 2: Probability of click and request by position

However, the bottom graph (Figure 2c) offers an interesting counterpoint. Conditional on being clicked, rooms in higher positions are not more likely to be requested than those in lower positions.[3] Specifically, conditional on having been clicked on, a top-ranked room and the 20-th ranked room are both requested with a probability close to 0.10. This suggests that the rooms' page ranking is to some extent orthogonal to the users' preferences over rooms, as it affects only whether users click on a room (which arguably depends on users' attention), but not whether they request a room (which depends on their preferences over the room's characteristics). Since rooms have to be clicked on before they can be requested, this also suggests that the concentration at the top positions observed in Figure 1 might be driven more by clicking behavior and attention than preferences.

Figure 3 presents initial evidence relating to the congestion on the platform across rooms. To illustrate congestion, we use Lorenz curves, which measure the degree of concentration, or congestion, across rooms in terms of the frequency with which they appear as search results, and are clicked on or requested. Higher congestion bends the Lorenz curve towards the plot's upper-left corner and indicates that the frequency is nonuniform across rooms; lower congestion moves the Lorenz curve towards the 45-degree line, indicating that the distribution across rooms is closer to uniform.

The top graph in Figure 3 shows the Lorenz curve for search results, which illustrates a high degree of congestion: the 20% of rooms that appear most often in users' search results (on the $x$-axis) accumulate around 65% of the total number of search results in the data (on the $y$-axis). This congestion only worsens when we consider clicked rooms, as from the middle graph (Figure 3b) we see that the top 20% of rooms in terms of clicks account for almost 70% of the total clicks. Finally, the bottom graph (Figure 3c) shows that the top 20% of rooms in terms of requests make up the entirety (100%) of all room requests.[4]

This suggests substantial congestion as the universe of clicks and requests are highly concentrated on around 20% of the rooms. At the same time, 80% of the rooms never receive a single request, which raises the possibility of allocative gains from reducing congestion. Since a room can be rented to at most one user, all else equal, renters are better off requesting rooms that receive fewer requests from other renters. These are topics which will guide our subsequent analysis.

Finally, Figure 4 illustrates users' price sensitivity. We plot CDFs of room prices across all search results (blue), search results that are clicked on (red), and search results that are requested (orange). We see that the CDF of prices across all search results stochastically dominates the other two; users tend to click and request cheaper rooms. The figure

---

[3]This is reminiscent of Ursu (2018), who finds a similar pattern analyzing data of Expedia.

[4]The dataset used in the analysis contains a random sample of 10% of the users in the market for rooms in Barcelona. We specify in detail how we construct the sample in Section 2.3.
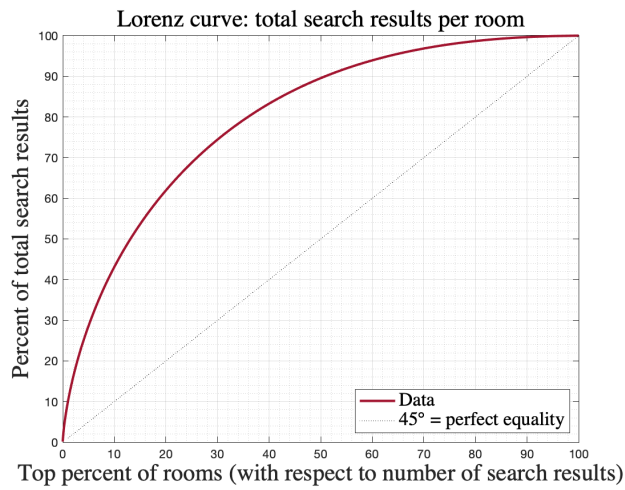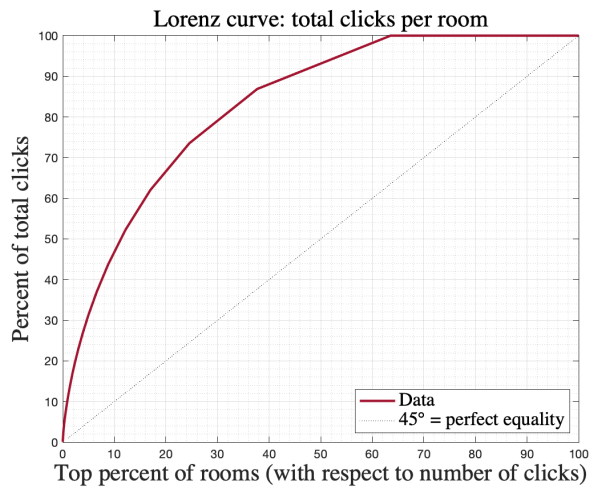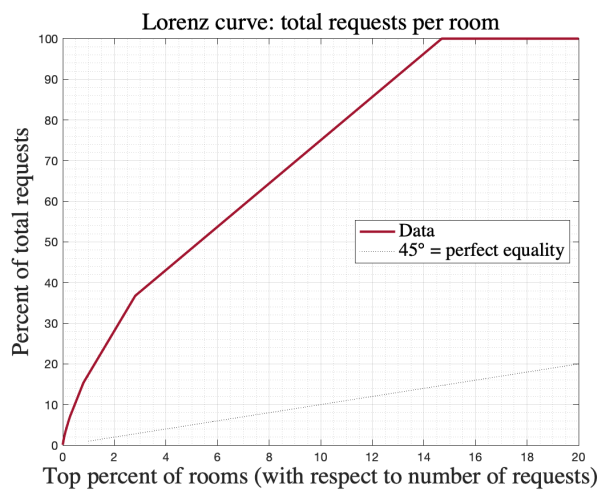
Figure (a) Searches



Figure (b) Clicks



Figure (c) Requests

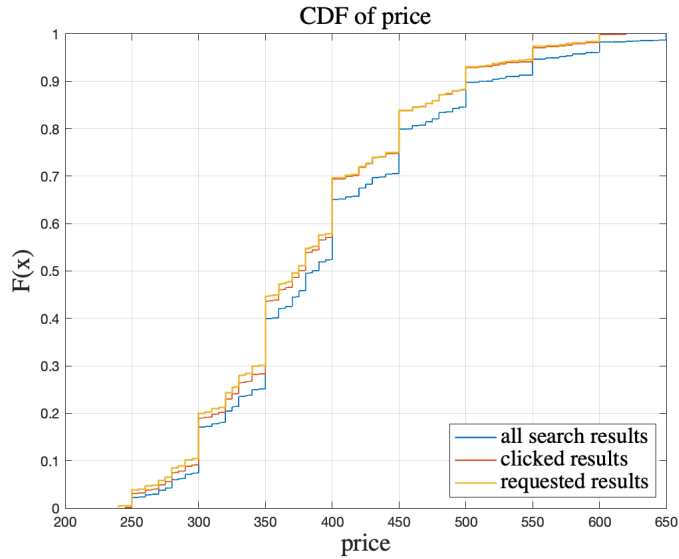Figure 3: Concentration of searches, clicks and requests

Figure 4: Price sensitivity during Customer Search Journey

also shows, albeit in a lower magnitude, that the distribution of prices across clicked rooms stochastically dominates that of rooms that are requested. In other words, as the consumer search journey progresses, the rooms users browse, click on, and eventually request, become cheaper. As we would expect, this shows that users are price sensitive, and may even update their evaluations of rooms after clicking on them and seeing their listings in more detail.

## 2  Model

Section 2 introduces the setup and notation which underlies our analysis. Users (or renters) are indexed by $i \in N = \{1, \ldots, n\}$. Each user is associated to a vector of observable characteristics $Z_i$. User $i$ conducts searches $S \in \mathcal{S}_i$, where $|\mathcal{S}_i| = n_i$ is the number of searches conducted by user $i$. A search $S$ is a collection of search results $s \in S = \{1, \ldots, n_S\}$, where $n_S$ is the number of search results in search $S$.

Each search result $s$ is associated to a vector of observable characteristics, $X_s = (X_{1s}, X_{2s}, pos_s)$, where $X_{1s}$ are the characteristics users observe in the page of search results, such as the price and the location of a room, $X_{2s}$ the ones they observe only upon clicking on the search result, such as the amenities of a room, and $pos_s$ is the position the room is shown

on the page of search results.[5] The decision variables observed in the data are:

$$k_{is} = 1\{\text{user } i \text{ clicks on search result } s\} \tag{1}$$

$$r_{is} = 1\{\text{user } i \text{ requests search result } s\} \tag{2}$$

We omit the dependence of $k_{is}$ and $r_{is}$ on $S \in \mathcal{S}_i$ for simplicity. Since to request a room one needs to click on it first, we have $r_{is} = 1 \Rightarrow k_{is} = 1$. Finally, we index rooms by $j \in J$, where $j_s$ denotes the room associated to search result $s$.

## 2.1   Users' clicks and room requests: A multinomial choice framework

Next we describe the framework we use to model users' observed click and room requests. The utility of requesting search result $s$ in a generic search $S$ made by user $i \in N$ is given by:

$$U_{is} = X_{1s}\beta_1 + X_{2s}\beta_2 + (X_{2s} \otimes Z_i)\beta_{XZ} + \varepsilon_{is}^r, \tag{3}$$

where $\varepsilon_S^r = (\varepsilon_{is}^r)_{s \in S}$ is an iid vector of extreme-value type 1 random utility shocks and $X_{2s} \otimes Z_i$ is a shorthand for interactions between $X_{2s}$ and $Z_i$ (which may not be exhaustive). For example, $X_{2s} \otimes Z_i$ includes whether user $i$ has the gender, age, or occupation preferred by the landlord of the listing in search result $s$.

The request data reveals a partial ordering of the utility of rooms conditional on being clicked, where rooms that are requested have a utility higher than those that are clicked on but not requested. For example, suppose that a user clicks on four rooms (search results) in a search, 1, 3, 4, and 7, with utilities $U_1$, $U_3$, $U_4$, and $U_7$, respectively. If they request, say, rooms 1 and 4, the request data reveals $\min\{U_1, U_4\} > \max\{U_3, U_7\}$. Essentially, in the estimation we allow the users to choose more than one item among their choice sets, and impose only that users obtain higher utility from the chosen options than the non-chosen options, while being agnostic about users' ranking among the chosen (and non-chosen) options. Given the extreme-value specification of the utility shocks, our model takes the form of a rank-ordered logit model with ties (Beggs, Cardell, and Hausman, 1981; Allison and Christakis, 1994).[6]

The propensity to click on a room depends on the position the room is shown in as a search result and on its expected utility conditional on the information the user has before clicking. Formally, the propensity to click on a search result $s$ in a generic search

---

[5]Specifically, a search $S$ is a collection of pages, each containing several search results. However, for simplicity we abstract from the page a search result is shown in, and just take into account the position it is shown in within the corresponding page. This should not have an impact on our counterfactual analysis if we assume that the changes in the ranking algorithm do not change the number of pages explored by users.

[6]We have also estimated semiparameteric versions of this model (Fox, 2007; Yan and Yoo, 2017), and our findings are qualitatively similar.

$S$ of user $i \in N$ is determined by:

$$I_{is} = g(pos_{is})\beta_{pos} + \mathbb{E}\left[U_{is} \mid X_{1s}, pos_{is}\right]\beta_U + pos_{is} \times \mathbb{E}\left[U_{is} \mid X_{1s}, pos_{is}\right]\beta_{pos \times U} + \varepsilon_{is}^k, \quad (4)$$

where $\varepsilon_S^k = (\varepsilon_{is}^k)_{s \in S}$ is an iid vector of extreme-value type 1 random shocks, $g(\cdot)$ is a known function (which we specify as a squared polynomial with point-masses at the first, second, and third positions), and $\mathbb{E}\left[U_{is} \mid X_{1s}, pos_{is}\right]$ is the room's expected utility conditional on the available information. Namely,

$$\mathbb{E}\left[U_{is} \mid X_{1s}, pos_{is}\right] = X_{1s}\beta_1 + \mathbb{E}\left[X_{2s} \mid X_{1s}, pos_{is}\right]\beta_2 + \mathbb{E}\left[X_{2s} \otimes Z_i \mid X_{1s}, pos_{is}\right]\beta_{XZ}. \tag{5}$$

As with requests, our model for clicks has the structure of a rank-ordered logit with ties.

## 2.2 Estimation

There are two stages in users' choice behavior. First, a user *clicks* a room from her search results. Second, after learning more about its features, she may decide to *request* the room by contacting the landlord. Analogously, there are two sequential stages in our estimation procedure, corresponding to users' click and request decisions. Proceeding backwards, we start by describing the estimation of the model for the request decision, and then describe how this feeds into our estimation of the earlier click decision.

In the first step, we estimate the parameters in the utility equation (3) via maximum likelihood: $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_{XZ}$. After estimating these parameters, we project the covariates in $X_{2s}$ onto $X_{1s}$ and $pos_{is}$ to estimate the expected requested utility. The estimated expected request utility is given by:

$$\hat{U}_i(X_s) = X_{1s}\hat{\beta}_1 + \hat{X}_{2s}\hat{\beta}_2 + (\hat{X}_{2s} \otimes Z_i)\hat{\beta}_{XZ}, \tag{6}$$

where $\hat{X}_{2s}$ are the fitted values of linear projections (in which we include numerous interactions to provide enough flexibility). Equation (6) is the estimated version of (5).

Next we proceed to estimate the click decision model, where we use the estimated expected request utility $\hat{U}_i(X_s)$ as a covariate and interact it with $pos_{is}$ (see equation (4)). The parameters in the click equation (4) are readily estimated via maximum likelihood: $\hat{\beta}_{pos}$, $\hat{\beta}_U$, and $\hat{\beta}_{pos \times U}$.

## 2.3 Empirical specification and summary statistics

In the empirical specification, we include the following variables:

- $X_{1s}$ = price, number of tenants in room (may be missing in which case we include a dummy), and location (we split Barcelona into six districts);

- $X_{2s}$ = room amenities (AC, Balcony, Dishwasher, Doorman, Elevator, Exterior View, Heating, Smoker Friendly, TV, Terrace), number of days the room has been published in the platform, and dummies indicating whether the user has the age, gender, and occupation (student or worker) preferred by the landlord.

Our dataset consists of a random sample comprising 10% of all the users who used the platform in a two-year period (January 2018 to February 2020). For each of these users, we observe the complete sequence of searches, clicks, and requests. There exists vast variation across users and searches in the data, with many users searching all over the globe and others searching without submitting any requests. In order to include in our analysis only the relevant choices made by users who are likely to be considering renting an apartment in Barcelona, we restrict our sample as follows. First, we say that a search is in Barcelona if at least 75% of its search results are in Barcelona, where a search result is said to be in Barcelona if the room's location is at most 30km away from Barcelona's city center. Then, we include in the estimation sample the searches in Barcelona made by users who (i) predominantly search in Barcelona (at least 75% of their searches are in Barcelona), and (ii) made at least one request to a room in Barcelona during the sample period.

Table 1 provides the size of our estimation sample and descriptive statistics of the users, rooms, searches, clicks, and requests in it. The sample contains 1,202 users. On average, users are 29 years old; 51% of them are females, 35% are students, and 81% are employed. There are 45,462 rooms, of which 80% specify they prefer renters within a specific age bracket (on average, they prefer renters that are between 20 and 36 years old). 27% of rooms also specify a preferred gender for renters, of which 23% prefer female renters and 4% male renters. Finally, 25% of rooms specify a preferred occupation for renters, of which 22% prefer *not* renting to students, and 3% prefer to rent to students *only*.

In terms of congestion, while all rooms appear at least once as a search result in the sample, only 64% of them receive at least one click, and 15% receive at least one request (see Figure 3 for a full picture of the concentration of searches, clicks, and requests in the sample). On average, each user conducts 144 searches and browses over 1,700 search results. Searches contain an average of 12 search results, with a maximum of 20. This adds up to a total of over 2 million search results. On average, each user clicks on 75 rooms and submits requests to 7 rooms.

Table 2 reports summary statistics of room characteristics. The left panel (first five columns) reports summary statistics across all search results in the sample.[7] The right panel (last three columns) reports mean differences across: (a) search results that are clicked on and all search results (clicked – all); (b) search results that are requested and

---

[7]We provide characteristics at the level of search results since this is the unit of observation in the analysis.

all search results (requested – all), and (c) search results that are requested and clicked on (requested – clicked). Analyzing these differences in means offers a quick insight into the presence of variation across the rooms that renters browse, click on, and ultimately request. Our model is designed to exploit this variation to estimate users' preferences for rooms.

On average, rooms displayed as search results are around €400 per month. Compared to all displayed rooms, on average, the ones that are clicked on are €12 cheaper, and the ones that are requested are €15 cheaper. We also see that renters tend to click on rooms that have been published for more days in the platform, but do not tend to request them. Renters tend to browse rooms for which their personal characteristics match the landlord's preferences: 81% of search results have the preferred gender of the user (or none), 65% their preferred age (or none), and 90% their preferred occupation (or none). Unsurprisingly, renters are more likely to click on and request listings in which their gender, age, and occupation match the landlord's preferences. There is rich variation in the data across room amenities: doorman, dishwasher, terrace, etc. In general, renters tend to click on and request rooms that have more amenities. In terms of location, the sample is relatively balanced across the six regions of the city we consider.[8] Perhaps unsurprisingly, renters seem to be more likely to click and request rooms near the city center, in Ciutat Vella, Saint Martí, or L'Eixample.

---

[8]The city of Barcelona proper has 10 districts, which we group into five by merging together Sarrià-Sant Gervasi and Gràcia, Ciutat Vella and Saint Martí, Sants-Montjuic and Les Corts, and Horta-Guinardó, Nou Barris and Sant Andreu (into North Barcelona). Additionally, we pool together rooms that are within a 30km radius from the city center but outside the 10 districts, into "Greater Barcelona."

Table 1: Sample description

| | |
|---|---|
| Number of users (renters) | 1,202 |
|     Average age | 29.44 |
|     Are females (%) | 51.4 |
|     Are students (%) | 34.61 |
|     Are employed (%) | 81.03 |
| | |
| Number of rooms | 45,462 |
|     Specify minimum tenant age (%) | 80.27 |
|       Average minimum tenant age | 20.57 |
|     Specify maximum tenant age (%) | 80.27 |
|       Average maximum tenant age | 36.06 |
|     Specify a preferred tenant gender (%) | 27.2 |
|       Prefer female tenants (%) | 23.13 |
|       Prefer male tenants (%) | 4.07 |
|     Specify a preferred tenant occupation (%) | 25.4 |
|       Prefer no student tenants (%) | 22.7 |
|       Prefer not employed tenants (%) | 2.68 |
|     Appear in at least one search result (%) | 100 |
|     Are clicked on at least one (%) | 63.56 |
|     Are requested at least one (%) | 14.69 |
| | |
| Total number of searches | 173,223 |
|     Average per user | 144.11 |
| | |
| Total number of search results | 2,066,147 |
|     Average per user | 1,718.92 |
|     Average per room | 45.45 |
|     Average per search | 11.93 |
| | |
| Total number of clicks | 89,624 |
|     Average per user | 74.56 |
|     Average per room | 1.97 |
|     Average per room $\mid \geq 1$ | 3.1 |
|     Average per search | 0.52 |
|     Prob. search result is clicked | 4.3% |
| | |
| Total number of requests | 8,542 |
|     Average per user | 7.11 |
|     Average per room | 0.19 |
|     Average per room $\mid \geq 1$ | 1.28 |
|     Average per search | 0.049 |
|     Prob. click is requested | 9.5% |
|     Prob. search result is requested | 0.41% |

*Notes:* Table reports sample sizes and descriptive statistics of the users (renters), rooms, searches, search results, clicks, and requests in the estimation sample.

Table 2: Summary statistics and mean differences

| | mean | sd | median | min | max | clicked − all | requested − all | requested − clicked |
|---|---|---|---|---|---|---|---|---|
| Price | 400.591 | 102.643 | 390 | 0 | 750 | -12.377*** | -15.288*** | -3.740*** |
| Missing number of tenants | 0.132 | 0.339 | 0 | 0 | 1 | -0.009*** | -0.005 | 0.004 |
| Number of tenants | 1.898 | 1.406 | 2 | 0 | 20 | 0.018*** | 0.047*** | 0.032** |
| Days since first published | 100.669 | 166.695 | 14 | 0 | 744 | 5.648*** | -5.017*** | -11.495*** |
| | | | | | | | | |
| User has gender preferred by landlord | 0.812 | 0.391 | 1 | 0 | 1 | 0.008*** | 0.034*** | 0.029*** |
| User has age preferred by landlord | 0.654 | 0.476 | 1 | 0 | 1 | 0.006*** | 0.026*** | 0.022*** |
| User has occupation preferred by landlord | 0.902 | 0.297 | 1 | 0 | 1 | 0.006*** | 0.013*** | 0.008** |
| | | | | | | | | |
| Has doorman | 0.142 | 0.349 | 0 | 0 | 1 | 0.003*** | -0.004 | -0.008* |
| Has dishwasher | 0.187 | 0.390 | 0 | 0 | 1 | 0.003** | -0.001 | -0.004 |
| Has terrace | 0.214 | 0.410 | 0 | 0 | 1 | 0.006*** | 0.001 | -0.005 |
| Has AC | 0.241 | 0.428 | 0 | 0 | 1 | -0.004*** | -0.017*** | -0.014*** |
| Is smoker friendly | 0.278 | 0.448 | 0 | 0 | 1 | 0.002 | 0.011** | 0.010** |
| Has elevator | 0.653 | 0.476 | 1 | 0 | 1 | -0.001 | 0.004 | 0.006 |
| Has exterior view | 0.492 | 0.500 | 0 | 0 | 1 | 0.014*** | 0.022*** | 0.009 |
| Has TV | 0.614 | 0.487 | 1 | 0 | 1 | 0.003* | -0.002 | -0.005 |
| Has balcony | 0.417 | 0.493 | 0 | 0 | 1 | 0.015*** | 0.040*** | 0.027*** |
| Has heating | 0.470 | 0.499 | 0 | 0 | 1 | 0.004** | 0.011** | 0.008 |
| | | | | | | | | |
| In Sarrià Saint-Gervasi or Gràcia | 0.130 | 0.336 | 0 | 0 | 1 | 0.007*** | 0.005 | -0.002 |
| In Ciutat Vella or Saint Martí | 0.198 | 0.399 | 0 | 0 | 1 | 0.003** | 0.043*** | 0.044*** |
| In L'Eixample | 0.263 | 0.440 | 0 | 0 | 1 | 0.016*** | 0.036*** | 0.022*** |
| In North of Barcelona | 0.124 | 0.329 | 0 | 0 | 1 | -0.022*** | -0.043*** | -0.025*** |
| In Sants-Montjuic or Les Corts | 0.160 | 0.366 | 0 | 0 | 1 | -0.008*** | -0.024*** | -0.018*** |
| In Greater Barcelona | 0.125 | 0.330 | 0 | 0 | 1 | 0.003** | -0.017*** | -0.022*** |

*Notes:* the table reports summary statistics of the variables used in the model. The first five columns report the mean, standard deviation, median, minimum, and maximum across all search results in the sample. The last three columns report mean differences between (a) clicked and all search results; (b) requested and all search results; (c) requested and clicked search results. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Variables *Price* and *Days since first published* are winsorized on the right at the 99th percentile. See Table 1 for sample sizes and additional descriptive statistics.

# 3 Estimation Results

In Section 3, we report and discuss the results of our estimation exercise. Table 3 and Table 4 report the estimated parameters of the request utility and the click propensity, respectively.

First, we discuss the parameter estimates of the request utility. As Table 3 shows, we estimate a negative and statistically significant price sensitivity across multiple specifications. The coefficients are only identified up to scale in our model. To interpret their magnitudes it is convenient to rescale them by dividing our estimates by the absolute value of the price coefficient, which allows us to interpret their magnitudes as willingness-to-pay in euros. For example, once normalized, the coefficient in the last specification for days since first published equals $-0.26$, which means that being listed for an additional day in the platform has an effect on the probability that a user submits a request to a room after having clicked on it equivalent to the room being €0.26 more expensive. While statistically significant different from zero, the magnitude of this coefficient does not point to a large economic significance. Table 5 in Section A.1 reports the normalized values of all coefficients. Seen through this lense, we observe that the magnitudes of the coefficients seem reasonable and economically significant, ranging from values close to zero up to a few hundred euros. For example, having the gender preferred by the landlord is valued positively at €521, having a doorman is valued negatively at €121, and having a balcony and heating are valued positively at €113 and €102, respectively.

In terms of statistical significance, few room characteristics appear to have a significant impact on consumers' preferences for rooms. Given the large sample size, this may be driven by limited variation in the rooms that users click on. However, as we shall see below, for counterfactuals, we focus on the overall predicted utility of rooms rather than on specific coefficients.

Next we discuss our parameter estimates of the click propensity. As Table 4 shows, we estimate a positive effect of the expected utility of a room, conditional on the available information, on the propensity to click on a room. While not statistically significant, this reflects that users are somewhat more likely than not to click on rooms they are more likely to request. The estimated effect of the position a room is shown as a search result on the propensity that a user clicks on it is negative and statistically significant. Users are more likely to click on rooms shown at the top. Moreover, the difference in click propensity from being shown at a higher position is higher for rooms with a higher utility, which amplifies this effect for more preferred rooms.[9] Finally, being shown at the first position has an additional positive and statistically significant effect on the click propensity. As with our estimates for the request utility, in order to assess

---

[9]Please bare in mind that "being shown at a higher position" is equivalent to a *lower* value for the variable *Position*: the top position has $Position = 1$, etc.

Table 3: Parameter estimates of request utility

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Price | -0.0007** | -0.0007** | -0.0008*** | -0.0009*** |
| | (0.0003) | (0.0003) | (0.0003) | (0.0003) |
| Missing number of tenants | | -0.0899 | -0.0786 | -0.0795 |
| | | (0.0803) | (0.0808) | (0.0810) |
| Number of tenants | | 0.0004 | 0.0009 | -0.0037 |
| | | (0.0160) | (0.0162) | (0.0162) |
| Days since first published | | -0.0003** | -0.0003** | -0.0002** |
| | | (0.0001) | (0.0001) | (0.0001) |
| User has gender preferred by landlord | | 0.4865*** | 0.4906*** | 0.4818*** |
| | | (0.0571) | (0.0572) | (0.0574) |
| User has age preferred by landlord | | -0.0068 | 0.0027 | 0.0095 |
| | | (0.0536) | (0.0538) | (0.0539) |
| User has occupation preferred by landlord | | 0.0771 | 0.0769 | 0.0799 |
| | | (0.0851) | (0.0853) | (0.0855) |
| Has doorman | | | -0.0924* | -0.1119** |
| | | | (0.0551) | (0.0554) |
| Has dishwasher | | | 0.0728 | 0.0729 |
| | | | (0.0492) | (0.0493) |
| Has terrace | | | -0.0191 | -0.0136 |
| | | | (0.0449) | (0.0450) |
| Has AC | | | 0.0082 | 0.0149 |
| | | | (0.0471) | (0.0473) |
| Is smoker friendly | | | 0.0017 | -0.0019 |
| | | | (0.0403) | (0.0404) |
| Has elevator | | | 0.0529 | 0.0416 |
| | | | (0.0413) | (0.0419) |
| Has exterior view | | | -0.0260 | -0.0161 |
| | | | (0.0399) | (0.0401) |
| Has TV | | | -0.0115 | -0.0125 |
| | | | (0.0374) | (0.0374) |
| Has balcony | | | 0.1149*** | 0.1049*** |
| | | | (0.0381) | (0.0382) |
| Has heating | | | 0.0892** | 0.0945** |
| | | | (0.0395) | (0.0397) |
| In Sarrià Saint-Gervasi or Gràcia | | | | -0.0710 |
| | | | | (0.1956) |
| In Ciutat Vella or Saint Martí | | | | -0.1014 |
| | | | | (0.1916) |
| In L'Eixample | | | | -0.0574 |
| | | | | (0.1905) |
| In North of Barcelona | | | | -0.5561*** |
| | | | | (0.1947) |
| In Sants-Montjuic or Les Corts | | | | -0.1696 |
| | | | | (0.1869) |
| Num Searches | 52,737 | 52,737 | 52,737 | 52,737 |
| Num Results | 89,624 | 89,624 | 89,624 | 89,624 |
| Pseudo-R2 | 0.0005 | 0.0087 | 0.0111 | 0.0150 |
| Log-likelihood | -5112.57 | -5070.45 | -5058.32 | -5038.15 |

*Notes:* Table reports rank-ordered logit parameter estimates. Estimation via maximum likelihood. Standard errors in parentheses; $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The category "In Greater Barcelona" is omitted.

Table 4: Parameter estimates of click propensity

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Position | -0.0538*** | -0.0834*** | -0.0698*** | -0.0698*** |
|  | (0.0010) | (0.0029) | (0.0052) | (0.0119) |
| Position$^2$ |  | 0.0016*** | 0.0010*** | 0.0010*** |
|  |  | (0.0002) | (0.0002) | (0.0002) |
| 1(Position = 1) |  |  | 0.0991*** | 0.0987*** |
|  |  |  | (0.0216) | (0.0217) |
| 1(Position = 2) |  |  | -0.0014 | -0.0010 |
|  |  |  | (0.0183) | (0.0183) |
| 1(Position = 3) |  |  | -0.0028 | -0.0027 |
|  |  |  | (0.0155) | (0.0155) |
| $\mathbb{E}[u|I]$ |  |  |  | 0.3027 |
|  |  |  |  | (0.2806) |
| $\mathbb{E}[u|I] \times$ Position |  |  |  | 0.0320 |
|  |  |  |  | (0.0217) |
| Num Searches | 173,223 | 173,223 | 173,223 | 173,223 |
| Num Results | 2,066,147 | 2,066,147 | 2,066,147 | 2,066,147 |
| Pseudo-R2 | 0.0143 | 0.0147 | 0.0149 | 0.0161 |
| Log-likelihood | -1.644e+05 | -1.644e+05 | -1.643e+05 | -1.641e+05 |

*Notes:* Table reports rank-ordered logit parameter estimates. Estimation via maximum likelihood. Bootstrapped standard errors in parentheses account for variation in the first-stage estimation (1,000 bootstrap draws); *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

the magnitudes of these coefficients, we can normalize them for them to be measured in euros. Table 6 in Section A.1 reports the normalized values of the coefficients in Table 4.[10] Accordingly, the negative effect that being shown in a lower position has on the click propensity can be valued at around €200–250 per position. And being shown at the top position has an additional positive effect equivalent to €352.

As we mentioned above, we use the estimated utility and click propensity indices in the counterfactuals below. Figure 5 plots these two indices against the position in which rooms are shown.[11] The top graph (Figure 5a) shows that the propensity to click on a room decreases monotonically with the position a room is shown. And this holds true in all rooms in the data, regardless of whether they are clicked on or requested. The bottom graph (Figure 5b) shows that the predicted request utility of rooms is relatively flat and does not depend heavily on the position in which they are shown. This can be seen especially in the predicted utility of the rooms that are requested in the data (green bars in Figure 5b). Altogether this shows that our estimates of the click propensity and request utility echo the raw data (cf. Figure 2): whereas users are more likely to click

---

[10]Normalizing the coefficients of the click propensity is done by dividing them by the absolute value of the utility coefficient multiplied by the absolute value of the price coefficient in the request utility. See that this is necessary in order to obtain the marginal effect of a change in the expected utility when it is measured in euros.

[11]In the plots, we normalize the location of the indices, so that the room at the 25th percentile has a zero index, both for the click propensity and the request utility.

at the top, once they have clicked on a room, the position a room is shown in has no effect on whether the user requests that room.
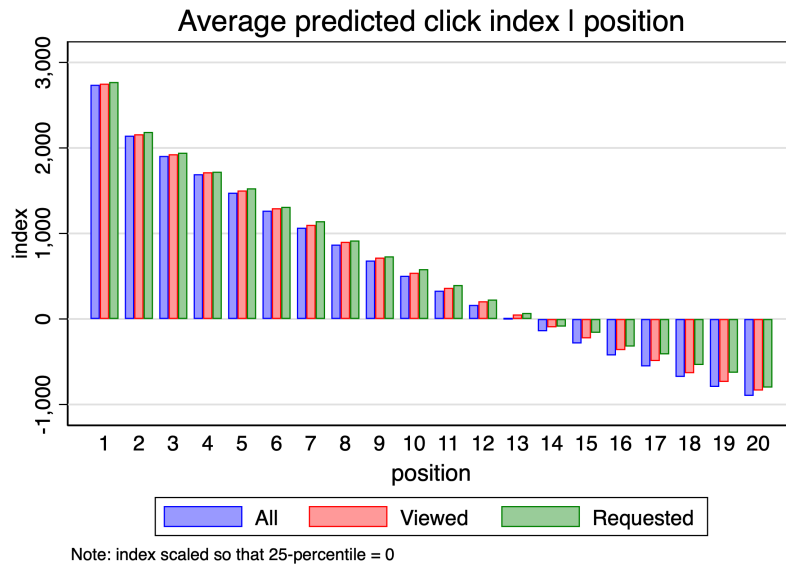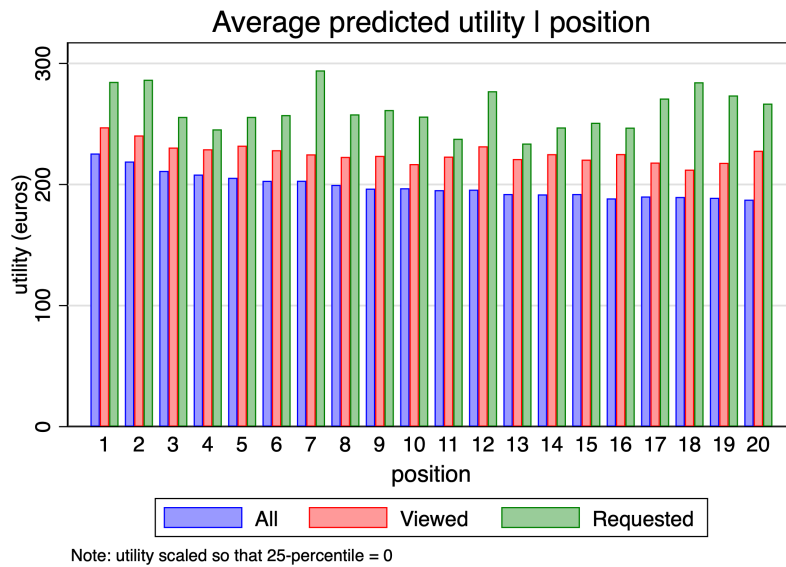


Figure (a) Click index



Figure (b) Request Utility

Figure 5: Estimated click index and request utility

## 4    Counterfactuals

Based on the estimation results above, we simulate a series of counterfactuals in which we consider the effect of changing the search ranking algorithm on the click and request

outcomes. The counterfactual exercises below consider the following question: What would the users have clicked on and requested had they observed the same set of search results but ranked in a different order? Throughout we assume that changing the ranking algorithm would not affect the users' unmodeled search behavior: the number of searches they conduct, clicks they make, and requests they submit.

Before presenting the results, we note that in these simulations we do not incorporate the utility shocks when predicting outcomes; we only make use of the estimated indices to predict clicks and requests. This approach is more robust because it does not rely on the parametric assumptions on the utility shocks, but solely on the heterogeneity present in the room and user characteristics. In this sense, our approach is conservative at capturing horizontal differentiation in preferences.

## 4.1 Different ranking algorithms: effects of congestion

In the counterfactuals we explore alternative ranking algorithms and compare their performance to the platform's current ranking system, which surfaces the same ordering of search results to every user. To compare congestion across counterfactual scenarios, we again use Lorenz curves, as in Figure 3 above. Figure 6 presents Lorenz curves across a number of scenarios, which we discuss in sequence. The top two graphs in Figure 6 show the Lorenz curves corresponding to the data and predicted values derived from simulations of the model at the estimated model parameters. The close relation between the data (red curve) and predicted values (blue curve) confirm that the estimated model fits the data well, both in terms of the clicks (Figure 6a) and requests (Figure 6b).

In the bottom two graphs of Figure 6, we consider two counterfactual scenarios that incorporate our parameter estimates of user preferences, and change the order in which search results are shown to users (keeping the selection of results within each search fixed).

First, we consider the case of **full personalization** in which the search results within each search are ordered according to each user's preferences. Since users tend to click on rooms shown at top positions, this algorithm steers users towards the rooms they prefer the most. The Lorenz curve from this full personalization scenario is graphed in green. Second, we consider the case of **random order** in which the order in which rooms are shown is fully randomized within every search. This algorithm destroys the correlation between the position at which rooms are shown across searches and users, and is aimed at reducing congestion. This scenario is shown in the orange-colored Lorenz curves.

The bottom two graphs show that these counterfactual ranking algorithms do change the concentration of clicks (Figure 6c) and requests (Figure 6d). The direction of the change in clicks and requests is qualitatively similar: randomizing the order in which rooms are shown (orange line) decreases congestion, if only slightly, while the fully personalized
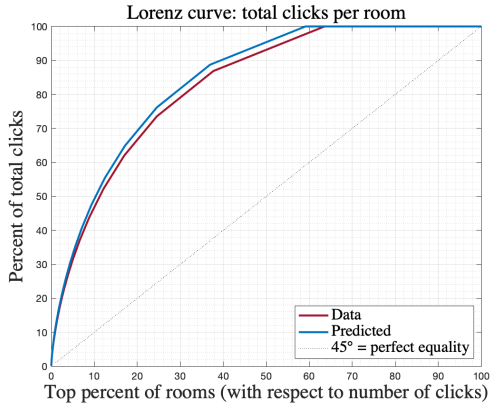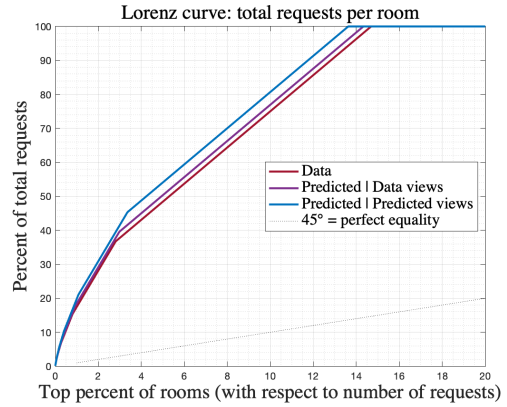
20

Figure (a) Predicted clicks
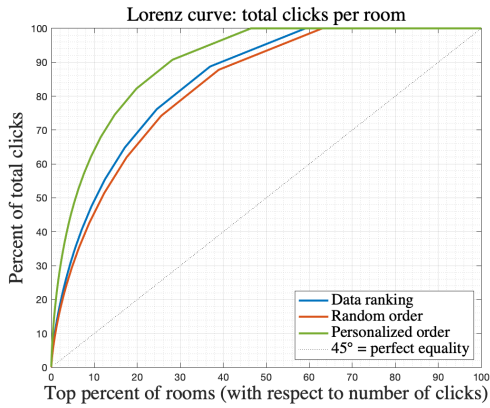
Figure (b) Predicted requests
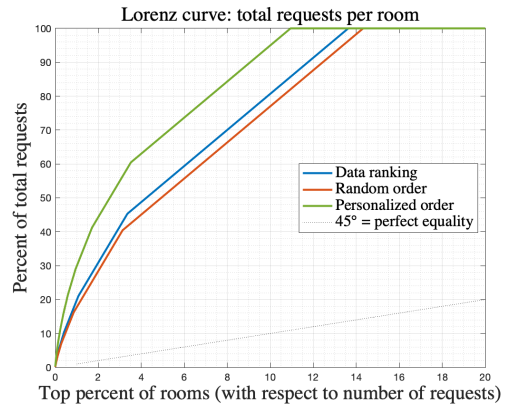
Figure (c) Counterfactual clicks

Figure (d) Counterfactual requests

Figure 6: Lorenz curves of the distribution across rooms of clicks and requests

one (green) increases congestion substantially. That is, using the information on user preferences to rank search results within searches leads to higher congestion compared to the data.

This suggests that across users, the rooms are primarily vertically differentiated, so that most users' preferences agree on which rooms offer the highest utility. This echoes the findings of Chen et al. (2023), who likewise found that surfacing dating prospects to users in order of their reported preferences results in excess congestion. From a congestion perspective, then, the homogeneous ranking algorithm currently used by the platform emerges as a preferable option to a fully personalized one.

## 4.2 Utility and congestion trade-off

The results above illustrate how congestion in this marketplace can be reduced by randomizing search results. However, random search results imposes a potential cost of lowering the utility that users receive from the rooms that they click on and request.

At the same time, personalizing search results to users by showing them rooms in the order of their preferences can optimize utility, but result in greater congestion.

This suggests a key trade-off in this market, between congestion and utility. To examine this trade-off more formally, we introduce a continuum of ranking algorithms indexed by the parameter $\alpha \in [0, 1]$, where $\alpha$ denotes the weight given to the "fully personalized" order, which ranks more preferred rooms at the top, and $1 - \alpha$ the weight given to the "random order," which ranks them randomly. In Figure 6 above, we only considered the polar cases of $\alpha = 1$ ("full personalization"), and $\alpha = 0$ ("random order"); here we expand to consider all intermediate alternatives.

The effects of incorporating information on user preferences to a random ranking algorithm, i.e., personalizing it, is illustrated in Figure 7. In the left two graphs, the $x$-axis plots $\alpha$ moving from 0 to 1. The blue curve shows how the average utility of the rooms requested by users increases with greater personalization, as $\alpha$ goes from 0 to 1. The red curve shows how congestion also increases with greater personalization. While personalizing the order in which rooms are shown to users steers them to rooms they prefer more, it leads to higher congestion. This pattern is present both in the distribution of clicks (Figure 7a) and requests (Figure 7c).

The utility and congestion levels in the data are indicated by black dots in the plots. In the left two graphs of Figure 7, one can see that the average utility in the data can be achieved by incorporating little information about user preferences to an otherwise random ranking algorithm, i.e., by setting a value for $\alpha$ lower than 0.1. Furthermore, the graphs also show that setting such value of $\alpha$ would result in lower congestion than what is observed in the data since the congestion observed in the data corresponds to higher values of $\alpha$, around 0.34 for clicks and 0.25 for requests.

The above results lead to two conclusions: (i) the slopes of the curves describing the change in average utility and congestion encode the trade-off between utility and congestion in this market, and (ii) the utility and congestion levels observed in the data seem to be inefficient, in that we can improve one without worsening the other. These conclusions can be appreciated in the right two graphs of Figure 7, which plot possibility frontiers in the (utility, congestion) space achievable by different values of $\alpha$. The black dot in these plots indicates the (utility, congestion) point observed in the data. Notably, the black dot lies *below* the frontier of both clicks (Figure 7b) and requests (Figure 7d), illustrating how the current ranking algorithm utilized by the platform is inefficient: holding congestion fixed, it is possible to improve utility by incorporating more personalization; holding utility fixed, it is possible to lower congestion by introducing more randomization.
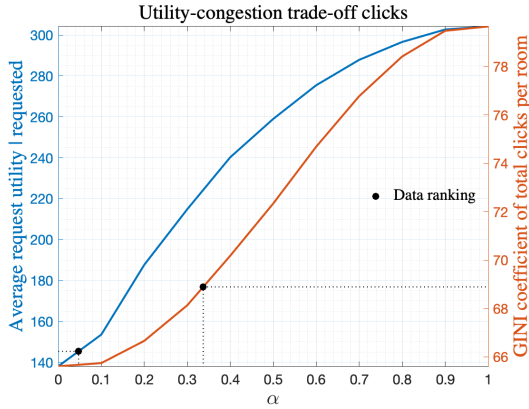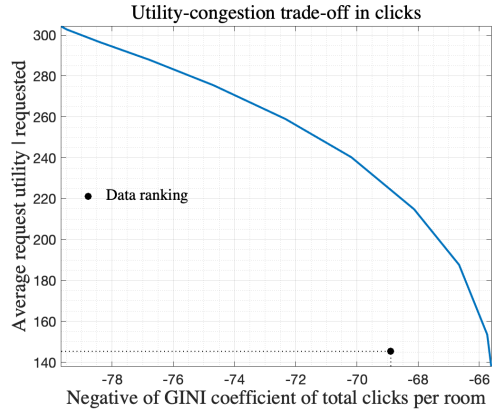
22

Figure (a) $\alpha$-plot clicks
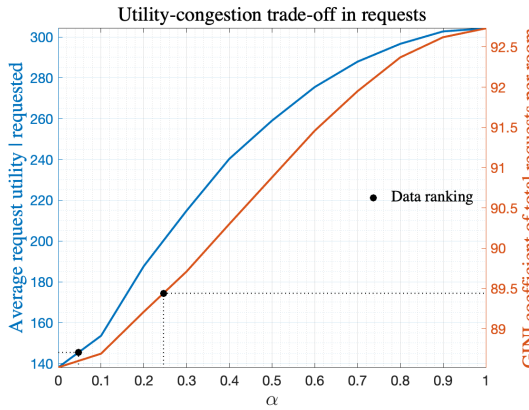
Figure (b) Efficiency frontier for clicks
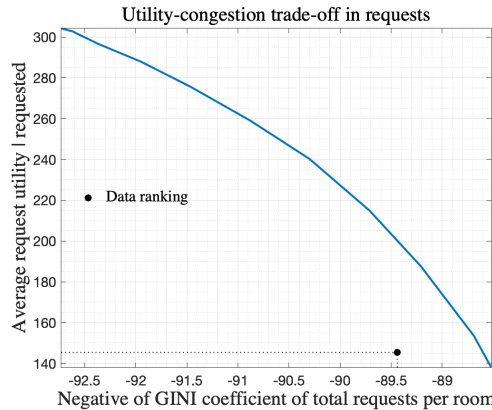
Figure (c) $\alpha$-plot requests

Figure (d) Efficiency frontier for requests

Figure 7: Utility and congestion trade-off of clicks and requests

## 4.3 Looking beyond the data: a hypothetical "garbling" scenario

The trade-off between utility and congestion depends on the degree of differentiation in consumer preferences. To emphasize this, we illustrate in Figure 8 a hypothetical scenario in which users' preferences are extremely horizontally differentiated. We achieve this by randomly relabeling the room IDs across searches. While this "garbling" keeps the utility of the alternatives in each search unchanged–including the utility of requested rooms–it destroys the correlation of preferences across users. In this case, we see that congestion no longer increases in $\alpha$, thereby eliminating the trade-off between utility and congestion (see the dotted red line in the left two graphs of Figure 8). The corresponding efficiency frontier approximates a vertical line (dotted blue line in right two graphs of Figure 8), indicating that at a fixed level of congestion, one can improve utility arbitrarily by personalizing search results to users.

Intuitively, when users' preferences are extremely horizontally differentiated, a "free lunch" emerges: high degrees of personalization, which lead to high user utility, do not
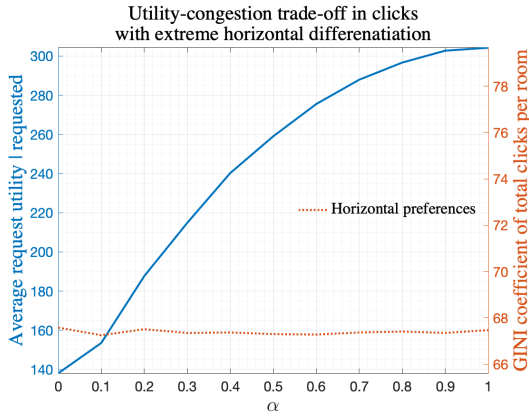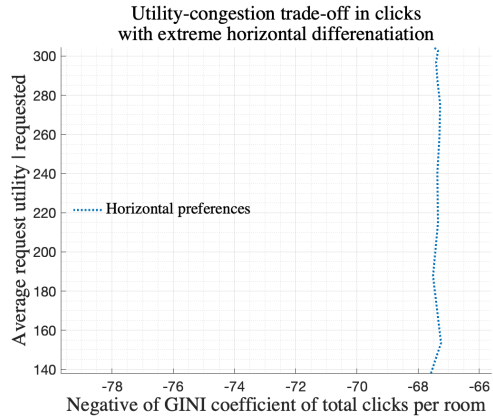
23

Figure (a) $\alpha$-plot clicks
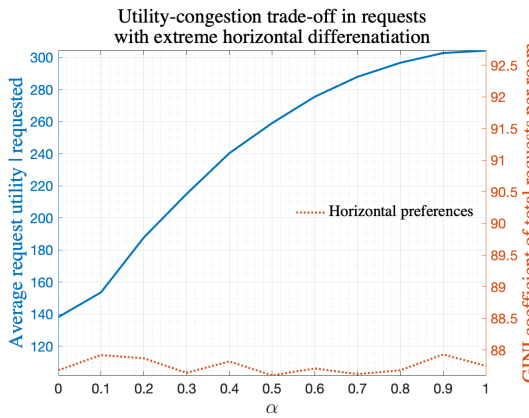


Figure (b) Efficiency frontier for clicks



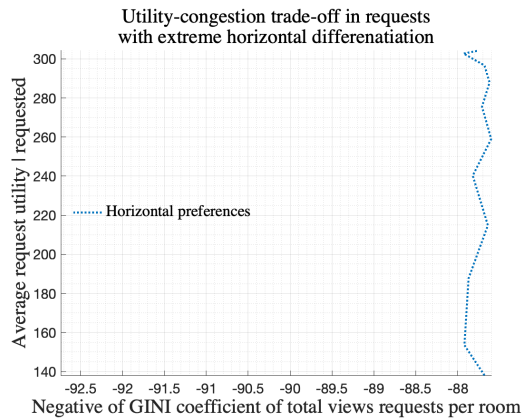Figure (c) $\alpha$-plot requests



Figure (d) Efficiency frontier for requests

Figure 8: Utility and congestion trade-off of clicks and requests
with extreme horizontal differentiation

increase congestion because user preferences are highly dissimilar. The nature of user preferences plays a central role in the direction of congestion. This exercise highlights that firms may gain from the ability to make user preferences more *horizontally differentiated*, perhaps by emphasizing features of products along which user preferences are more likely to disagree. We will study these issues in future work.

# 5    Conclusions

Two-sided platforms have the opportunity to reduce search costs of match seekers. However, unnecessary congestion may be generated if users are steered towards the same alternatives, for example, by showing alternatives in the same order to users. This may be a concern especially if the platform has limited information on user preferences. Using detailed data from a room-rental platform, we estimate preferences over rooms and demonstrate that, for this market, there is a key trade-off between congestion and utility,

in that ranking algorithms which reduce congestion must also reduce personalization by steering users to lower-utility rooms.

This arises from the fact that estimated preferences are primarily vertically differentiated, so that customers tend to agree on the most desirable features in the available rooms. Indeed, in hypothetical simulations, we show that if preferences exhibited sufficient horizontal differentiation, then the trade-off would reverse, leading to a "free lunch" whereby personalized algorithms which present customers their most desired rooms would not reduce congestion.

Clearly, platforms cannot change customers' preferences. However, they do have a role to play in designing their websites to emphasize different room features at different stages of the customer search journey. Can platforms effectively "horizontalize" their users' preferences by emphasizing the horizontal rather than vertical features of available rooms? These and related questions will be explored in future work.

# References

ALLISON, P. D. AND N. CHRISTAKIS (1994): "Logit models for sets of ranked items," in *Sociological Methodology*, ed. by P. V. Marsden, Oxford: Blackwell, vol. 24, 123–126. 10

ARNOSTI, N., R. JOHARI, AND Y. KANORIA (2021): "Managing congestion in matching markets," *Manufacturing & Service Operations Management*, 23, 620–636. 3

ASHLAGI, I., A. K. KRISHNASWAMY, R. MAKHIJANI, D. SABAN, AND K. SHIRAGUR (2022): "Assortment planning for two-sided sequential matching markets," *Operations Research*, 70, 2784–2803. 3

BAI, J., M. X. CHEN, J. LIU, X. MU, AND D. Y. XU (2022): "Stand Out from the Millions: Market Congestion and Information Friction on Global E-Commerce Platforms," *mimeo*. 3

BEGGS, S., S. CARDELL, AND J. HAUSMAN (1981): "Assessing the potential demand for electric cars," *Journal of Econometrics*, 17, 1–19. 10

CHEN, K.-M., Y.-W. HSIEH, AND M.-J. LIN (2023): "Reducing Recommendation Inequality via Two-Sided Matching: A Field Experiment of Online Dating," *International Economic Review*, n/a. 3, 21

CHEN, Y. AND S. YAO (2017): "Sequential search with refinement: Model and application with click-stream data," *Management Science*, 63, 4345–4365. 4

DANG, C. I., R. URSU, AND P. K. CHINTAGUNTA (2022): "Product Revisits and Consumer Consideration Sets," Available at SSRN: https://ssrn.com/abstract=3626451. 4

DE LOS SANTOS, B. AND S. KOULAYEV (2017): "Optimizing click-through in online rankings with endogenous search refinement," *Marketing Science*, 36, 542–564. 4

FARRONATO, C., A. FRADKIN, AND A. MACKAY (2023): "Self-Preferencing at Amazon: Evidence from Search Rankings," *AEA Papers and Proceedings*, 113, 239–43. 4

FOX, J. T. (2007): "Semiparametric estimation of multinomial discrete-choice models using a subset of choices," *The RAND Journal of Economics*, 38, 1002–1019. 10

FRADKIN, A. (2017): "Search, matching, and the role of digital marketplace design in enabling trade: Evidence from airbnb," Available at SSRN: https://ssrn.com/abstract=2939084. 3

GHOSE, A., P. G. IPEIROTIS, AND B. LI (2014): "Examining the impact of ranking on consumer behavior and search engine revenue," *Management Science*, 60, 1632–1654. 4

HODGSON, C. AND G. LEWIS (2020): "You Can Lead a Horse to Water: Spatial Learning and Path Dependence in Consumer Search," Cowles Foundation Discussion Papers 2246, Cowles Foundation for Research in Economics, Yale University. 4

HORTON, J. J. (2019): "Buyer uncertainty about seller capacity: Causes, consequences, and a partial solution," *Management Science*, 65, 3518–3540. 3

IMMORLICA, N., B. LUCIER, V. MANSHADI, AND A. WEI (2021): "Designing approximately optimal search on matching platforms," in *Proceedings of the 22nd ACM Conference on Economics and Computation*, 632–633. 3

KOULAYEV, S. (2014): "Search for differentiated products: identification and estimation," *The RAND Journal of Economics*, 45, 553–575. 4

MEKONNEN, T. (2023): "Random vs. Directed Search for Scarce Resources," *mimeo*. 4

ROTH, A. E. (2008): "What Have We Learned from Market Design?" *The Economic Journal*, 118, 285–310. 3

ROTH, A. E. AND X. XING (1997): "Turnaround Time and Bottlenecks in Market Clearing: Decentralized Matching in the Market for Clinical Psychologists," *Journal of Political Economy*, 105, 284–329. 3

URSU, R. M. (2018): "The Power of Rankings: Quantifying the Effect of Rankings on Online Consumer Search and Purchase Decisions," *Marketing Science*, 37, 530–552. 4, 7

YAN, J. AND H. I. YOO (2017): "Semiparametric Estimation of the Random Utility Model with Rank-Ordered Choice Data," Working Papers April 2017, Durham University Business School. 10

# A  Appendix

## A.1  Additional tables

Table 5 and Table 6 report the normalized coefficients of the request utility and click propensity in the main text. More specifically, the parameters in Table 5 correspond to the parameters in Table 3, and the parameters in Table 6 correspond to those in Table 4.

Table 5: Normalized parameter estimates of request utility

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Price | -1.00 | -1.00 | -1.00 | -1.00 |
| Missing number of tenants | | -130.93 | -96.72 | -85.94 |
| Number of tenants | | 0.53 | 1.13 | -4.02 |
| Days since first published | | -0.40 | -0.32 | -0.26 |
| User has gender preferred by landlord | | 708.38 | 603.54 | 520.57 |
| User has age preferred by landlord | | -9.85 | 3.38 | 10.28 |
| User has occupation preferred by landlord | | 112.34 | 94.65 | 86.36 |
| Has doorman | | | -113.70 | -120.95 |
| Has dishwasher | | | 89.54 | 78.72 |
| Has terrace | | | -23.48 | -14.73 |
| Has AC | | | 10.03 | 16.12 |
| Is smoker friendly | | | 2.10 | -2.08 |
| Has elevator | | | 65.10 | 44.97 |
| Has exterior view | | | -31.93 | -17.42 |
| Has TV | | | -14.12 | -13.54 |
| Has balcony | | | 141.31 | 113.34 |
| Has heating | | | 109.70 | 102.14 |
| In Saint-Gervasi or Gracia | | | | -76.72 |
| In Ciutat Vella or Saint Marti | | | | -109.53 |
| In Eixample | | | | -61.97 |
| In North of Barcelona | | | | -600.84 |
| In Sants-Montjuic of Les Corts | | | | -183.20 |

*Notes:* Coefficients normalized to be in euros (divided by absolute value of price coefficient). The category "In Greater Barcelona" is omitted.

Table 6: Normalized parameter estimates of click propensity

|  | (1) |
|---|---|
| Position | -249.14 |
| Position$^2$ | 3.74 |
| 1(Position = 1) | 352.42 |
| 1(Position = 2) | -3.70 |
| 1(Position = 3) | -9.68 |
| $\mathbb{E}[u\|I]$ | 1.00 |
| $\mathbb{E}[u\|I] \times$ Position | 114.38 |

*Notes:* Coefficients normalized to be in euros (divided by absolute value of the normalized utility coefficient).