The Economics of Recommender Systems: Evidence from a Field Experiment on MovieLens*

Guy Aridor† Duarte Gonçalves‡
Daniel Kluver§ Ruoyan Kong¶ Joseph Konstan‖

November 25, 2022

Abstract

We conduct a field experiment on a movie-recommendation platform to identify if and how recommendations affect consumption. We use within-consumer randomization at the good level and elicit beliefs about unconsumed goods to disentangle exposure from informational effects. We find recommendations increase consumption beyond its role in exposing goods to consumers. We provide support for an informational mechanism: recommendations affect consumers' beliefs, which in turn explain consumption. Recommendations reduce uncertainty about goods consumers are most uncertain about and induce information acquisition. Our results highlight the importance of recommender systems' informational role when considering policies targeting these systems in online marketplaces.

Keywords: Recommender Systems; Information Acquisition; Field Experiment.
JEL Classifications: D83, D47, D12, L15, M37.

*We thank Áureo de Paula, Yeon-Koo Che, Mark Dean, Brett Gordon, Ilya Morozov, Imran Rasul, Silvio Ravaioli, Shan Sikdar, Andrey Simonov, Jack Willis, Michael Woodford, and audiences at several seminars / conferences for helpful comments which greatly improved the paper. Daniel, Ruoyan, and Joseph are members of the GroupLens lab, which runs and maintains the MovieLens platform. This study was approved under both the Columbia IRB under protocol number AAAT2659 and the University of Minnesota IRB under protocol number STUDY00010960. The experimental design, the primary experimental hypotheses and analyses were pre-registered in the AEA RCT Registry as AEARCTR-0007545. This work was partially supported by NSF grant CNS-2016397.

†Department of Marketing, Northwestern Kellogg; guy.aridor@kellogg.northwestern.edu.
‡Department of Economics, University College London; duarte.goncalves@ucl.ac.uk.
§Department of Computer Science and Engineering, University of Minnesota - Twin Cities; kluve018@umn.edu.
¶Department of Computer Science and Engineering, University of Minnesota - Twin Cities; kong0135@umn.edu.
‖Department of Computer Science and Engineering, University of Minnesota - Twin Cities; konstan@umn.edu.

1. Introduction

Recommendation systems are nearly ubiquitous in the digital economy. They have a wide set of applications ranging from e-commerce to curation feeds on social media platforms to cultural goods on streaming platforms such as Spotify and Netflix, and to the articles served on news platforms. The ubiquity of these systems has raised a large number of economic and social questions about the influence that they have on consumption choices and online marketplaces, which has spurred a number of legislative initiatives targeting these systems.[1] This has motivated researchers to explore how recommender systems influence aggregate consumption choices (Fleder and Hosanagar, 2009), the antitrust implications of their power to steer consumption choices (Hagiu and Jullien, 2011; De Corniere and Taylor, 2019; Aridor and Gonçalves, 2022) and the consequences of the dynamic interaction between consumers and these systems (Chaney, Stewart and Engelhardt, 2018).

A key assumption that researchers need to make in order to speak about these broader issues is how recommendations influence consumer choices and there have been a variety of approaches taken to model this link. However, despite this being a crucial input to guide modeling choices and to assess the impact of policies targeting these systems, it has proved challenging to disentangle the mechanisms that drive the influence of such systems.

In this paper, we analyze data from a pre-registered field experiment we conducted on a movie-recommendation platform, MovieLens, in order to identify the causal effect these systems have on consumption, as well as the mechanisms that drive their influence. We find recommendations significantly drive up consumption beyond mere exposure effects. Furthermore, after establishing that higher expected quality and lower uncertainty about quality are associated with greater consumption probability, we provide causal evidence that recommendations impact beliefs and induce further information acquisition.

We propose a theoretical framework in which recommendation impacts consumption through two main channels: it forces consideration of a good (*exposure*) and provides information about its idiosyncratic quality (*information*). Consumers make a sequence of choices over a set of experience goods about whose idiosyncratic quality they are uncertain. As in many settings in which recommender systems are deployed, the feasible set is large and at any given point consumers may not consider all goods available. Recommendation of a good can then steer behavior by rendering it part of the consideration set, thereby naturally increasing the likelihood it is chosen Additionally, recommendation also provides information, a signal about the good's idiosyncratic

[1]The European Union and the United Kingdom have focused on legislation targeting algorithms (CMA, 2021) and artificial intelligence specifically (e.g. `https://artificialintelligenceact.eu`) with significant portions of this legislation directly targeting recommender systems (e.g. see Schwemer (2021)).

quality. Then, this informational channel operates through affecting consumers' beliefs about goods' quality.

There are two main challenges in identifying the causal effect of recommendations on consumption and decomposing the relevant mechanisms that guide our experimental design.

The first challenge refers to identifying the causal impact of recommendations on consumption. For this, it is necessary to be able to compare outcomes between goods that are equally likely to be credibly recommended. In our intervention, we generate a *control*, set-aside group, and a *recommendation* group of recommendable goods that are ex-ante equally likely to be recommended. We do this through within-subject, good-level block randomization by exploiting the recommendation system's estimate of consumer-specific quality — which guides the preexisting recommendation algorithm. This enables us to compare consumption frequency holding fixed not only the likelihood of recommendation, but also idiosyncratic consumer quality.

The second main challenge is to disentangle exposure from informational effects of recommendation. We study the informational effects of recommendations both by examining how it impacts beliefs relative to otherwise ex-ante identical non-recommended goods, and how it affects consumption beyond mere exposure. In our experimental intervention, we elicit beliefs about unconsumed goods' quality. A byproduct of this elicitation is that it makes consumers consider the good, without providing the informational content of recommendation. As a result, we leverage this within our experimental design and, through our block randomization, we create a third *exposure-only* group of goods. Both goods in the exposure-only and recommendation groups are utilized for belief elicitation, but only those in the latter are selected for recommendation. We rely on this additional variation both to decompose the effects of exposure and information on consumption and examine the impact of recommendation on consumers' beliefs.

Our experiment is conducted on a movie-recommendation platform. The platform we conduct the study on is noncommercial and devoted to producing helpful recommendations (it does not host movies), ensuring there are no confounding strategic aspects arising from potential misalignment of preferences between consumers and the platform. The platform also features open-sourced data and algorithm implementation, and its data constitutes a central benchmark in the recommender system community for the development and evaluation of new recommender system algorithms, used in thousands of papers.[2] While our experiment is in the movie domain, the hypotheses are grounded on a general theoretical framework of consumer decision-making in environments where recommendation systems are typically deployed. We believe the core

[2]Two recent examples in economics are Chen et al. (2010) and Rossi (2021). However, the vast majority of papers using the data rely on the resulting ratings dataset to evaluate the performance of new recommendation system algorithms (see Harper and Konstan (2015) for an overview).

mechanisms of exposure and information are similar across many environments such as content on Instagram, music on Spotify, or articles on Google News.

The experimental intervention is simple: we generate random variation in recommendations in order to study its causal effect on consumption, and we elicit belief data about good quality prior to consumption to examine the underlying mechanisms. We partition each user's pre-experiment top 750 recommended movies into three groups: *recommendation*, *exposure-only*, and *control*. For a period of six months, we track movie consumption and recommendation of over 1,000 users, and elicit their beliefs through periodic surveys. During the experimental intervention only goods from the *recommendation* group show up in recommendations and we elicit beliefs about goods from both the *exposure-only* and *recommendation* group. In order to assess the causal effect of recommendation on beliefs, we elicit beliefs about a good both before and after recommendation/exposure. We compare the fraction of goods consumed across groups, how beliefs map to consumption, and how these beliefs are impacted by exposure and recommendations.

Our first main finding is that recommendation induces a significant increase in consumption. The experimental design isolates the role of consideration through comparison between exposed and non-exposed goods and we show it plays a meaningful role. However, the quantitative magnitude of the difference in consumption between recommendation and exposure indicates that recommendations impact consumption well beyond the consideration channel. Under our preferred specification we find that exposure alone leads to a 0.2 p.p. increase in consumption relative to a baseline of 1.2 p.p. consumption of goods in the control group. In contrast, recommendation leads to a 1.8 p.p. increase in consumption relative to the control group, indicating recommendation nearly doubles the probability of consumption relative to exposure alone.

Our second set of results highlights the relevance of the informational effects of recommendations. We examine if and how beliefs relate to consumption and recommendations affect beliefs. We observe that elicited beliefs explain consumption, with higher expected quality and lower uncertainty about a movie's quality being strongly associated with higher probability of its consumption. Then, we find that, when prior uncertainty is high, recommending a good lowers uncertainty about its quality and drives their expectations about quality closer to the platform's prediction. In other words, there is little scope for information provision to play a role if the consumer is already very certain of a good's quality; it is then when the consumer is more uncertain that recommendations' informational impact is most significant. This is the case even when controlling for additional information acquisition, such as visiting the details page on the platform, indicating that either the platform's recommendation and the provision of user-specific rating induces a shift in beliefs. Our test is conservative since we only observe beliefs after recommendation if recommendation does not shift beliefs enough to induce consumption.

Recommendations also promote additional information acquisition by consumers. Recommending goods not only informs consumers of their match value, but also reduces the cost of acquiring information on their details. Examining platform logs, we show that recommendation increases the probability of visiting a good's detail page by 4.5 p.p. relative to a baseline of 1.2 p.p., while exposure has the same estimated visit probability as the control group. This additional information acquisition influences consumers' beliefs, both reducing uncertainty and moving the expected quality closer to the platform-predicted quality, with larger magnitudes than those estimated for the unconditional effect of recommendation. These results point to recommendation having a smaller, though sizeable, direct impact on consumer beliefs as well as an effect mediated through spurring further information acquisition. Overall, our results provide support for an informational mechanism whereby recommendation provides information which influences beliefs that then influences consumption.

Our last set of results concerns the dynamic implications of consumption and recommendation through information spillover effects. In our theoretical framework, learning about a good's quality also informs the consumer about the quality of similar goods. A unique aspect of our setting is that we have a rich, high-dimensional dataset of good-specific attributes derived from content tags on the platform which enables us to have a precise measurement of good similarity. Using this, we document that there is a strong and robust correlation between movie similarity and beliefs. We find evidence for informational spillovers from consumption, but observe none arising from recommendations.

In summary, this paper shows that (i) recommendations determiningly affect consumption decisions, (ii) recommendations affect consumers' beliefs which in turn guide consumption, and that (iii) recommendations cause a greater propensity to acquire further information on recommended goods. Additionally, we establish the spatial correlation of beliefs and provide suggestive evidence on the existence of informational spillover effects from consumption.

These results have important implications for policy and societal considerations surrounding these systems. First, a large theoretical literature has spawned motivated by antitrust considerations of such systems in digital markets (Hagiu and Jullien, 2011; De Corniere and Taylor, 2019; Aridor and Gonçalves, 2022; Calvano et al., 2022). Our results highlight that the informational role played by such systems is first-order in modeling their impact and for evaluating policy interventions. Second, the documented spatial correlation in beliefs suggests the importance of dynamics in evaluation not just in terms of algorithmic confounding,[3] but also in terms of an analogous feedback loop on the consumer side. In particular, our results imply that recommendations im-

[3] That is, recommendation algorithms affect consumption which in turn affects the resulting data used to train the recommendation algorithm and influences future recommendations, as in Chaney, Stewart and Engelhardt (2018).

pact current consumption, which reduces uncertainty about similar goods, thereby increasing the probability of consuming these in the future. Thus, when measuring the effects of proposed policy interventions it is important to characterize not just their immediate impact, but also the future changes on consumption, especially in contexts where recommendation is optimized for firm profitability and not only consumer welfare.

Related Work

This paper contributes to a burgeoning literature on recommender systems and the impact of recommendations.

Evidence on the impact of recommender systems. A recent literature has examined whether recommendation systems impact consumption patterns. Senecal and Nantel (2004), Das et al. (2007), Freyne et al. (2009), Zhou, Khemmarat and Gao (2010), Claussen, Peukert and Sen (2019), Holtz et al. (2020), and Donnelly, Kanodia and Morozov (2022) show that recommendation meaningfully impacts consumption patterns on hypothetical choices in a lab experiment, Google News, a social network, a news website, YouTube, Spotify, and Wayfair respectively. An important question raised in the literature is not just if recommended goods are more likely to be consumed, but also whether they change the type of goods consumed. Specifically, existing work has studied the implications of recommender systems for individual consumption diversity (Fleder and Hosanagar, 2009; Nguyen et al., 2014), as well as aggregate consumption diversity (Van Alstyne and Brynjolfsson, 2005; Brynjolfsson, Hu and Simester, 2011; Hosanagar et al., 2013; Lee and Hosanagar, 2019; Holtz et al., 2020; Donnelly, Kanodia and Morozov, 2022), with relevant implications for market competition and product variety, as well as the emergence of filter bubbles and echo chambers supporting opinion polarization. Furthermore, the existence and importance of spatial correlation in beliefs has been explored in the context of consumer search by Hodgson and Lewis (2019).

Our paper differs from these both by identifying the effect of recommendations on consumption via good-level randomization and by examining the mechanisms through which recommendation operates. For instance, a typical approach is to compare consumption in a personalized recommendation system to a popularity ranking, making it difficult to discern whether differences are due to information or exposure, as well as resulting in a treatment group with systematically higher idiosyncratic quality relative to the control group. Our experimental design resolves this concern by having a control and an exposure-only group of goods *that would have been recommended*, while the beliefs data we collect allows us to explain why recommendations may influence consumption — crucial for understanding how recommendation influences consumption diversity. Finally, the beliefs data allows us to directly measure spatial correlation and spillovers in beliefs as opposed to inferring it from an economic model.

Theoretical work on recommender systems. The importance of understanding mechanisms is further highlighted by the literature that studies the economic implications of recommendation systems on competition and antitrust issues. One line of theoretical work models recommendation as primarily forcing consideration of certain items, so that recommendation provides little informational value beyond this (Armstrong, Vickers and Zhou, 2009; Hagiu and Jullien, 2011; De Corniere and Taylor, 2019; Bourreau and Gaudin, 2022).[4] Another class of models conceives recommendations as directly providing Bayesian consumers with information on the match-value of products in order to shift their beliefs and therefore the items they consume (Bergemann and Ozmen, 2006; Che and Hörner, 2017; Aridor and Gonçalves, 2022; Lee and Wright, 2021; Lee, 2022; Calvano et al., 2022). Our paper directly allows us to understand the importance of each of these mechanisms in driving the influence of recommender systems, which is important for guiding modeling assumptions within this policy-relevant literature.

Advertising. Finally, there is a connection between the mechanisms driving the role of recommendation systems and advertising in consumer choice. There is a vast literature on the economics of advertising, but, as several examples that align with the mechanisms we consider, some (e.g. Honka, Hortaçsu and Vitorino, 2017; Tsai and Honka, 2021) argue that advertising in U.S. banking and television ads acts through an exposure/consideration channel, and others (e.g. Grossman and Shapiro, 1984; Meurer and Stahl II, 1994; Ackerberg, 2003; Sahni and Nair, 2020) focus on its informational effects. While there are similarities in the underlying mechanisms, they target distinct aspects of the consumer choice process, are generated using different methods and by agents with differing incentives. Recommendation systems aggregate consumer data on a platform to provide predictions for multiple goods present on the platform, while advertisements are aimed at persuading consumers to purchase one good or goods from one brand.

2. Framework and Hypotheses

2.1. Framework

In this section, we lay out our theoretical framework which guides the hypotheses that our experiment will test. Our experiment will allow us to test both assumptions and implications of the model.

[4]See Cheung and Masatlioglu (2021) for a recent decision-theoretic approach to modeling recommendation's impact on choice based on consideration.

Consumers. At each period $t = 0, 1, ...$, consumer i makes a sequence of consumption choices $\{c_{i,t}\}_t$, in which the consumer chooses a good from a finite set of goods X or takes an outside option. Each good x is fully described by its attributes, which take value in $\mathscr{A} \subset \mathbb{R}^n$, where $\mathscr{A}$ is compact. We denote goods consumed up to time t by consumer i by $C_{i,t} := \{c_{i,\ell} \in X, \ell \leq t\}$ and assume that consumers choose each good at most once.

Quality and Consumers' Beliefs. The quality of a good for consumer i specifies an idiosyncratic match value, $q_i : \mathscr{A} \to \mathbb{R}$. In many environments in which recommendation systems are deployed, the goods are experience goods. We assume consumers are uncertain about how they map attributes to quality, with q_i being drawn from the space of sample paths of a Gaussian process with a continuous prior mean function $q_i^b : \mathscr{A} \to \mathbb{R}$ and a symmetric, positive semi-definite covariance function $\sigma_i^b : \mathscr{A} \times \mathscr{A} \to \mathbb{R}$ such that $\sigma_i^b(x, x')$ is continuous. This induces Gaussian beliefs about quality: for any finite set of goods, $A \subseteq \mathscr{A}$, their quality is jointly normally distributed, with mean $(q_i^b(x))_{x \in A}$ and covariance matrix $\Sigma_{i,A}$ with element (x, x') given by $\sigma_i(x, x')$. We further assume consumers believe that similar goods have similar quality, that is, $\sigma_i^b(x, x')$ is nonnegative and decreasing in the distance between the goods, $\|x - x'\|$. This implies that learning about a good's quality is more informative about that of more similar goods (goods closer in the attribute space) than of goods with very dissimilar attributes.

Consideration and Exposure. Consumers may not consider all goods at a given time. In other words, at time t, consumer i considers a subset of available goods, and their consideration set is denoted by $\Gamma_{i,t}$. The discrepancy between $\Gamma_{i,t}$ and X may arise because consumers may have limited working memory and therefore consider only a limited subset of available goods or are unaware of which goods are available for consumption and form their consideration set through potentially costly search.

Recommender systems manipulate consideration sets by exposing consumers to some goods, thereby forcing them to consider them. We consider such manipulations through exposure and write $e_{i,x,t} = 1$ to denote the situation in which the recommender system forces exposure of consumer i to good x at time t. Naturally, exposure implies consideration — $e_{i,x,t} = 1 \Longrightarrow x \in \Gamma_{i,t}$ — but the converse does not hold in general: consumers may consider goods that they are not exposed to on the recommendation platform.

Learning and Recommendation. We allow for consumers to acquire (potentially costly) information about $q_i(x)$ prior to consuming the good. Recommender systems typically provide information through various channels prior to consumption by providing a noisy estimate $q_i^p(x)$ about the quality of a subset of items — e.g. by featuring these as being of high potential quality ('top picks for you') or by providing consumer-specific predicted quality — directly through the inferences consumers make from having the good be explicitly recommended or by lowering the cost of information acquisition for some goods. We denote a recommendation of good x to consumer i at time t by $r_{i,t,x} = 1$ and, for simplicity, information is modeled as a Gaussian signal that is centered on $q_i(x)$. Finally, upon consuming a good, consumer i learns its quality $q_i(x)$.

Choice. Consumer i evaluates quality according to an increasing and concave utility function $u_i : \mathbb{R} \to \mathbb{R}$, capturing the consumer's attitudes toward uncertainty. We assume that the utility associated with the outside option available at time t, $u_i(o)$, is distributed according to some distribution F, independently from $q_i(x)$ and across periods, and known to the consumer prior to making their choice at t. Then, given their information at time t, consumer i at time t chooses to maximize current expected utility, i.e.

$$c_{i,t} \in \arg\max_{y \in \Gamma_{i,t} \cup \{o\}} \mathbb{E}_t[u_i(y)].$$

2.2. Hypotheses

As our framework points out, recommender systems affect consumption directly through two mechanisms: exposure of consumers to a good and recommendation of that good. Note that if good x is recommended to consumer i, then consumer i is also exposed to it, that is, $r_{i,x,t} = 1 \implies e_{i,x,t} = 1$, and therefore, the good is considered, $(e_{i,x,t} = 1 \implies x \in \Gamma_{i,t})$. This motivates our first hypothesis, which distinguishes between the effect of recommendation and exposure on consumption:

Hypothesis 1 (Impact of Recommender Systems on Consumption) *(1) Exposure to a good increases the likelihood of its consumption relative to no exposure. (2) Recommendation of a good increases the likelihood of its consumption relative to simple exposure.*

We then turn to examining the mechanisms through which recommendation acts on consumption. Underlying our theoretical framework is the assumption that recommendation affects consumption by affecting consumers' beliefs, which, in turn, explain consumption patterns. In particular, one would expect that consumers are more likely to choose goods they believe have a higher expected quality, $q_{i,x,t} := \mathbb{E}_t[q_i(x)]$, which in our framework corresponds to utility being increasing in quality. On the other hand, if consumers are uncertainty averse (concave utility), then, all else equal, they are more likely to choose goods about whose quality they are less uncertain. Our second hypothesis summarizes these predictions:

Hypothesis 2 (Beliefs Explain Consumption) *The likelihood of consuming a good increases in expected quality and decreases in uncertainty.*

If consumption is driven by consumers' beliefs, then the informational impact of recommendations can be assessed by studying how recommendations affect consumers' beliefs. In our theoretical framework we assumed that the recommendation of good x to consumer i at a given time provides a noisy signal of the true quality, that is, $q^p_{i,x,t} = q_i(x) + \epsilon_{i,x,p}$, where $\epsilon_{i,x,p} \sim N(0, \sigma^2_{i,x,p})$. This leads not only to the consumer to be more certain about the quality of good x but — owing to the normality structure of the setting — also drives their expected quality toward the signal.[5] We will term such signal the *platform's predicted quality*. We then posit the following:

Hypothesis 3 (Impact of Recommendations on Beliefs) *Recommendation of a good (1) decreases the uncertainty relative to the good's quality, and (2) drives consumers' expected quality toward the platform's predicted quality.*

Finally, we consider the effects of spatial correlation. One key feature of our model is that beliefs about goods' quality are correlated across the attribute space:

Hypothesis 4 (Beliefs' Spatial Correlation) *The distance between two goods' attributes is correlated with consumers' beliefs about the goods' quality, in particular, (1) consumers' expected quality, and (2) uncertainty associated to the goods' quality.*

This suggests that information about a good's quality impacts beliefs about other goods' quality differentially, depending on their location on the attribute space. Specifically, information about a particular good's quality is more informative about the quality of more similar goods. Since both consumption and recommendation of a good provide information, this implies our next prediction:

Hypothesis 5 (Information Spillover Effects) *(1) Consumption of a good decreases uncertainty more for goods closer in attributes. (2) Recommendation of a good decreases uncertainty more for goods closer in attributes.*

3. Experimental Design

In order to study whether and how recommender systems impact consumption we rely on an experimental intervention on a movie recommendation online platform, MovieLens. Our intervention has two main features: (i) we generate random variation in recommendations in order to

[5] Note that, while we left search and information acquisition unspecified to focus attention on the main features of the model, whether recommendation acts through direct information provision or by reducing costs to acquire information would lead to similar predictions in this setting, even if with potentially different magnitudes. Indeed, our setting we will be able to quantify the magnitudes of a subset of the different channels.

study its causal effect on consumption, and (ii) we elicit belief data about good quality prior to consumption to examine the mechanisms through which recommendation acts. In this section we provide background information on the platform and describe our experimental procedures.

3.1. Background on the Recommendation Platform

MovieLens is a movie-recommendation platform created in 1997 in which consumers can obtain information about movies as well as personalized movie recommendations based on their ratings. The platform has since been widely used as a movie discovery platform by consumers, and the resulting data on movie ratings is a central benchmark in the recommender system community for the evaluation of new recommender system algorithms, used in thousands of papers in this literature.[6]

The platform's home page displays movies organized by categories in rows, with the very first showing eight 'top picks', the platform's top recommended movies for the user. Movies are set in a grid fashion, with their poster, title, and the platform-predicted rating for the user.[7] When hovering over a movie title, users see the genres of the movies, the predicted rating according to the recommendation algorithm, and the number and average of community ratings for the movie. Rows following the first correspond to recent releases, and other categories of potential interest (e.g. 'favorites from the past year' or 'new additions').

The platform is mainly used as a movie discovery tool; it does not provide consumption opportunities (it does not host movies to stream nor does it direct users to other platforms). Consumers periodically use the platform to find movies to watch and rate them after watching. As the platform is free to use and noncommercial, users have no reason to not truthfully report their ratings as it is in their benefit to provide the platform with truthful information in order to get the best possible recommendations. If a user clicks into the movie page they have access to detailed information about the movie, including its trailer, synopsis, cast, tags associated with the movie (Vig, Sen and Riedl, 2012), and similar movies.

The recommendation system used by the platform is of high quality and ideal from both a user and a researcher perspective for several reasons. First, the underlying training data is of high quality, since a subset of the underlying ratings data is open source and used as a benchmark dataset for the evaluation of new algorithms for such systems. Second, the set of algorithms used are also open source and constitute canonical implementations of widely used item-item or SVD

[6]For instance, the search expression "MovieLens dataset" or "MovieLens data" returns over 9,000 entries on Google Scholar, whereas "Netflix dataset" or "Netflix data" — which includes both proprietary data and the public access Netflix data associated to Netflix open competition for the best collaborative filtering algorithm to predict user ratings for films — returns less than half the number of entries.

[7]Screenshots of the platform's interface are included in the Appendix D.

collaborative filtering algorithms (see Ekstrand et al., 2011). Third, the noncommercial aspect of the platform ensures that the recommendation system is designed to maximize consumer welfare as opposed to considering platform profitability.

3.2. Experimental Intervention

The platform provides a natural setting to test our hypotheses. We take ratings as a measure of realized good quality and the platform's predicted rating as a noisy signal; we will use the terms rating and quality interchangeably. We take the platform's 'top picks' category as a recommendation of the specific goods listed there. To understand how consumers' beliefs about expected quality are affected by recommendation, we incorporate a regular survey on the platform to elicit this data.

We identify the effect of recommendations by inducing exogenous randomness and comparing outcomes for recommended goods to those which would otherwise be recommended. At the start of the experiment, consumer i's $3N$ goods (not previously consumed) with highest platform-predicted quality are split into three sets: a control set $X_{i,C}$, an exposure-only set $X_{i,E}$, and a recommendation set $X_{i,R}$. We elicit beliefs about quality only for goods in the exposure-only set ($X_{i,E}$) and the recommendation set ($X_{i,R}$), and we restrict recommendations to goods in the latter. So as to have recommendations of goods in the recommended set to be meaningful while N to be large, we set $N = 250$.

In order to enable us to identify the effects of recommendations, the control, exposure-only, and recommendation sets are consumer-specific and constructed by stratified block randomization in a manner that controls for consumer's idiosyncratic quality assessment. Specifically, we assign the n-th, $(n+1)$-th, and $(n+2)$-th goods with highest platform predicted quality and match these one-to-one with $X_{i,C}$, $X_{i,E}$, and $X_{i,R}$ uniformly at random. Note it is not possible to elicit beliefs about a good's quality without exposing the consumer to such a good. We leverage the unavoidable consequence of belief elicitation implying exposure to identify the effect of exposure by comparing outcomes for goods in the exposure-only set and those in the control set.

Every day the consumer enters the platform, we elicit their beliefs about the quality of 10 goods not previously consumed by time t through a survey.[8] Following the completion of the survey, the consumer is taken to the platform's home page, in which they are presented with a set of 8 goods that are recommended (their 'top picks'). Letting t denote the time of the t-th belief elicitation, the following procedure summarizes how we choose the goods for belief elicitation and recommendation.

[8] The survey can be deferred to the following visit to the platform. The set of platform recommendations remains the same until the survey is completed.

Procedure 1 (Belief Elicitation and Recommendation)

Step 1. *Elicit which goods were consumed since $t-1$ and remove consumed goods from the control, exposure-only, and recommendation sets: $X_{i,\ell,t} := X_{i,\ell} \setminus C_{i,t}$, $\ell = C, E, R$.*

Step 2. *Generate a new sorting for each of the resulting sets by platform-predicted quality plus i.i.d. Gaussian noise.*

Step 3. *For each $\ell = E, R$, define $S_{i,\ell,t}$ as a selection of 2 of from the first 8 goods in $X_{i,\ell,t}$, drawn uniformly at random.*

Step 4. *Elicit beliefs on 10 goods: the 4 in $S_{i,E,t}$ and $S_{i,R,t}$, those in $S_{i,E,t-1}$ and $S_{i,R,t-1}$ (for $t \geq 2$), and the remaining being drawn uniformly at random from other goods in $X_{i,E,t}$.*

Step 5. *Recommend the first 8 goods in $X_{i,R,t}$ (including $S_{i,R,t}$) as per the sorting generated in 2.*

Our sampling procedure generates exogenous variation both on the goods which are recommended and those about which beliefs are elicited. Since, absent an intervention, the 8 goods with highest platform-predicted quality are recommended, recommending the top 8 goods from the recommendation set according to noisy predicted quality (Steps 2 and 5) allows us to retain platform credibility by having meaningful recommendations while exogenously varying which goods are recommended.[9,10] As recommended goods are of high predicted quality, our sampling procedure elicits beliefs on goods of similar predicted quality in the exposure-only set (Step 3). This, together with having beliefs about goods being elicited in two subsequent periods (Step 4), enables us to identify how recommendations impact beliefs. Finally, by randomly selecting two goods from the exposure-only set to elicit beliefs on (Step 4), we are able to learn consumers' beliefs about goods on a broader subset of the attribute space.

3.3. Measurements

In this section we provide details on belief data and elicitation, consumption measurement, attribute space, and search patterns.

Beliefs about Quality. In order to elicit consumers' beliefs about quality, they are asked to fill out a survey whenever they return to the platform.

[9] In particular, the variance of the additive noise was set to generate enough variation in the set of presented movies, without compromising the platform's necessity of keeping recommendations meaningful for the consumers. In consultation with the platform's experts, variance of Gaussian additive noise was set to .2.

[10] In order to validate that recommendations retained credibility, we conduct a difference-in-differences analysis between the set of targeted users that opted into the study versus those that opted not to participate. We find no evidence that the consumption rate of the opted in consumers was lower than those who opted not to be in the experiment.

In the survey, we elicit beliefs about the quality of 10 movies, selected according to the procedure described in Section 3.2. For each of the 10 selected movies, the consumer is asked whether or not they had watched it. The survey needed to be complete to be submitted. The survey interface matches closely the platform's interface for ratings, except in omitting the predicted rating and not displaying additional information other than the name and movie poster.[11] Additionally, unlike on the homepage, it is not possible for the consumer to either hover over or click into the details page in order to acquire additional information on the good. If they have declared to have watched it, we elicit their rating — corresponding to $q_{i,x}$, the idiosyncratic realized quality — and an approximate date of when they watched it. If they declare not to have watched it, we elicit their expected rating, which we will denote as expected quality, based on the movie poster, corresponding to $q^b_{i,x,t}$ on the same scale as used to rate movies in the platform (a 10-point Likert scale). We also ask how certain they are of their reported expected rating on a 5-point Likert scale, which we standardize to the unit interval and take as a measure of uncertainty $\sigma^b_{i,x,t} \in [0,1]$, where higher values are associated with higher uncertainty about the expected rating.

We assume belief data was truthfully reported. While the survey was not incentivized, consumers do not have an incentive to misreport since the platform is non-commercial — there are no strategic aspects to the interaction between the platform and the consumers — and helping the platform improve on its recommendation system is in the consumers' own interest.

The analysis of the belief data provides assurance that the belief data is internally consistent and consistent with the consumers' behavior on the platform. In Appendix A, we show that (i) consumers' expected quality on average equals the realized quality rating they provide after consumption, and (ii) the distance between consumers' expected and realized quality is increasing in our uncertainty measure. We also show that (iii) uncertainty about expected quality is decreasing in the movie's popularity — proxied by number of community ratings — and on whether the movie is a sequel or part of a franchise.

Good Consumption. The platform is primarily used for good discovery and there is no direct observation of consumption. We rely on two proxies, a consumption survey and the timing of ratings.

The most natural proxy for consumption is to rely on ratings that consumers input on the platform to determine what is consumed and when. We call this first measure 'Rating'. One possible concern is that the rating does not refer to a recent experience.

We construct a more robust measure of consumption by relying on two elements to obtain an approximate date of consumption: a consumption survey and the belief elicitation survey. The

[11] See Appendix D for interface screenshots.

consumption survey is presented just before every instance of the belief elicitation survey and allows us to directly track consumption during our intervention. The consumer is asked whether they have watched any movie since they were last on the platform. The consumer can search for the movie as they can on the platform and a number of options appear. If the consumer declares they have seen a movie, they need to rate the movie and provide an approximate date of when they watched it. The belief elicitation survey itself also requires the consumer to input an approximate consumption date.

Based on these elements, we construct our restrictive consumption measure. It determines a consumer watched a movie at time t if the movie was rated at time t and had been previously marked unwatched, or if the approximate consumption date provided in the survey was time t. We call this second measure 'Robust'.

Attribute Space. A unique aspect of our setting and the MovieLens platform in particular is that we have a rich, high-dimensional attribute space that comes from the tag genome. Each good x is associated with a vector of attributes in $\mathbb{R}^n$ that uniquely identify the movie; for reliability, x is required to have at least 50 community ratings.[12] The usefulness and details about how these tags are constructed is discussed extensively in Vig, Sen and Riedl (2012).

We report our findings using the Euclidean distance on the (exogeneously defined) attribute space, similar to previous studies that have used this data (Nguyen et al., 2014). Our results are qualitatively robust to considering cosine 'distance,' a notion of attribute (dis)similarity commonly used in the literature.[13]

We sample several movies and report the three most similar movies to these to highlight our similarity measure accords with intuition. For instance, the three most similar movies to *John Wick* are *John Wick: Chapter Two*, *The Equalizer*, and *John Wick: Chapter Three*. The three most similar movies to *Lady Bird* are *Booksmart*, *Eighth Grade*, and *Wildlife*. In Appendix B, we provide evidence on the validity of the attribute space by showing that indeed sequels and franchises are closer to each other than to other movies.

Search Patterns. We document when consumers acquire further information about a good on the platform. For this, we use consumers' granular on-platform search data and encode visits to each movie's details page as for the purpose of acquiring information when these occur not within 30 minutes prior to rating the movie or after having rated the movie. This ensures we are

[12]For each good, each of the 1128 tags is attributed a relevance score, from 0 to 1, and an irrelevance score, with a good being described by the vector of relevance and irrelevance scores, normalized to have Euclidean norm of 1, i.e. $\|x\| = 1$. Missing tag-relevance measures are imputed utilizing machine-learning techniques; the underlying methods are discussed here: `https://grouplens.org/datasets/movielens/tag-genome-2021/`.

[13]The 'cosine distance' between two goods x, y in our attribute space is given by $\|x - y\|^2/2$.

capturing any instances of (within-platform) information acquisition before they watched the movie and rules out visits to the details page with the purpose of providing a rating.

3.4. Recruitment

In our intervention, we target a random sample from a subset of the platform's users.[14] In order to mitigate the heterogeneity of treatment effects across consumers arising from differences in the quality of the recommendations, we restrict our sample to users who satisfy the following conditions: (i) having rated more than 100 movies in total; (ii) having rated fewer than 3,000 movies in total; and (iii) over the previous $m = 1, 2, 3, 4$ months, having rated a minimum of $\lceil 1.5m \rceil$ movies. The first condition is a minimum data requirement so that the recommender system algorithm utilized by the platform is able to provide valuable recommendations. This is especially important given that, throughout the duration of the intervention, the assignment of movies to treatments is held fixed and therefore so is the set of movies that can be recommended. The second excludes high-powered users, as these are likely to would constitute outliers. The last condition, seeks to guarantee that the targeted user is minimally active on the platform over the recent past. These criteria were chosen in consultation with the platform's experts in order to ensure that the data is representative of the overall platform population, with stable users who are familiar with the platform's recommender system.

The roll-out of the study is phased in order to control for implementation issues and targets 4,572 users. On March 29, 2021, 100 eligible users were randomly selected to participate in the study. On April 5, the study is expanded to an additional 500 randomly selected eligible users. On April 15, 3,972 additional eligible users were randomly selected to take part in the experiment. The data collection was pre-registered to conclude on October 31, 2021. The length of the study period was selected based on power calculations and considering the possibly slow rate of consumption of movies over time.

From this set of users, 1,452 consumers decide to enroll in the study and 1,033 consumers filled out at least one belief elicitation survey. Over the 6 month study period, 290 consumers explicitly opted out of the study at some point with 107 opting out before being in the study for 7 days. Our primary analysis focuses on the 1,006 consumers who completed at least one survey and were enrolled in the study for at least 7 days.

[14] The use of or access to the platform is prohibited to individuals under the age of 18, as per the platform's terms of service.

3.5. Pre-Registration

The experimental design, data collection phase (including the length of the intervention), and hypotheses were pre-registered using the AEA RCT Registry, with ID AEARCTR-0007545.

4. The Impact of Recommendations on Consumption

In this section, we test Hypothesis 1: whether exposure induces additional consumption and whether recommendation further influences consumption beyond exposure. We study two alternative specifications, the first of which is as follows:

$$c_{i,x} = \beta_0 + \beta_1\{x \in X_{i,E} \cup X_{i,R}\} + \beta_2\{x \in X_{i,R}\} + \epsilon_{i,x} \tag{1a}$$

where $c_{i,x} = 1$ if by the end of the intervention consumer i reported to have consumed good x (and zero if otherwise), $X_{i,E}$ and $X_{i,R}$ denote consumer i's exposure-only and recommendation sets as defined at the start of the intervention, and $X_{i,E} \cup X_{i,R}$, the exposure set, given by their union.

Exposure through belief elicitation of good x only occurs if x is in the exposure-only or recommendation sets, and platform recommendation of good x only occurs if x is in the recommendation set. Naturally, the consumer may have other recommendation sources and, even through the platform, be exposed to goods in either of these sets. Since we stratified our randomization by consumer-specific tastes, exposure and recommendation to goods via other channels should be orthogonal to treatment assignment.

While this specification enables a clear and straightforward causal estimate of the impact of exposure and recommendation on consumption, it is potentially too conservative. Specifically, it does not take into account the fact that some goods in the recommendation set are never explicitly recommended and some goods in the exposure-only set are never selected for beliefs elicitation. In order to obtain better estimates on the treatment effects, we consider two additional strategies.

First, we consider the same specification (1a), but restricting the sample to goods in exposed strata. While we would like to compare the effect of actual exposure and actual recommendation to the control group, realizations of exposure and recommendation are not independent of platform-predicted quality.[15] In order to resolve this issue, we take not only goods that the consumers were actually exposed to through belief elicitation or recommended but also goods in their strata. Recall that, at the outset of the intervention, goods with the ($3n$+1)-th, ($3n$+2)-th, or ($3n$+3)-th highest platform-predicted quality for consumer i are bundled into the same stratum, n, and

[15] A design restriction, since recommendations need to remain useful and meaningful to consumers.

block-randomized into the control, exposure-only, and recommendation sets. Then, if good x was recommended ($r_{i,x} = 1$) to consumer i or exposed to it through our experimental intervention ($e_{i,x} = 1$), we include x in the sample, as well as the other goods in the same stratum used for block-randomization. We then estimate the same specification (1a), but obtaining more precise estimates on the average treatment effect of exposure and recommendation on consumption.

Second, we restrict to estimating the average treatment effect of recommendations relative to exposure by estimating the following alternative specification:

$$c_{i,x} = \beta_0 + \beta_1 r_{i,x} + \epsilon_{i,x} \tag{1b}$$

where $r_{i,x} = 1$ if consumer i was recommended good x (zero if otherwise). Here we rely exclusively on goods that were selected for belief-elicitation as per step 3 in Procedure 1, which — together with our block randomization — guarantees that, in this sample, actual recommendation $r_{i,x}$ is orthogonal to the good's characteristics.

Our first result is the resounding positive empirical support for Hypothesis 1(1) and 1(2). In Table 1, we present the estimates for specification (1a) (columns (1)-(4)) and specification (1b) (columns (5)-(6)), using both of the consumption measures described in Section 3.3.2. In both of these specifications, the coefficient on exposure set and those on recommendation set and recommendation measure, respectively, denote the average treatment effects (ATE) of exposure and recommendation (beyond exposure) on consumption as per Hypothesis 1(1) and 1(2).[16]

In the conservative specifications reported in columns (1) and (2), we find that exposure increases consumption probability by 0.1 to 0.3 percentage points (p.p.), and recommendation leads to a further increase of over 0.6 p.p. relative to exposure — about a 1 percentage point increase relative to the control group. Columns (3) and (4) restrict focus to strata in which a good was used for belief elicitation or recommended, and, as expected, we find stronger treatment effects: exposure increases consumption probability by 0.2-1.1 p.p. relative to the control group, while recommendation adds to it about 1.3-1.8 p.p. more, with a combined effect of over 2 percentage points increase in consumption probability vis-á-vis the control group. Consistent with this, columns (5) and (6) estimate the effect of recommendation as leading to a 1.2 p.p. increase beyond exposure, by considering only goods that would be recommended (of high predicted quality) and used for belief elicitation. Overall, even the magnitude of the impact of recommendation on consumption probability is fairly consistent across the different estimation strategies. We note that while we do not include consumer fixed effects as our randomization is within consumer, their inclusion does not affect the results (see Appendix C).

[16]Throughout the paper we will report results from the linear probability model specification, but our results are robust and with similar estimated marginal effect sizes if we estimate a logistic regression as well.

	Consumption $(c_{i,x})$					
	(Rating)	(Robust)	(Rating)	(Robust)	(Rating)	(Robust)
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure Set	0.003***	0.001**	0.011***	0.002***		
$(\mathbb{1}(x \in X_{i,E} \cup X_{i,R}))$	(0.0003)	(0.0002)	(0.001)	(0.001)		
Recommendation Set	0.006***	0.008***	0.013***	0.018***		
$(\mathbb{1}(x \in X_{i,R}))$	(0.001)	(0.001)	(0.002)	(0.002)		
Recommendation					0.012***	0.012***
$(r_{i,x})$					(0.003)	(0.003)
Constant	0.007***	0.007***	0.012***	0.012***	0.017***	0.015***
	(0.0004)	(0.0004)	(0.001)	(0.001)	(0.002)	(0.002)
Observations	754,500	750,978	93,321	92,001	11,852	11,686
R^2	0.051	0.051	0.149	0.153	0.109	0.110

Clustered standard errors at the consumer level in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1: The Impact of Recommendation on Consumption (Hypothesis 1)

Notes: This table tests if exposure and recommendation impact consumption probability. Each column displays the average treatment effect of exposure and recommendation on consumption for the different measures of consumption. Each observation corresponds to a pair (consumer i, good x). Columns (1)-(4) correspond to specification (1a) and columns (5)-(6) to specification (1b). For columns (1)-(2), we include all consumers i and all goods x in the consumer-specific control, exposure-only, and recommendation sets. In columns (3)-(4), for each consumer we include the goods to which they were exposed through the belief elicitation survey, and all the goods in the same consumer-specific stratum. Columns (5)-(6) restrict to goods exposed to the consumer through belief elicitation and sampled as per step 3 in Procedure 1.

5. Drivers of Recommendation's Influence

The results from Section 4 indicate that exposure increased consumption probability compared to the control group, and that recommendation further increases consumption probability. In this section, we explore whether recommendation affects consumption by impacting consumers' beliefs.

5.1. Beliefs Explain Consumption

We start by evaluating the extent to which the belief data explains consumption behavior. We test Hypothesis 2, which — in line with our theoretical framework — suggests the likelihood of consumption is increasing in the expected quality and decreasing in reported uncertainty.

	Consumption $(c_{i,x})$			
	(Rating)	(Robust)	(Rating)	(Robust)
	(1)	(2)	(3)	(4)
Expected Quality	0.011^{***}	0.010^{***}	0.015^{***}	0.013^{***}
$(q^b_{i,x})$	(0.001)	(0.001)	(0.003)	(0.003)
Uncertainty	-0.030^{***}	-0.023^{***}	-0.038^{***}	-0.028^{***}
$(\sigma^b_{i,x})$	(0.006)	(0.005)	(0.009)	(0.008)
Recommendation			0.019	0.017
$(r_{i,x})$			(0.012)	(0.011)
Rec. × Exp. Quality			0.003	0.003
$(r_{i,x} \cdot q^b_{i,x})$			(0.003)	(0.003)
Rec. × Uncertainty			-0.025^{**}	-0.024^{**}
$(r_{i,x} \cdot \sigma^b_{i,x})$			(0.011)	(0.010)
Consumer FEs	Yes	Yes	Yes	Yes
Observations	14,362	14,327	11,852	11,809
R^2	0.085	0.084	0.125	0.121

Clustered standard errors at the consumer level in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Beliefs Explain Consumption (Hypothesis 2)

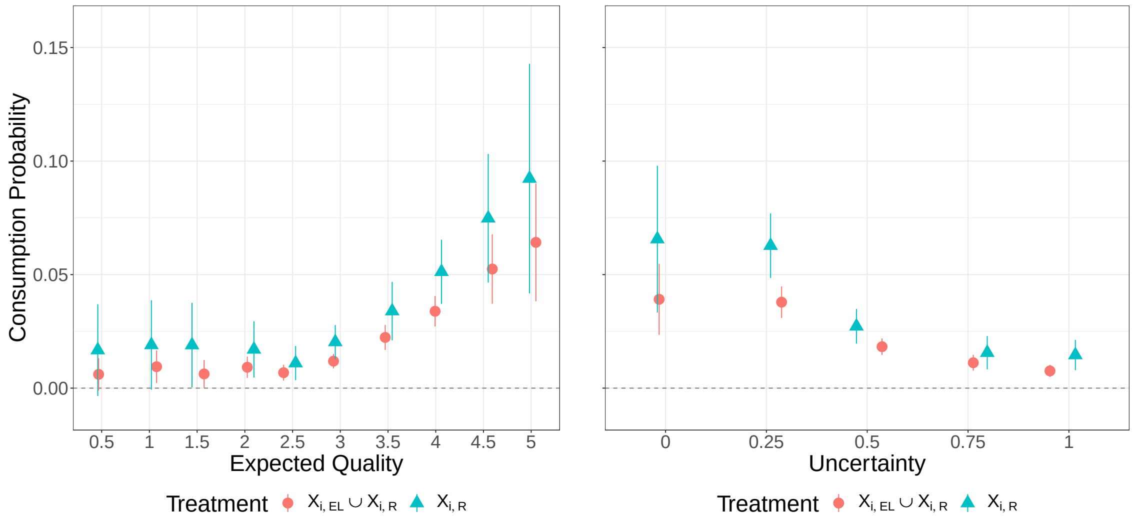
Notes: In this table we ascertain if beliefs explain consumption, i.e. if consumption probability increases in expected quality and decreases in uncertainty. Each observation corresponds to a pair (consumer i, good x). In columns (1)-(2), corresponding to specification (2a), the sample is restricted to goods x in the exposure-only set — to avoid confounds due to recommendation — and uses the first elicitation for each good and consumer. In columns (3)-(4), corresponding to specification (2b), the sample is restricted to goods selected for belief elicitation and sampled into $S_{i,E,t}$ or $S_{i,R,t}$ (step 3 of Procedure 1) to control for good characteristics across treatment. All columns rely on the first belief elicitation, with odd columns using the standard consumption measure, while even ones use the robust consumption measure.

We evaluate the relationship through the following regression:

$$c_{i,x} = \beta_1 q^b_{i,x} + \beta_2 \sigma^b_{i,x} + \text{FE}_i + \epsilon_{i,x} \tag{2a}$$

where $q^b_{i,x}$ and $\sigma^b_{i,x}$ denote consumer i's expected quality and uncertainty associated with good x and FE_i denote consumer fixed effects. We highlight that this is a test of association — that consumers' beliefs are associated with their consumption decisions in a particular manner; we take as self-evident from economic theory and the prevailing patterns in the data the causal relation from consumers' quality beliefs to their consumption decisions.

Our results, displayed in columns (1)-(2) of Table 2, support Hypothesis 2. For both specifications



(a) Expected Quality

(b) Uncertainty

Figure 1: Beliefs Explain Consumption: Heterogeneous Effect by Treatment

Notes: The figure compares consumption frequency between goods assigned to the recommendation set (blue, triangle marker; $X_{i,R}$) and those assigned to the exposure-only set (red, circle marker; $X_{i,E} \cup X_{i,R}$) by relying on the first belief elicitation. The markers represent the fraction consumed out of the goods with a given expected quality (panel (a)) or uncertainty level (panel (b)); the whiskers denote 95% confidence intervals using standard errors clustered at the consumer level.

we rely on the data from first time consumer i's beliefs about x are elicited, restricting to goods that were in the exposure-only bin. We exclude cases of recommendation explicitly, so as to avoid the estimates from being contaminated by any effect of recommendations on beliefs.[17] Both estimates suggest that higher expected quality increases consumption, while higher uncertainty decreases it. The estimates are economically meaningful: relying on column (1), a one unit increase in the expected quality (measured 0 to 5) leads to an increase in consumption probability of about 1 p.p., while going from the lowest to the highest uncertainty level decreases consumption probability by approximately 3 p.p.

In Figure 1, we explore whether the relationship between beliefs and consumption decisions is affected by recommendation-specific heterogeneous effects. The figure suggests that, aside from recommendation increasing consumption probability, the effect of expected quality on consumption is virtually unchanged by recommendation (a level change), while recommendations strengthen the relationship between uncertainty and consumption (a steeper relationship). We test these suggestive results by estimating the same specification (2a), but allowing for potential

[17]However, regardless of whether one considers the last or first belief elicitation and excludes or not goods in the recommendation set, the estimates are similar and the conclusions on the effects the same.

recommendation-specific effects:

$$c_{i,x} = \beta_1 q^b_{i,x} + \beta_2 \sigma^b_{i,x} + \gamma_1 1(x \in X_{i,R}) + \gamma_2 1(x \in X_{i,R}) \cdot q^b_{i,x} + \gamma_3 1(x \in X_{i,R}) \cdot \sigma^b_{i,x} + \text{FE}_i + \epsilon_{i,x} \quad (2b)$$

where $q^b_{i,x}$ and $\sigma^b_{i,x}$ denote consumer i's expected quality and uncertainty associated with good x. In order to control for ex-ante good characteristics across treatment, we restrict the sample to goods sampled for belief elicitation as per step 3 of our Procedure 1. Columns (3)-(4) of Table 2 confirm the conjecture.

5.2. The Impact of Recommendations on Beliefs

Now that we have established that the belief data displays reasonable properties and correlates positively with consumption decisions, we explore the extent to which recommendation impacts these beliefs as a possible explanation for the increase in consumption resulting from recommendation. We characterize the direct effect of recommendation on the level of expected quality and degree of uncertainty as well as the indirect effect of recommendation on inducing consumer-driven exploration and information acquisition.

Direct Impact of Recommendations on Beliefs. The primary effect of information is to decrease uncertainty. If recommendations have an informational role, they ought, first and foremost, to decrease uncertainty. This is the crux of Hypothesis 3(1).

To evaluate whether recommendations affect beliefs, we examine how beliefs change over two consecutive belief elicitation surveys. Since goods selected for belief elicitation as per step 3 in Procedure 1 are of the same expected (high) quality regardless of whether they are recommended or not, we are able to identify a causal effect of recommendations on consumer beliefs. We test whether recommendations affect uncertainty by estimating these two specifications:

$$\sigma^b_{i,x,t} = \beta_0 + \beta_1 \sigma^b_{i,x,t-1} + \gamma_1 r_{i,x,t-1} + \epsilon_{i,x,t} \quad (3a)$$

$$\sigma^b_{i,x,t} = \beta_0 + \beta_1 \sigma^b_{i,x,t-1} + \gamma_1 r_{i,x,t-1} + \gamma_2 r_{i,x,t-1} \cdot \sigma^b_{i,x,t-1} + \epsilon_{i,x,t} \quad (3b)$$

where $\sigma^b_{i,x,t}$ denotes consumer i's level of uncertainty about good x's quality at time t, and $r_{i,x,t-1}$ a dummy indicating whether or not i was recommended x at $t-1$. The first specification tests if recommendations decrease uncertainty ($\gamma_1 < 0$), whereas the second explores the existence of heterogeneous effects of recommendations on uncertainty. In particular, one could reasonably conjecture that recommendations are all the more effective in reducing uncertainty the greater the uncertainty level one starts from ($\gamma_2 < 0$) — an idea expressing the fact that the effect of information on beliefs is decreasing in the degree of certainty. In other words, if the consumer is

	Uncertainty (t) ($\sigma^b_{i,x,t}$)	
	(1)	(2)
Recommendation	-0.004	0.017^{*}
($r_{i,x,t-1}$)	(0.004)	(0.009)
Rec. × Uncert.($t-1$)		-0.034^{**}
($r_{i,x,t-1} \cdot \sigma^b_{i,x,t-1}$)		(0.016)
Uncertainty ($t-1$)	0.650^{***}	0.667^{***}
($\sigma^b_{i,x,t-1}$)	(0.017)	(0.018)
Constant	0.199^{***}	0.189^{***}
	(0.011)	(0.012)
Observations	8,854	8,854
R^2	0.421	0.421

Clustered standard errors at the consumer level in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: The Impact of Recommendation on Beliefs: Uncertainty (Hypothesis 3(1))

Notes: This table tests whether recommendations cause lower uncertainty, as well as for heterogeneous treatment effects. Each observation corresponds to a (consumer i, good x, elicitation period t). The dependent variable is the level of uncertainty report by the consumer at period t. The sample is restricted to goods not previously used for belief elicitation and, to control for good characteristics across treatments, sampled into $S_{i,E,t}$ or $S_{i,R,t}$ (step 3 of Procedure 1).

already very certain of a good's quality, then the role for recommendation to increase certainty is limited; whereas if they are very uncertain, then there is possibly a more significant role.

It is worth mentioning that this is a conservative test of whether recommendation decreases uncertainty. As consumption is negatively related to uncertainty, and as recommended goods are more likely to lead to consumption, we are biasing against recommendation decreasing uncertainty by focusing on goods not consumed (a form of selection bias).

Our results, shown in Table 3, provide a nuanced conclusion regarding Hypothesis 3(1). Column (1) shows a null average treatment effect of recommendation on uncertainty. However, column (2) suggests that recommendations do decrease uncertainty when prior uncertainty is high with a heterogeneous effect of -0.034 between minimal and maximal uncertainty.

We then adopt an analogous strategy toward Hypothesis 3(2) and estimate the effect of recom-

	|Expect.– Predict. Qual.| (t) ($\lvert q^b_{i,x,t} - q^p_{i,x,t-1}\rvert$)		
	(1)	(2)	(3)
Recommendation	-0.015	0.046^{*}	0.050^{*}
($r_{i,x,t-1}$)	(0.012)	(0.026)	(0.027)
Rec. × Uncertainty ($t-1$)		-0.101^{**}	-0.100^{**}
($r_{i,x,t-1} \cdot \sigma^b_{i,x,t-1}$)		(0.042)	(0.043)
Rec. × |Expect. – Predict. Qual.| ($t-1$)			-0.003
($r_{i,x,t-1} \cdot \lvert q^b_{i,x,t-1} - q^p_{i,x,t-1}\rvert$)			(0.014)
|Expect.– Predict. Qual.| ($t-1$)	0.819^{***}	0.818^{***}	0.820^{***}
($\lvert q^b_{i,x,t-1} - q^p_{i,x,t-1}\rvert$)	(0.014)	(0.014)	(0.016)
Uncertainty ($t-1$)		0.090^{**}	0.090^{**}
($\sigma^b_{i,x,t-1}$)		(0.041)	(0.042)
Constant	0.250^{***}	0.198^{***}	0.196^{***}
	(0.019)	(0.027)	(0.026)
Observations	8,847	8,847	8,847
R^2	0.655	0.656	0.656

Clustered standard errors at the consumer-good level in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: The Impact of Recommendation on Beliefs: Expected Quality (Hypothesis 3(2))

Notes: This table tests whether recommendations cause the distance between expected quality and platform-predicted quality to decrease, as well as for heterogeneous treatment effects. Each observation corresponds to a (consumer i, good x, elicitation period t). The dependent variable is the absolute difference between the expected quality reported by the consumer at period t and the platform-predicted quality at period $t-1$. The sample is restricted to goods not previously used for belief elicitation and, to control for good characteristics across treatments, sampled into $S_{i,E,t}$ or $S_{i,R,t}$ (step 3 of Procedure 1).

mendations on expected quality assessments by estimating the following equations:

$$\left|q^b_{i,x,t} - q^p_{i,x,t}\right| = \beta_0 + \beta_1 \left|q^b_{i,x,t-1} - q^p_{i,x,t-1}\right| + \gamma_1 r_{i,x,t-1} + \epsilon_{i,x,t} \tag{3c}$$

$$\left|q^b_{i,x,t} - q^p_{i,x,t}\right| = \beta_0 + \beta_1 \left|q^b_{i,x,t-1} - q^p_{i,x,t-1}\right| + \gamma_1 r_{i,x,t-1} + \gamma_2 r_{i,x,t-1} \cdot \sigma^b_{i,x,t-1} + \epsilon_{i,x,t} \tag{3d}$$

$$\begin{aligned}\left|q^b_{i,x,t} - q^p_{i,x,t}\right| = \beta_0 + \beta_1 \left|q^b_{i,x,t-1} - q^p_{i,x,t-1}\right| + \beta_2 \sigma^b_{i,x,t-1} + \gamma_1 r_{i,x,t-1} + \gamma_2 r_{i,x,t-1} \cdot \sigma^b_{i,x,t-1} \\ + \gamma_3 r_{i,x,t-1} \cdot \lvert q^b_{i,x,t-1} - q^p_{i,x,t-1}\rvert + \epsilon_{i,x,t}\end{aligned} \tag{3e}$$

where $q^b_{i,x,t}$ denotes consumer i's belief about the expected quality of good x at time t and $q^p_{i,x,t}$ the platform's prediction for consumer i's realized quality assessment of good x. Equation (3c) tests for whether recommendations decrease the distance between the consumer's quality expectation for good x and the platform's consumer-specific prediction ($\gamma_1 < 0$). Equations (3d) and (3e) explore the heterogeneity of treatment effects across different prior beliefs. Similarly to before, we conjecture that the more uncertain a consumer is, the greater the effect of recommendations — again a prediction implied by our theoretical framework.

In column (1) of Table 4 we report a negative result: we find that recommendations do not, on average, decrease the distance between consumer expectations and platform predictions.[18] However, we find significant heterogeneous treatment effects of recommendations with respect to the prior level of uncertainty. In columns (2) and (3) we find a negative and significant interaction between consumer i's prior uncertainty $\sigma^b_{i,x,t-1}$ about good x's quality and recommendation. This implies expected quality drifts more toward the platform's prediction the more uncertain the consumer was prior to obtaining the recommendation.

Recommendations, Information Acquisition, and Beliefs. While recommendation provides a direct informational gain — either implicitly through the fact that a good is stated to be recommended or explicitly through the platform-predicted rating — it also can have an indirect impact. We explore how recommendation induces *additional* information acquisition by the consumer and how this further impacts beliefs. In our context, consumers can explicitly click-through to learn more about a movie from the homepage or explicitly search for movies, terms, or tags that link to this detail page. We first characterize what drives this exploration as well as the causal effect of recommendation on fostering this additional information acquisition. Then, we study how this impacts the resulting changes in consumer beliefs. In other words, this section examines the alternative indirect mechanism to explain how recommendation affects consumption via an informational channel:

$$\text{Recommendation} \longrightarrow \text{Additional Information Acquisition} \longrightarrow \Delta \text{ Beliefs.}$$

We define $\text{Info}_{i,x,t}$ and $\text{Info}_{i,x}$ as whether consumer i viewed the details page for movie x in time period t or at any point in the study period, respectively. The first two specifications we consider are identical to the empirical strategy used for identifying the causal impact of recommendation on consumption in Section 4. In particular, we consider the average treatment effect estimated over the full set of goods in the control, exposure-only, and recommendation sets, and then restricting the sample to exposed strata, estimating the equation

$$\text{Info}_{i,x} = \beta_0 + \beta_1\{x \in X_{i,E} \cup X_{i,R}\} + \beta_2\{x \in X_{i,R}\} + \epsilon_{i,x} \tag{4a}$$

[18] Note that this result is not at odds with Cosley et al. (2003) as we are eliciting ratings for unseen as opposed to seen movies.

	Information Acquisition				
	Info. Acq. ($\text{Info}_{i,x}$)		Info. Acq. $\geq t-1$ ($\text{Info}_{i,x,\geq t-1}$)		Info. Acq. $[t-1,t]$ ($\text{Info}_{i,x,[t-1,t]}$)
	(1)	(2)	(3)	(4)	(5)
Exposure Set	-0.0001	-0.001			
($\mathbb{1}(x \in X_{i,E} \cup X_{i,R})$)	(0.0002)	(0.001)			
Recommendation Set	0.011***	0.041***			
($\mathbb{1}(x \in X_{i,R})$)	(0.001)	(0.005)			
Recommendation			0.045***	0.045***	0.024***
($r_{i,x}$)			(0.005)	(0.005)	(0.003)
Exp. Quality				0.013***	
($q^b_{i,x}$)				(0.003)	
Uncertainty				-0.013	
($\sigma^b_{i,x}$)				(0.009)	
Constant	0.007***	0.016***	0.012***	-0.018**	0.003***
	(0.001)	(0.002)	(0.003)	(0.008)	(0.001)
Observations	754,500	93,321	8,847	8,847	8,847
R^2	0.002	0.013	0.015	0.020	0.010

Clustered standard errors at the consumer level in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Recommendations Impact Information Acquisition

Notes: This table tests whether recommendations cause consumers to acquire more information about the recommended goods. Each column displays the average treatment effect of exposure and recommendation on information acquisition (visit a movie's page or watch its trailer); $\text{Info}_{i,x}$ (columns (1)-(2)), $\text{Info}_{i,x,\geq t-1}$ (columns (3)-(4)), and $\text{Info}_{i,x,[t-1,t]}$ (column (5)) are binary variables that correspond, respectively, to consumer i having ever acquired information about good x, having acquired it after $t-1$, and between $t-1$ and t. Each observation corresponds to a pair (consumer i, good x). Columns (1)-(2) correspond to specification (4a); columns (3), (4), and (5) to specifications (4b), (4c), and (4d), respectively. For column (1), we include all consumers i and all goods x in the consumer-specific control, exposure-only, and recommendation sets. In column (2), for each consumer we include the goods to which they were exposed through the belief elicitation survey, and all the goods in the same consumer-specific stratum. Columns (3)-(5) restrict to goods exposed to the consumer through belief elicitation and sampled as per step 3 in Procedure 1.

While this specification provides clear causal estimates of the average treatment effect, it neglects the time when the information acquisition occurs as well as how it is driven by consumers' beliefs.

	Uncertainty (t) ($\sigma^b_{i,x,t}$)				
	(1)	(2)	(3)	(4)	(5)
Info. Acq. $[t-1,t]$	-0.055^{**}	0.016	-0.054^{**}	-0.054^{**}	0.016
($\text{Info}_{i,x,[t-1,t]}$)	(0.022)	(0.043)	(0.022)	(0.022)	(0.043)
Info. Acq. $[t-1,t]$ × Uncert.($t-1$)		-0.121^{*}			-0.121^{*}
($\text{Info}_{i,x,[t-1,t]} \cdot \sigma^b_{i,x,t-1}$)		(0.073)			(0.073)
Uncertainty ($t-1$)	0.650^{***}	0.652^{***}	0.650^{***}	0.667^{***}	0.652^{***}
($\sigma^b_{i,x,t-1}$)	(0.017)	(0.017)	(0.017)	(0.018)	(0.017)
Recommendation			-0.002	0.018^{*}	-0.002
($r_{i,x,t-1}$)			(0.004)	(0.009)	(0.004)
Rec. × Uncertainty ($t-1$)				-0.034^{**}	
($r_{i,x,t-1} \cdot \sigma^b_{i,x,t-1}$)				(0.016)	
Constant	0.198^{***}	0.197^{***}	0.199^{***}	0.189^{***}	0.198^{***}
	(0.011)	(0.011)	(0.011)	(0.012)	(0.011)
Observations	8,854	8,854	8,854	8,854	8,854
R^2	0.421	0.422	0.421	0.422	0.422

Table 6: Information Acquisition and Changes in Beliefs: Uncertainty

Clustered standard errors at the consumer level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table examines if information acquisition explains change in consumers' beliefs. Columns (1)-(5) in correspond to variations on specifications (5a)-(5b). Observations correspond to pairs of (consumer i, good x) restrict to goods exposed to the consumer through belief elicitation and sampled as per step 3 in Procedure 1 to control for good quality selection; we use the first belief elicitation of good x's quality from consumer i.

We estimate three additional specifications to account for this:

$$\text{Info}_{i,x,\geq t-1} = \beta_0 + \beta_1 r_{i,x,t-1} + \epsilon_{i,x,t} \tag{4b}$$

$$\text{Info}_{i,x,\geq t-1} = \beta_0 + \beta_1 r_{i,x,t-1} + \beta_2 \sigma^b_{i,x,t-1} + \beta_3 q^b_{i,x,t-1} + \epsilon_{i,x,t} \tag{4c}$$

$$\text{Info}_{i,x,[t-1,t]} = \beta_0 + \beta_1 r_{i,x,t-1} + \epsilon_{i,x,t} \tag{4d}$$

The results are reported in Table 5. Columns (1) and (2) indicate that recommendation induces, respectively, a 1.1 percentage points and 4.1 p.p. increase in the probability of acquiring information over exposure. This increase in the likelihood of information acquisition only comes from recommendation as there is a null effect of the impact of exposure. Furthermore, note that the baselines are 0.7 and 1.6 p.p. respectively, indicating that the effect size is quite substantial.

We then rely exclusively on goods that were selected for belief-elicitation as per step 3 in Procedure 1, enabling a causal interpretation of the effect of recommendation of the specifications (4b)-(4d). Columns (3) and (4) display the estimated effect restricting focus to information ac-

	|Expect.– Predict. Qual.| (t) ($\lvert q^b_{i,x,t} - q^p_{i,x,t-1}\rvert$)				
	(1)	(2)	(3)	(4)	(5)
Info. Acq. $[t-1,t]$	-0.189^{***}	-0.030	-0.185^{***}	-0.183^{***}	-0.027
($\text{Info}_{i,x,[t-1,t]}$)	(0.056)	(0.077)	(0.056)	(0.056)	(0.077)
Info. Acq. $[t-1,t]$ × |Expect. – Predict. Qual.| $(t-1)$		-0.121			-0.120
($\text{Info}_{i,x,[t-1,t]} \cdot \lvert q^b_{i,x,t-1} - q^p_{i,x,t-1}\rvert$)		(0.074)			(0.074)
|Expect. – Predict. Qual.| $(t-1)$	0.819^{***}	0.821^{***}	0.819^{***}	0.818^{***}	0.821^{***}
($\lvert q^b_{i,x,t-1} - q^p_{i,x,t-1}\rvert$)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Recommendation			-0.015	0.046^{*}	0.050^{*}
($r_{i,x,t-1}$)			(0.011)	(0.027)	(0.011)
Rec. × Uncertainty $(t-1)$				-0.100^{**}	
($r_{i,x,t-1} \cdot \sigma^b_{i,x,t-1}$)				(0.042)	
Uncertainty $(t-1)$				0.089^{**}	
($\sigma^b_{i,x,t-1}$)				(0.041)	
Constant	0.246^{***}	0.243^{***}	0.251^{***}	0.199^{***}	0.248^{***}
	(0.018)	(0.018)	(0.019)	(0.027)	(0.019)
Observations	8,847	8,847	8,847	8,847	8,847
R^2	0.656	0.656	0.656	0.656	0.656

Table 7: Information Acquisition and Changes in Beliefs: Expected Quality

Clustered standard errors at the consumer level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table examines if information acquisition explains change in consumers' beliefs. Columns (1)-(5) correspond to variations on specifications (5c)-(5d). Observations correspond to pairs of (consumer i, good x) restrict to goods exposed to the consumer through belief elicitation and sampled as per step 3 in Procedure 1 to control for good quality selection; we use the first belief elicitation of good x's quality from consumer i.

quired after elicitation and find similar effect sizes to the previous specifications with the result being robust to controlling for the underlying beliefs. Finally, column (5) shows that the result is robust to focusing on movie details pages viewed between elicitations.

Having established that recommendation induces additional information acquisition, we characterize the extent to which information acquisition induces changes in beliefs. We use similar specifications and sample selection as previously for studying how information acquisition influences the underlying degree of uncertainty:

$$\sigma^b_{i,x,t} = \beta_0 + \beta_1 \text{Info}_{i,x,[t-1,t]} + \beta_2 \sigma^b_{i,x,t-1} + \beta_3 r_{i,x,t-1} + \epsilon_{i,x,t} \tag{5a}$$

$$\sigma^b_{i,x,t} = \beta_0 + \beta_1 \text{Info}_{i,x,[t-1,t]} + \beta_2 r_{i,x,t-1} + \beta_3 \sigma^b_{i,x,t-1} \times \text{Info}_{i,x,[t-1,t]} + \beta_4 \sigma^b_{i,x,t-1} + \epsilon_{i,x,t} \tag{5b}$$

The results are reported in Table 6. Column (1) shows that viewing the details page reduces uncertainty by 0.055. However, column (2) reveals that this average effect is heterogeneous depending on the initial level of uncertainty: the impact is larger when prior uncertainty is higher, entailing a 0.121 reduction in uncertainty when the initial belief indicates maximal uncertainty. Columns (3)-(5) validate that both the sign and magnitude of the effects are robust to controlling for whether the good was recommended or not. Furthermore, column (4) can be taken as showing that recommendations do decrease uncertainty when prior uncertainty is high, even when controlling for whether the consumer acquired additional information.

Finally, we study a similar specification to understand how viewing the details page impacts the expected quality:

$$\left|q^b_{i,x,t} - q^p_{i,x,t-1}\right| = \gamma_0 + \gamma_1 \text{Info}_{i,x,[t-1,t]} + \beta_1 \left|q^b_{i,x,t-1} - q^p_{i,x,t-1}\right| + \epsilon_{i,x,t} \tag{5c}$$

$$\begin{aligned}\left|q^b_{i,x,t} - q^p_{i,x,t-1}\right| = \gamma_0 + \gamma_1 \text{Info}_{i,x,[t-1,t]} + \gamma_2 \text{Info}_{i,x,[t-1,t]} \cdot \left|q^b_{i,x,t-1} - q^p_{i,x,t-1}\right| \\ + \beta_1 \left|q^b_{i,x,t-1} - q^p_{i,x,t-1}\right| + \beta_2 r_{i,x,t-1} + \epsilon_{i,x,t}\end{aligned} \tag{5d}$$

Table 18 summarizes our findings. Columns (1) and (2) indicate that viewing the details page similarly induces a level effect on the resulting expected quality, while there is no estimated heterogeneity across the initial distance from the predicted rating. Columns (3)-(5) validate that the same results hold when we control for recommendation. As for uncertainty, column (4) also highlights that the result that recommendations decrease the distance between the consumer's expected quality and the platform's prediction when prior uncertainty is high holds even after controlling for additional information acquisition.

Overall, these results suggest two distinct avenues through which recommendations entail informational gains that shifts expected quality and reduces uncertainty: directly through the recommendation itself and indirectly, by inducing additional information acquisition.

6. Spillover Effects of Recommendation

The collected data also enables us to explore the correlation structure of beliefs and evaluate the presence of informational spillovers.

We first use cross-sectional data on beliefs to assess Hypothesis 4, that is, if beliefs about goods are spatially correlated, with beliefs about goods closer in the attribute space being more similar than to those with more dissimilar attributes. After establishing this, we evaluate Hypothesis 5 by using cross-sectional data on consumption, as well as by exploiting the panel nature of our data and the repeated elicitation of the same good over time.

	Abs. Difference between Goods	
	in Uncertainty	in Expected Quality
	$\lvert\sigma^b_{i,x} - \sigma^b_{i,y}\rvert$	$\lvert q^b_{i,x} - q^b_{i,y}\rvert$
	(1)	(2)
Distance between Goods	0.226^{***}	0.902^{***}
($\lVert x - y\rVert$)	(0.025)	(0.084)
Consumer FEs	Yes	Yes
Good FEs	Yes	Yes
Observations	243,524	243,524
R^2	0.302	0.259

Clustered standard errors at the consumer level in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Beliefs' Spatial Correlation (Hypothesis 4)

Notes: This table displays the relationship between good attribute distance and distance in expected quality and uncertainty. Each observation is a tuple (consumer i, good x, good y). The sample is restricted to the first instance of each tuple.

6.1. Spatially Correlated Beliefs

We test if the similarity between beliefs about two goods is correlated with their attribute similarity; i.e. if beliefs about goods' quality are more similar the more similar goods are in attributes (Hypothesis 4). For instance, intuition would suggest that beliefs about *The Godfather* are closer to those about *The Godfather: Part II* than about *Frozen*. We consider the following cross-sectional regressions for consumer i and goods x and y:

$$\left|\sigma^b_{i,x} - \sigma^b_{i,y}\right| = \beta_1 \lVert x - y\rVert + FE_i + FE_x + FE_y + \epsilon_{i,x,y} \tag{6a}$$

$$\left|q^b_{i,x} - q^b_{i,y}\right| = \beta_1 \lVert x - y\rVert + FE_i + FE_x + FE_y + \epsilon_{i,x,y} \tag{6b}$$

where FE_i, FE_x, FE_y denote consumer, good x, and good y fixed effects, and the remaining notation is consistent with previous empirical specifications. Hypothesis 4 implies $\beta_1 > 0$, that is, the closer two goods are in the attribute space (smaller distance), the more similar beliefs about their quality. The results, reported in Table 8, confirm the conjecture of a strong association between beliefs and distance.

6.2. Informational Spillovers

This section tests for the existence of informational spillovers by examining how prior consumption and recommendation relates to beliefs of unconsumed goods. If beliefs are spatially correlated, then information about good x may also be informative about other goods, especially

	Uncertainty (t) ($\sigma^b_{i,x,t}$)	
	(1)	(2)
Avg. Distance to Consumed Goods	6.466***	
($\sum_{y \in C_{i,t-1}} \|y - x\| / \|C_{i,t-1}\|$)	(0.679)	
Distance to Closest Consumed Good		3.134***
($\min_{y \in C_{i,t-1}} \|y - x\|$)		(0.351)
Consumer FEs	Yes	Yes
Good FEs	Yes	Yes
Observations	18,006	18,006
R^2	0.572	0.572

(a) Consumption Spillovers

	Uncertainty (t) ($\sigma^b_{i,x,t}$)	
	(1)	(2)
Avg. Distance to Consumed Goods	0.695	
($\sum_{y \in R_{i,t-1}} \|y - x\| / \|R_{i,t-1}\|$)	(0.938)	
Distance to Closest Consumed Good		−0.272
($\min_{y \in R_{i,t-1}} \|y - x\|$)		(0.500)
Consumer FEs	Yes	Yes
Good FEs	Yes	Yes
Observations	6,769	6,769
R^2	0.705	0.705

(b) Recommendation Spillovers

Table 9: Information Spillover Effects (Hypothesis 5)

Clustered standard errors at the consumer level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: This table examines if consumption and recommendation induce spillovers on beliefs. Subtable 9a and Subtable 9b correspond to specifications (7a) and (7b). Columns (1) and (2) use, respectively, the average and minimum of the distance between good x and goods that were consumed (in Subtable 9a) or recommended (in Subtable 9b). In both subtables, observations correspond to pairs of (consumer i, good x). For Subtable 9b, we restrict focus to goods from the exposure-only bin; we use the first belief elicitation of good x's quality from consumer i for both regressions. For Subtable 9a, we use the full history of consumption, whereas for Subtable 9b we consider recommendations during the experimental intervention.

those that are more similar to good x. We consider two sources of information: consumption and recommendation. Specifically, we consider whether having previously consumed or been recommended more similar goods to good x is associated with a lower uncertainty about good x's quality (Hypothesis 5). We rely on the exogenous randomness in the selection of goods utilized to elicit beliefs about, though note that our analyses here are tests of association.

Let $C_{i,t}$ and $R_{i,t}$ denote the set of goods consumer i, respectively, consumed and was recommended prior to time t. For each of these, $A = C_{i,t}, R_{i,t}$, we consider both the average distance to x, i.e. $\sum_{y\in A} \|y-x\|/|A|$, and the minimum distance, $\min_{y\in A} \|y-x\|$. The former allows us to capture the average effect over the full set of previously consumed/recommended goods, while the latter measures the effect of the most similar one. We use the following specifications in order to test for informational spillovers in the form of lower uncertainty (Hypothesis 5):

$$\sigma^b_{i,x,t} = \beta_1 s(x, C_{i,t-1}) + FE_i + FE_x + \epsilon_{i,x,t} \tag{7a}$$

$$\sigma^b_{i,x,t} = \beta_1 s(x, R_{i,t-1}) + FE_i + FE_x + \epsilon_{i,x,t} \tag{7b}$$

where $s(x, A)$ corresponds to one of the two measures of (dis)similarity, average or minimum distance between x and A. Hypothesis 5 would then imply $\beta_1 > 0$. We again restrict focus to the first time a consumer's beliefs about a good are elicited in order to minimize any possible changes in beliefs arising from the experimental interventions; only goods recommended during the intervention period are considered.

We report the results in Table 9. Subtable 9a shows that the uncertainty is increasing in both the average and minimum distance to previously consumed goods. In contrast, we fail to detect informational spillovers from recommendation (Subtable 9b).

7. Conclusion

In this paper we report the results of a field experiment aimed at understanding the mechanisms that drive the impact that recommender systems have on consumption choices. We monitored consumers for a period of over 6 months and collected the set of movies they consumed as well as elicited their beliefs about unseen movies. Our randomization was at the good-level, so that we can directly identify the effects that recommendation has on consumption choices. We find that exposing consumers to a good significantly increases its consumption probability relative to a held-out control group, and that recommendations double the effect of exposure on consumption. We use the collected belief data to show that recommendations causally affect consumers' beliefs, and these, in turn, guide consumption choices. Furthermore, we establish that beliefs are spatially correlated and provide suggestive evidence of informational spillover effects.

Our results provide insight into several higher-level debates about the broader impact of recommender systems. First, contrary to several popular modeling approaches, the informational role of recommender systems emerges as a first-order aspect of their influence. Second, since recommendations provide information, shifting spatially correlated beliefs and guiding consumption, they may have informationally-driven dynamic consequences. In other words, if recommendations have informational value and are therefore prima facie beneficial to consumers, they can also be used to steer consumer behavior toward parts of the product space supplied by or more profitable for the platform.

As argued in Aridor, Gonçalves and Sikdar (2020), explicitly collecting data on consumer beliefs enables a more fine-grain understanding of the impact of these systems and can shed light on many social and economic questions surrounding them. While we view the domain of movie recommendation to be fruitful and illuminating about the general mechanisms that drive the influence of recommender systems, an exciting avenue for future work is collecting similar types of data and running similar experimental designs across the different domains where recommender systems are deployed. In particular, while our experiment was conducted on a platform with plausibly unbiased information provision, it would be interesting for future work to characterize how the magnitude of information provision is degraded by the degree of platform self-preferencing. Overall, we believe our approach and results are insightful for understanding the broader economic implications of recommender systems and for shaping the evaluation of their impact.

References

Ackerberg, Daniel A. 2003. "Advertising, learning, and consumer choice in experience good markets: an empirical examination." *International Economic Review*, 44(3): 1007–1040.

Aridor, Guy, and Duarte Gonçalves. 2022. "Recommenders' Originals: The Welfare Effects of the Dual Role of Platforms as Producers and Recommender Systems." *International Journal of Industrial Organization*, 102845.

Aridor, Guy, Duarte Gonçalves, and Shan Sikdar. 2020. "Deconstructing the Filter Bubble: User Decision-Making and Recommender Systems." *Fourteenth ACM Conference on Recommender Systems*, 82–91.

Armstrong, Mark, John Vickers, and Jidong Zhou. 2009. "Prominence and consumer search." *The RAND Journal of Economics*, 40(2): 209–233.

Bergemann, Dirk, and Deran Ozmen. 2006. "Optimal pricing with recommender systems." *Proceedings of the 7th ACM Conference on Electronic Commerce*, 43–51.

Bourreau, Marc, and Germain Gaudin. 2022. "Streaming platform and strategic recommendation bias." *Journal of Economics & Management Strategy*, 31(1): 25–47.

Brynjolfsson, Erik, Yu Hu, and Duncan Simester. 2011. “Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales.” *Management Science*, 57(8): 1373–1386.

Calvano, Emilio, Giacomo Calzolari, Vincenzo Denicolò, and Sergio Pastorello. 2022. “Artificial Intelligence, product recommendations and concentration.” *Working Paper.*

Chaney, Allison, Brandon Stewart, and Barbara Engelhardt. 2018. “How algorithmic confounding in recommendation systems increases homogeneity and decreases utility.” *Proceedings of the 12th ACM Conference on Recommender Systems*, 224–232.

Chen, Yan, F Maxwell Harper, Joseph Konstan, and Sherry Xin Li. 2010. “Social comparisons and contributions to online communities: A field experiment on movielens.” *American Economic Review*, 100(4): 1358–98.

Cheung, Paul, and Yusufcan Masatlioglu. 2021. “Decision Making with Recommendation.”

Che, Yeon-Koo, and Johannes Hörner. 2017. “Recommender systems as mechanisms for social learning.” *Quarterly Journal of Economics*, 133(2): 871–925.

Claussen, Jörg, Christian Peukert, and Ananya Sen. 2019. “The Editor vs. the Algorithm: Targeting, Data and Externalities in Online News.”

CMA, United Kingdom. 2021. “Algorithms: How They Can Reduce Competition and Harm Consumers.”

Cosley, Dan, Shyong K. Lam, Istvan Albert, Joseph A. Konstan, and John Riedl. 2003. “Is seeing believing? How recommender system interfaces affect users’ opinions.” *Proceedings of the SIGCHI conference on Human factors in computing systems*, 585–592.

Das, Abhinandan S, Mayur Datar, Ashutosh Garg, and Shyam Rajaram. 2007. “Google news personalization: scalable online collaborative filtering.” *Proceedings of the 16th international conference on World Wide Web*, 271–280.

De Corniere, Alexandre, and Greg Taylor. 2019. “A model of biased intermediation.” *The RAND Journal of Economics*, 50(4): 854–882.

Donnelly, Robert, Ayush Kanodia, and Ilya Morozov. 2022. “Welfare Effects of Personalized Rankings.” *Available at SSRN 3649342.*

Ekstrand, Michael D., Michael Ludwig, Jack Kolb, and John T. Riedl. 2011. “LensKit: a modular recommender framework.” *Proceedings of the fifth ACM conference on Recommender systems*, 349–350.

Fleder, Daniel, and Kartik Hosanagar. 2009. “Blockbuster culture’s next rise or fall: The impact of recommender systems on sales diversity.” *Management Science*, 55(5): 697–712.

Freyne, Jill, Michal Jacovi, Ido Guy, and Werner Geyer. 2009. “Increasing engagement through early recommender intervention.” *Proceedings of the third ACM conference on Recommender systems*, 85–92.

Grossman, Gene M., and Carl Shapiro. 1984. "Informative advertising with differentiated products." *The Review of Economic Studies*, 51(1): 63–81.

Hagiu, Andrei, and Bruno Jullien. 2011. "Why do intermediaries divert search?" *The RAND Journal of Economics*, 42(2): 337–362.

Harper, F. Maxwell, and Joseph A. Konstan. 2015. "The MovieLens datasets: History and context." *ACM Transactions on Interactive Intelligent Systems (TIIS)*, 5(4): 1–19.

Hodgson, Charles, and Gregory Lewis. 2019. "You Can Lead a Horse to Water: Spatial Learning and Path Dependence in Consumer Search." *Proceedings of the 2019 ACM Conference on Economics and Computation*, 215.

Holtz, David, Ben Carterette, Praveen Chandar, Zahra Nazari, Henriette Cramer, and Sinan Aral. 2020. "The engagement-diversity connection: Evidence from a field experiment on spotify." *Proceedings of the 21st ACM Conference on Economics and Computation*, 75–76.

Honka, Elisabeth, Ali Hortaçsu, and Maria Ana Vitorino. 2017. "Advertising, consumer awareness, and choice: Evidence from the US banking industry." *The RAND Journal of Economics*, 48(3): 611–646.

Hosanagar, Kartik, Daniel Fleder, Dokyun Lee, and Andreas Buja. 2013. "Will the global village fracture into tribes? Recommender systems and their effects on consumer fragmentation." *Management Science*, 60(4): 805–823.

Lee, Changhwa. 2022. "Optimal Recommender System Design." *Unpublished manuscript, University of Pennsylvania.*

Lee, Dokyun, and Kartik Hosanagar. 2019. "How do recommender systems affect sales diversity? A cross-category investigation via randomized field experiment." *Information Systems Research*, 30(1): 239–259.

Lee, Gunhaeng, and Julian Wright. 2021. "Recommender systems and the Value of User Data." *Working Paper.*

Meurer, Michael, and Dale O Stahl II. 1994. "Informative advertising and product match." *International Journal of Industrial Organization*, 12(1): 1–19.

Nguyen, Tien T., Pik-Mai Hui, F. Maxwell Harper, Loren Terveen, and Joseph A. Konstan. 2014. "Exploring the filter bubble: the effect of using recommender systems on content diversity." *Proceedings of the 23rd International Conference on World Wide Web*, 677–686.

Rossi, Michelangelo. 2021. "Quality Disclosures and Disappointment: Evidence from the Academy Awards." *Proceedings of the 22nd ACM Conference on Economics and Computation*, 790–791.

Sahni, Navdeep S., and Harikesh S. Nair. 2020. "Does advertising serve as a signal? Evidence from a field experiment in mobile search." *The Review of Economic Studies*, 87(3): 1529–1564.

Schwemer, Sebastian Felix. 2021. "Recommender Systems in the EU: from Responsibility to

Regulation?" *Proceedings of the FAccTRec Workshop*, 21.

Senecal, Sylvain, and Jacques Nantel. 2004. "The influence of online product recommendations on consumers' online choices." *Journal of Retailing*, 80(2): 159–169.

Tsai, Yi-Lin, and Elisabeth Honka. 2021. "Informational and Noninformational Advertising Content." *Marketing Science*, 40(6): 1030–1058.

Van Alstyne, Marshall, and Erik Brynjolfsson. 2005. "Global village or cyber-balkans? Modeling and measuring the integration of electronic communities." *Management Science*, 51(6): 851–868.

Vig, Jesse, Shilad Sen, and John Riedl. 2012. "The tag genome: Encoding community knowledge to support novel interaction." *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 2(3): 1–44.

Zhou, Renjie, Samamon Khemmarat, and Lixin Gao. 2010. "The impact of YouTube recommendation system on video views." *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, 404–410.

Appendix

A. Patterns in Belief Data

In this appendix we demonstrate the belief data we collect not only exhibits reasonable patterns, but also is informative about the resulting good quality. In other words, we provide evidence that consumers have well-formed beliefs about movies and that survey-based measures can accurately capture them.

First, we show that consumers' beliefs are an *unbiased* statistic for their quality assessments after consumption, arguably settling any question about the validity of the expected quality measure. We estimate

$$q_{i,x} = \beta_1 q^b_{i,x,t} + \epsilon_{i,x}$$

where $q_{i,x}$ denotes the realized quality of good x for consumer i, and, in order to capture the initial beliefs of consumers before any intervention, $q^b_{i,x,t}$ denotes the consumers' first belief elicitation about good x for consumer i. The results, in column (1) of Table 10, show the estimated coefficient β_1 is a precisely estimated 1: prior beliefs of consumers are on average correct.

Second, we show the (Euclidean) distance between expected quality assessment and the realized quality is increasing in the reported uncertainty level. Specifically, we estimate

$$\left| q_{i,x} - q^b_{i,x,t} \right| = \beta_0 + \beta_1 \sigma^b_{i,x,t} + \epsilon_{i,x,t}.$$

In order to capture the initial beliefs of consumers before any intervention, we use the first elicitation of the belief of a good for a given consumer. Column (2) of Table 10 reports $\beta_1 > 0$, a positive relationship between uncertainty and the resulting difference, validating that larger uncertainty results in less aligned belief and actual quality.

We provide additional checks on the belief data by linking it to additional data sources. For instance, we would expect that consumers are more likely to have higher expected quality for popular movies and be more sure about what they think about them, since they are more likely to have been exposed at some point to these movies either on or off the platform. Another reasonable conjecture would be that consumers are more sure of their expected quality about sequels.

We evaluate this by constructing a measure of popularity from the MovieLens data and joining our data with IMDb data to identify which movies are sequels. We proxy popularity from the aggregated MovieLens data by the overall number of ratings the movie has on the platform. We run the following regression to assess the association between consumers' reported beliefs and

	Realized Quality ($q_{i,x}$) (1)	|Expect.– Realized. Qual.| ($\lvert q^b_{i,x,t} - q_{i,x}\rvert$) (2)
Exp. Quality	1.007^{***}	
($q^b_{i,x}$)	(0.011)	
Uncertainty		0.555^{***}
($\sigma^b_{i,x}$)		(0.119)
Constant		0.308^{***}
		(0.054)
Observations	402	402
R^2	0.459	0.065

Clustered standard errors at the consumer level in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 10: Expected Quality and Uncertainty

Notes: This table demonstrates the sensible patterns in the belief data. The sample is restricted to the first elicitation of consumer i's beliefs about good x. Column (1) displays the regression of the realized quality on the expected quality. Column (2) displays the regression of the distance between the expected and realized quality on the measured consumer uncertainty.

popularity:

$$\sigma^b_{i,x,t} = \beta_1 \log(\text{number ratings}_x) + \beta_2 \text{Sequel}_x + \text{FE}_i + \epsilon_{i,x,t} \tag{8a}$$

$$q^b_{i,x,t} = \beta_1 \log(\text{number ratings}_x) + \beta_2 \text{Sequel}_x + \text{FE}_i + \epsilon_{i,x,t} \tag{8b}$$

where the notation is similar to the previous specifications. In order to isolate the possible role of the experimental intervention in modifying beliefs, we restrict focus to the first belief elicitation of a given good for each consumer.

Table 11 displays the results, confirming the conjectures. As expected, greater popularity is associated with lower uncertainty as well as with higher expected quality. Moreover, sequels are associated with significantly lower uncertainty about movie quality.

Finally, we explore the relationship between expected quality and uncertainty. Figure 2 plots the mean uncertainty level for each expected quality and shows an inverted U-shaped relationship. This accords with the intuition that consumers are more sure of extreme quality assessments (i.e. close to 0 or 5 stars) relative to more moderate ones (i.e. 3 star).

	Uncertainty (t) ($\sigma^b_{i,x,t}$)		Expected Quality (t) ($q^b_{i,x,t}$)	
	(1)	(2)	(3)	(4)
Popularity	-0.026^{***}	-0.035^{***}	0.114^{***}	0.127^{***}
(log(number ratings$_x$))	(0.001)	(0.002)	(0.004)	(0.006)
Sequel		-0.073^{***}		-0.028
(Sequel$_x$)		(0.013)		(0.057)
Consumer FEs	Yes	Yes	Yes	Yes
Observations	21,281	9,502	21,281	9,502
R^2	0.411	0.447	0.435	0.450

Clustered standard errors at the subject level in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 11: Popularity and Uncertainty

Notes: This table displays the relationship between consumer expected quality and log(number of ratings) on the MovieLens platform and whether the movie is a sequel or not as the covariates. The sample is restricted to the first elicitation of consumer i's beliefs about good x.

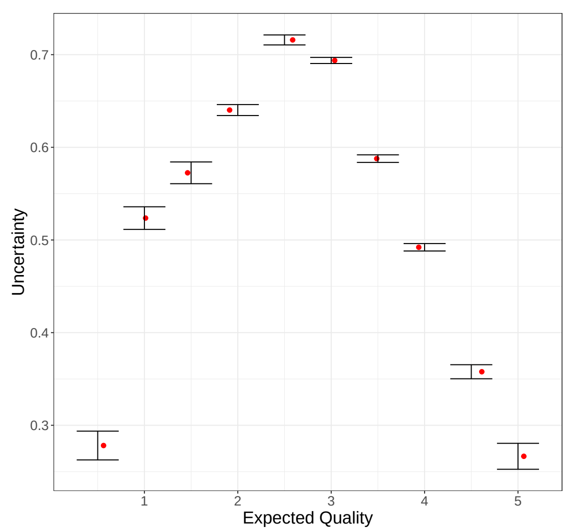


Figure 2: Uncertainty by Expected Quality

Notes: This figure shows the expected quality on the x-axis and the associated conditional average uncertainty score on the y-axis as well as the associated 95% confidence interval. The sample is restricted to the first belief elicitation.

	Pairwise Distance ($\|x-y\|$)	
	(1)	(2)
Same Franchise	-0.115^{***}	-0.137^{***}
($\text{franchise}_{x,y}$)	(0.005)	(0.007)
Constant	0.269^{***}	
	(0.001)	
Good x FE	No	Yes
Good y FE	No	Yes
Observations	19,307,597	19,307,597
R^2	0.0001	0.726

Clustered standard errors at each good level in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 12: Movie Franchises and Attribute Distance

Notes: This regression is estimated on the dataset of all pairs of goods x and y in the MovieLens dataset with tag data. Same Franchise equals 1 if the goods are part of the same franchise.

B. Validating Similarity Measures

We provide a validation exercise to show that the similarity measure captures movies that we intuitively expect to be similar to each other in content — movie franchises. We join our dataset of movies with a dataset of movie franchises from Opus Data[19] (i.e. The Godfather, Harry Potter, The Matrix, John Wick, etc.) and assess whether, according to our similarity measure, movies which are part of a franchise are more similar to those in the same franchise than to other movies using the following specification:

$$\|x-y\| = \beta_1 \text{franchise}_{x,y} + FE_x + FE_y + \epsilon_{x,y}$$

where FE_x and FE_y denote good fixed effects and $\text{franchise}_{x,y}$ is an indicator for whether movies x and y are part of the same franchise.[20] In order to make the exercise computationally feasible, we randomly sample 15,000 movies from the set of movies and conduct the exercise over this set. The results are displayed in Table 12 and validate that movies in the same franchise are substantially closer than others: as column (1) indicates the baseline distance for movies not in the same franchise is 0.269, while movies in the same franchise on average have a distance of 0.154.

[19]Data available from www.opusdata.com.
[20]Note that we cannot use the same sequel data that we used before from IMDb since this data only tells us whether a movie is a sequel and not which movie it is a sequel of.

Online Appendix

C. Robustness Checks

C.1. Recommendation and Beliefs

In this section we extend the specifications in Tables 1, 3, 4, 5, 6, and 18 to include fixed effects and report the results.

We opted for not to include fixed effects in estimating the causal effects of recommendations throughout since (1) our causal identification of treatment effects of recommendation does not require controlling for fixed effects, and (2) especially because, instead of improving estimates of treatment effects, including fixed effects would be expected to negatively impact the precision of our estimates by removing meaningful variation in belief data. However, when considering tests of association we include these fixed effects when relevant.

As our belief elicitation is drawn from consumer-specific sets (recall recommendations are consumer-specific), it includes goods that align with each consumer's idiosyncratic tastes. This results in the fact that, for most goods, beliefs about their quality are elicited from a single consumer. These are the most relevant goods when examining the effect of recommendation on beliefs, as consumers are potentially less exposed to them outside of the platform. Then, good fixed effects for specifications referring to belief data, instead of improving the quality of our estimates, would remove meaningful variation coming from consumer-specific goods. A similar issue occurs with consumer fixed effects. Indeed, on average, consumers completed 4 belief surveys and by including consumer fixed effects we would be absorbing a significant fraction of the belief variation for the majority of the consumers.

The above being said, we note that the results are consistent with the tables reported in the main text.

	Consumption $(c_{i,x})$					
	(Rating)	(Robust)	(Rating)	(Robust)	(Rating)	(Robust)
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure Set	0.003***	0.001**	0.013***	0.004***		
$(1(x \in X_{i,E} \cup X_{i,R}))$	(0.0003)	(0.0002)	(0.001)	(0.001)		
Recommendation Set	0.006***	0.008***	0.013***	0.017***		
$(1(x \in X_{i,R}))$	(0.001)	(0.001)	(0.002)	(0.002)		
Recommendation					0.012***	0.011***
$(r_{i,x})$					(0.003)	(0.003)
Consumer FEs	Yes	Yes	Yes	Yes	Yes	Yes
Good FEs	Yes	Yes	Yes	Yes	No	No
Observations	754,500	750,978	93,321	92,001	11,852	11,686
R^2	0.051	0.051	0.149	0.153	0.109	0.110

Clustered standard errors at the consumer level in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 13: The Impact of Recommendation on Consumption (Hypothesis 1); with Fixed Effects

Notes: This table tests if exposure and recommendation impact consumption probability. Each column displays the average treatment effect of exposure and recommendation on consumption for the different measures of consumption. Each observation corresponds to a pair (consumer i, good x). Columns (1)-(4) correspond to specification (1a) and columns (5)-(6) to specification (1b). For columns (1)-(2), we include all consumers i and all goods x in the consumer-specific control, exposure-only, and recommendation sets. In columns (3)-(4), for each consumer we include the goods to which they were exposed through the belief elicitation survey, and all the goods in the same consumer-specific stratum. Columns (5)-(6) restrict to goods exposed to the consumer through belief elicitation and sampled as per step 3 in Procedure 1.

	Uncertainty (t) $(\sigma^b_{i,x,t})$	
	(1)	(2)
Recommendation	-0.003	0.017^{*}
$(r_{i,x,t-1})$	(0.004)	(0.009)
Rec. × Uncert.$(t-1)$		-0.033^{**}
$(r_{i,x,t-1} \cdot \sigma^b_{i,x,t-1})$		(0.015)
Uncertainty $(t-1)$	0.475^{***}	0.491^{***}
$(\sigma^b_{i,x,t-1})$	(0.016)	(0.017)
Consumer FEs	Yes	Yes
Good FEs	No	No
Observations	8,854	8,854
R^2	0.554	0.554

Clustered standard errors at the consumer level in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 14: The Impact of Recommendation on Beliefs: Uncertainty (Hypothesis 3(1)); with Fixed Effects

Notes: This table tests whether recommendations cause lower uncertainty, as well as for heterogeneous treatment effects. Each observation corresponds to a (consumer i, good x, elicitation period t). The dependent variable is the level of uncertainty report by the consumer at period t. The sample is restricted to goods not previously used for belief elicitation and, to control for good characteristics across treatments, sampled into $S_{i,E,t}$ or $S_{i,R,t}$ (step 3 of Procedure 1).

	|Expect.– Predict. Qual.| (t) ($\|q^b_{i,x,t} - q^p_{i,x,t-1}\|$)		
	(1)	(2)	(3)
Recommendation	-0.015	0.052^{**}	0.056^{**}
($r_{i,x,t-1}$)	(0.011)	(0.026)	(0.026)
Rec. × Uncertainty ($t-1$)		-0.113^{***}	-0.112^{***}
($r_{i,x,t-1} \cdot \sigma^b_{i,x,t-1}$)		(0.040)	(0.041)
Rec. × |Expect. – Predict. Qual.| ($t-1$)			-0.003
($r_{i,x,t-1} \cdot \|q^b_{i,x,t-1} - q^p_{i,x,t-1}\|$)			(0.014)
|Expect.– Predict. Qual.| ($t-1$)	0.651^{***}	0.643^{***}	0.644^{***}
($\|q^b_{i,x,t-1} - q^p_{i,x,t-1}\|$)	(0.013)	(0.014)	(0.015)
Uncertainty ($t-1$)		0.218^{**}	0.217^{**}
($\sigma^b_{i,x,t-1}$)		(0.039)	(0.040)
Consumer FEs	Yes	Yes	Yes
Good FEs	No	No	No
Observations	8,847	8,847	8,847
R^2	0.723	0.724	0.724

Clustered standard errors at the consumer-good level in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 15: The Impact of Recommendation on Beliefs: Expected Quality (Hypothesis 3(2)); with Fixed Effects

Notes: This table tests whether recommendations cause the distance between expected quality and platform-predicted quality to decrease, as well as for heterogeneous treatment effects. Each observation corresponds to a (consumer i, good x, elicitation period t). The dependent variable is the absolute difference between the expected quality reported by the consumer at period t and the platform-predicted quality at period $t-1$. The sample is restricted to goods not previously used for belief elicitation and, to control for good characteristics across treatments, sampled into $S_{i,E,t}$ or $S_{i,R,t}$ (step 3 of Procedure 1).

	Information Acquisition				
	Info. Acq. ($\text{Info}_{i,x}$)		Info. Acq. $\geq t-1$ ($\text{Info}_{i,x,\geq t-1}$)		Info. Acq. $[t-1,t]$ ($\text{Info}_{i,x,[t-1,t]}$)
	(1)	(2)	(3)	(4)	(5)
Exposure Set	-0.0001	0.0001			
($\mathbb{1}(x \in X_{i,E} \cup X_{i,R})$)	(0.0002)	(0.001)			
Recommendation Set	0.010***	0.037***			
($\mathbb{1}(x \in X_{i,R})$)	(0.001)	(0.004)			
Recommendation			0.045***	0.045***	0.023***
($r_{i,x}$)			(0.005)	(0.005)	(0.003)
Exp. Quality				0.006**	
($q^b_{i,x}$)				(0.003)	
Uncertainty				-0.007	
($\sigma^b_{i,x}$)				(0.010)	
Consumer FEs	Yes	Yes	Yes	Yes	Yes
Observations	754,500	93,321	8,847	8,847	8,847
R^2	0.108	0.268	0.159	0.160	0.120

Clustered standard errors at the consumer level in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 16: Recommendations Impact Information Acquisition; with Fixed Effects

Notes: This table tests whether recommendations cause consumers to acquire more information about the recommended goods. Each column displays the average treatment effect of exposure and recommendation on information acquisition (visit a movie's page or watch its trailer); $\text{Info}_{i,x}$ (columns (1)-(2)), $\text{Info}_{i,x,\geq t-1}$ (columns (3)-(4)), and $\text{Info}_{i,x,[t-1,t]}$ (column (5)) are binary variables that correspond, respectively, to consumer i having ever acquired information about good x, having acquired it after $t-1$, and between $t-1$ and t. Each observation corresponds to a pair (consumer i, good x). Columns (1)-(2) correspond to specification (4a); columns (3), (4), and (5) to specifications (4b), (4c), and (4d), respectively. For column (1), we include all consumers i and all goods x in the consumer-specific control, exposure-only, and recommendation sets. In column (2), for each consumer we include the goods to which they were exposed through the belief elicitation survey, and all the goods in the same consumer-specific stratum. Columns (3)-(5) restrict to goods exposed to the consumer through belief elicitation and sampled as per step 3 in Procedure 1.

	Uncertainty (t) ($\sigma^b_{i,x,t}$)				
	(1)	(2)	(3)	(4)	(5)
Info. Acq. $[t-1,t]$	-0.048^{**}	0.042	-0.047^{**}	-0.047^{**}	0.043
($\text{Info}_{i,x,[t-1,t]}$)	(0.020)	(0.042)	(0.020)	(0.020)	(0.042)
Info. Acq. $[t-1,t]$ × Uncert.$(t-1)$		-0.155^{**}			-0.155^{**}
($\text{Info}_{i,x,[t-1,t]} \cdot \sigma^b_{i,x,t-1}$)		(0.071)			(0.072)
Uncertainty $(t-1)$	0.475^{***}	0.477^{***}	0.475^{***}	0.491^{***}	0.477^{***}
($\sigma^b_{i,x,t-1}$)	(0.016)	(0.016)	(0.016)	(0.017)	(0.016)
Recommendation			-0.002	0.018^{*}	-0.002
($r_{i,x,t-1}$)			(0.004)	(0.009)	(0.004)
Rec. × Uncertainty $(t-1)$				-0.033^{**}	
($r_{i,x,t-1} \cdot \sigma^b_{i,x,t-1}$)				(0.015)	
Consumer FEs	Yes	Yes	Yes	Yes	Yes
Observations	8,854	8,854	8,854	8,854	8,854
R^2	0.554	0.554	0.554	0.554	0.554

Table 17: Information Acquisition and Changes in Beliefs: Uncertainty; with Fixed Effects

Clustered standard errors at the consumer level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table examines if information acquisition explains change in consumers' beliefs. Columns (1)-(5) in correspond to variations on specifications (5a)-(5b). Observations correspond to pairs of (consumer i, good x) restrict to goods exposed to the consumer through belief elicitation and sampled as per step 3 in Procedure 1 to control for good quality selection; we use the first belief elicitation of good x's quality from consumer i.

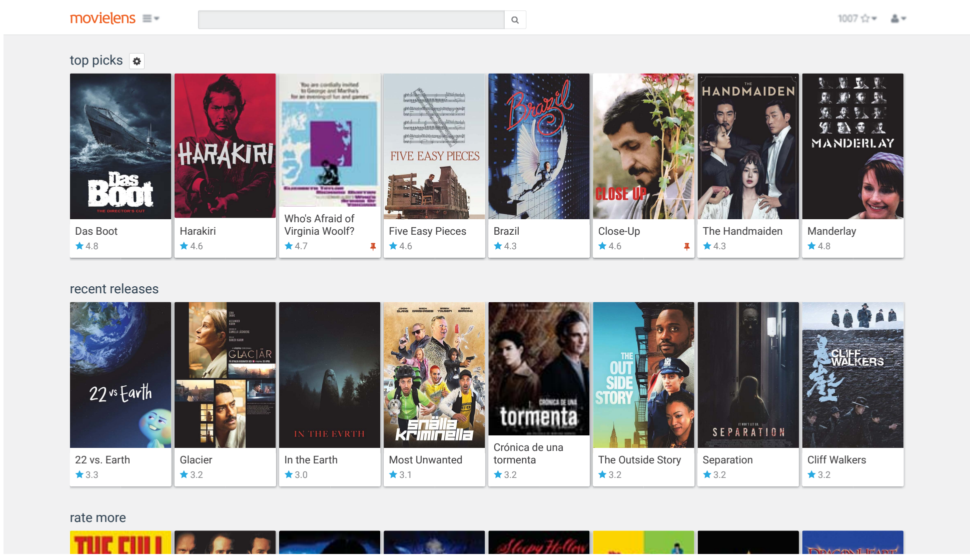
	|Expect.– Predict. Qual.| (t) ($\lvert q^b_{i,x,t} - q^p_{i,x,t-1}\rvert$)				
	(1)	(2)	(3)	(4)	(5)
Info. Acq. $[t-1,t]$	-0.177^{***}	-0.029	-0.172^{***}	-0.170^{***}	-0.026
($\text{Info}_{i,x,[t-1,t]}$)	(0.057)	(0.083)	(0.057)	(0.057)	(0.083)
Info. Acq. $[t-1,t]$ × |Expect. – Predict. Qual.| ($t-1$)		-0.112			-0.111
($\text{Info}_{i,x,[t-1,t]} \cdot \lvert q^b_{i,x,t-1} - q^p_{i,x,t-1}\rvert$)		(0.076)			(0.076)
|Expect. – Predict. Qual.| ($t-1$)	0.651^{***}	0.652^{***}	0.651^{***}	0.643^{***}	0.652^{***}
($\lvert q^b_{i,x,t-1} - q^p_{i,x,t-1}\rvert$)	(0.014)	(0.013)	(0.014)	(0.014)	(0.013)
Recommendation			-0.011	0.056^{**}	-0.011
($r_{i,x,t-1}$)			(0.011)	(0.026)	(0.011)
Rec. × Uncertainty ($t-1$)				-0.113^{***}	
($r_{i,x,t-1} \cdot \sigma^b_{i,x,t-1}$)				(0.040)	
Uncertainty ($t-1$)				0.217^{***}	
($\sigma^b_{i,x,t-1}$)				(0.039)	
Consumer FEs	Yes	Yes	Yes	Yes	Yes
Observations	8,847	8,847	8,847	8,847	8,847
R^2	0.723	0.723	0.723	0.725	0.723

Table 18: Information Acquisition and Changes in Beliefs: Expected Quality; with Fixed Effects

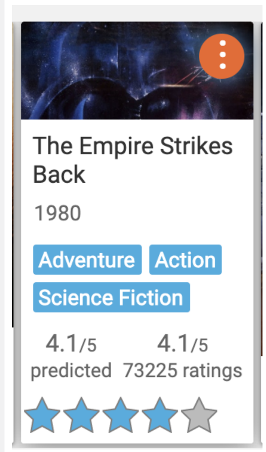
Clustered standard errors at the consumer level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table examines if information acquisition explains change in consumers' beliefs. Columns (1)-(5) correspond to variations on specifications (5c)-(5d). Observations correspond to pairs of (consumer i, good x) restrict to goods exposed to the consumer through belief elicitation and sampled as per step 3 in Procedure 1 to control for good quality selection; we use the first belief elicitation of good x's quality from consumer i.

D. Experimental Instructions and Interface Figures



(a)



(b)

Figure 3: MovieLens Interface

Notes: Panel (a) exhibits the MovieLens home page, where the “top picks” or recommended movies are always at the top. Panel (b) shows the interface experienced when a user hovers over a movie.

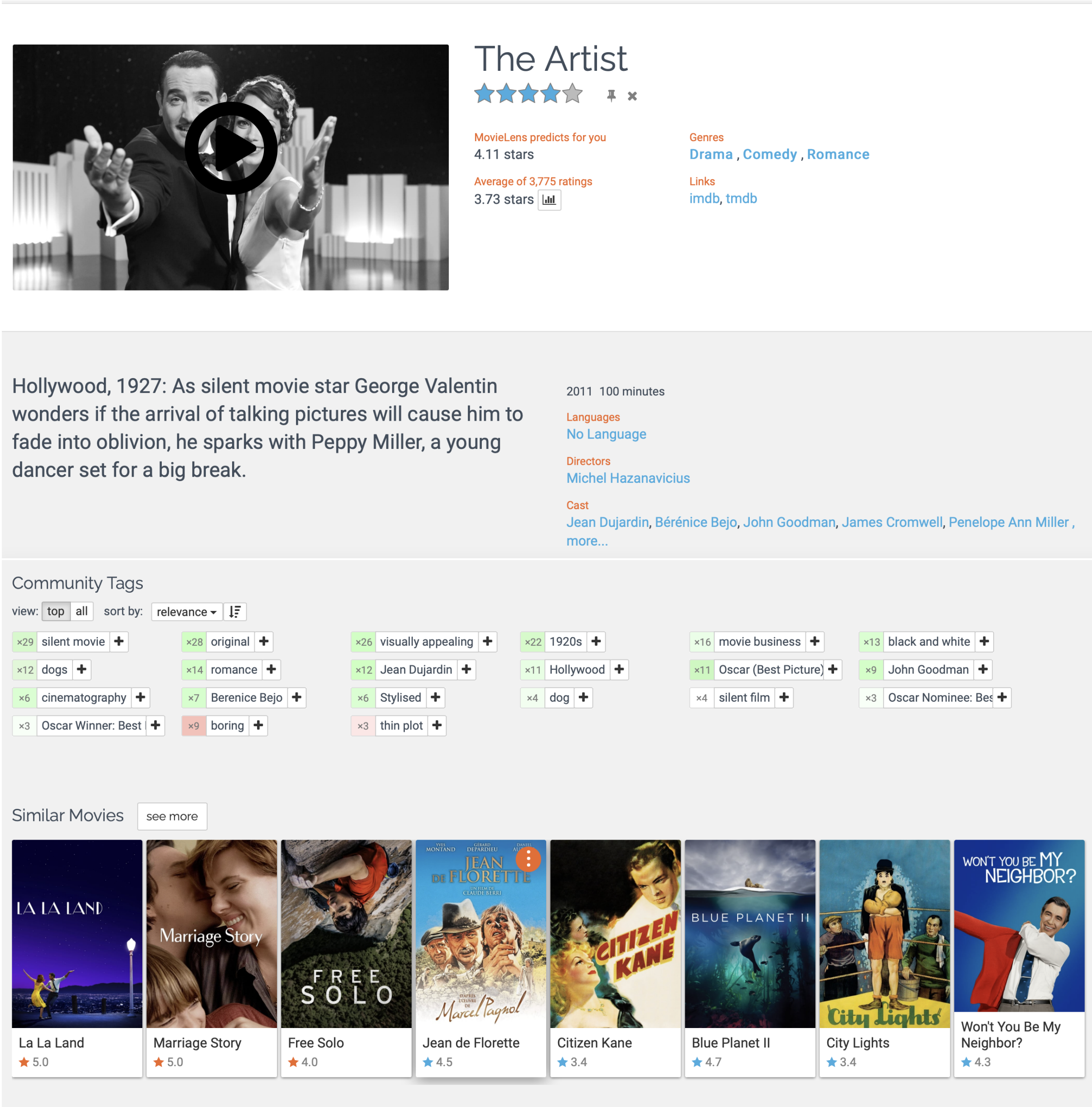
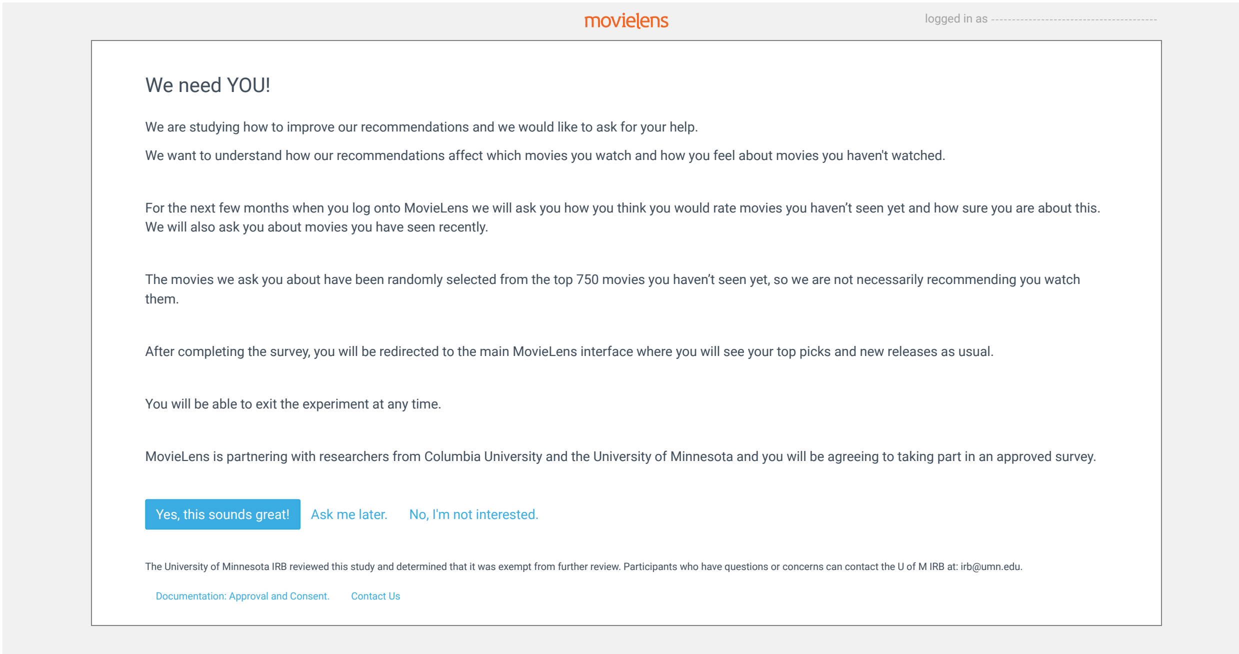


Figure 4: Movie Details Page


movielens

logged in as --

We need YOU!

We are studying how to improve our recommendations and we would like to ask for your help.

We want to understand how our recommendations affect which movies you watch and how you feel about movies you haven't watched.

For the next few months when you log onto MovieLens we will ask you how you think you would rate movies you haven't seen yet and how sure you are about this. We will also ask you about movies you have seen recently.

The movies we ask you about have been randomly selected from the top 750 movies you haven't seen yet, so we are not necessarily recommending you watch them.

After completing the survey, you will be redirected to the main MovieLens interface where you will see your top picks and new releases as usual.

You will be able to exit the experiment at any time.

MovieLens is partnering with researchers from Columbia University and the University of Minnesota and you will be agreeing to taking part in an approved survey.

Yes, this sounds great! Ask me later. No, I'm not interested.

The University of Minnesota IRB reviewed this study and determined that it was exempt from further review. Participants who have questions or concerns can contact the U of M IRB at: irb@umn.edu.

Documentation: Approval and Consent. Contact Us

Figure 5: Consent Form

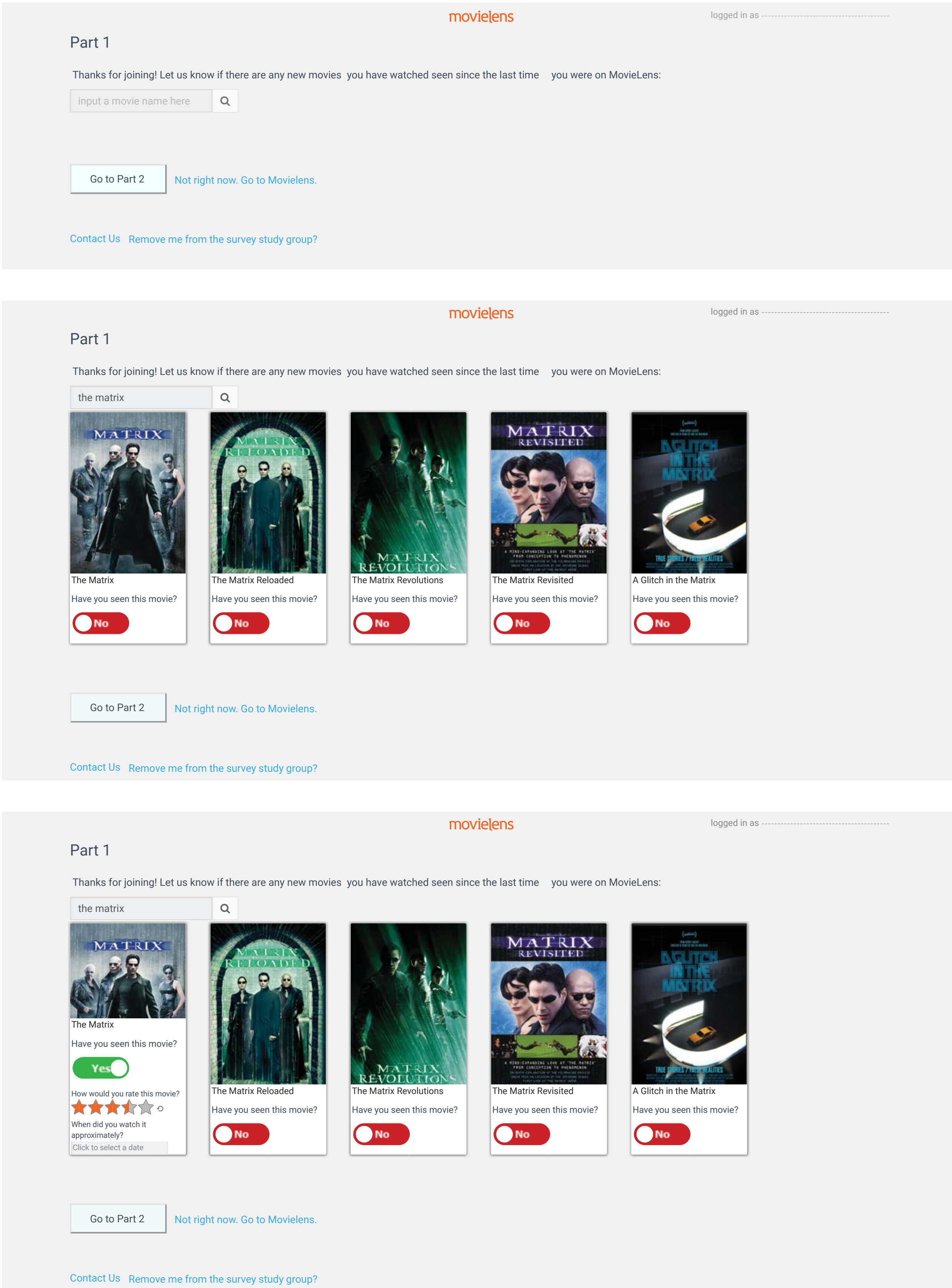


Figure 6: Tracking Movies Watched

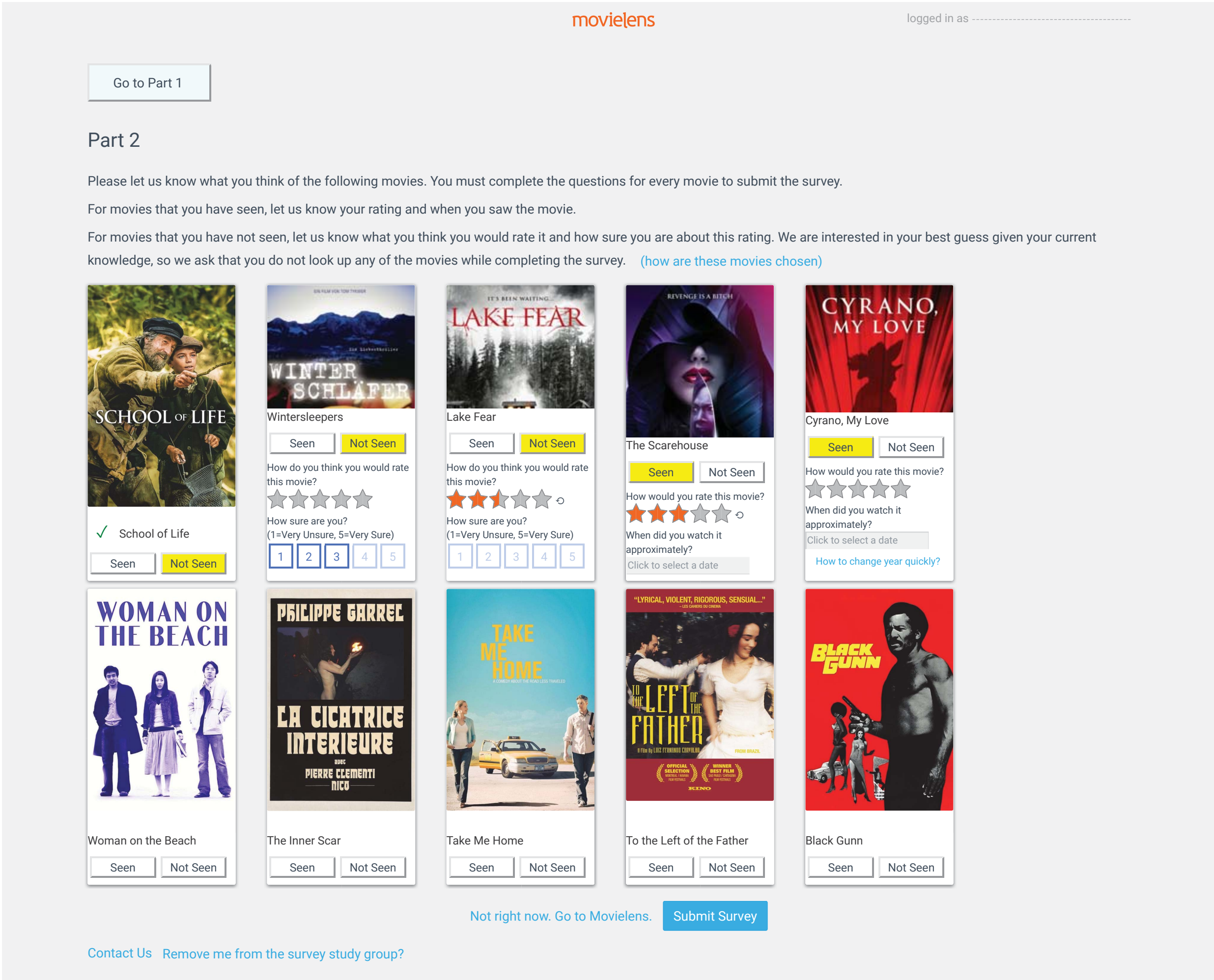


Figure 7: Belief Elicitation

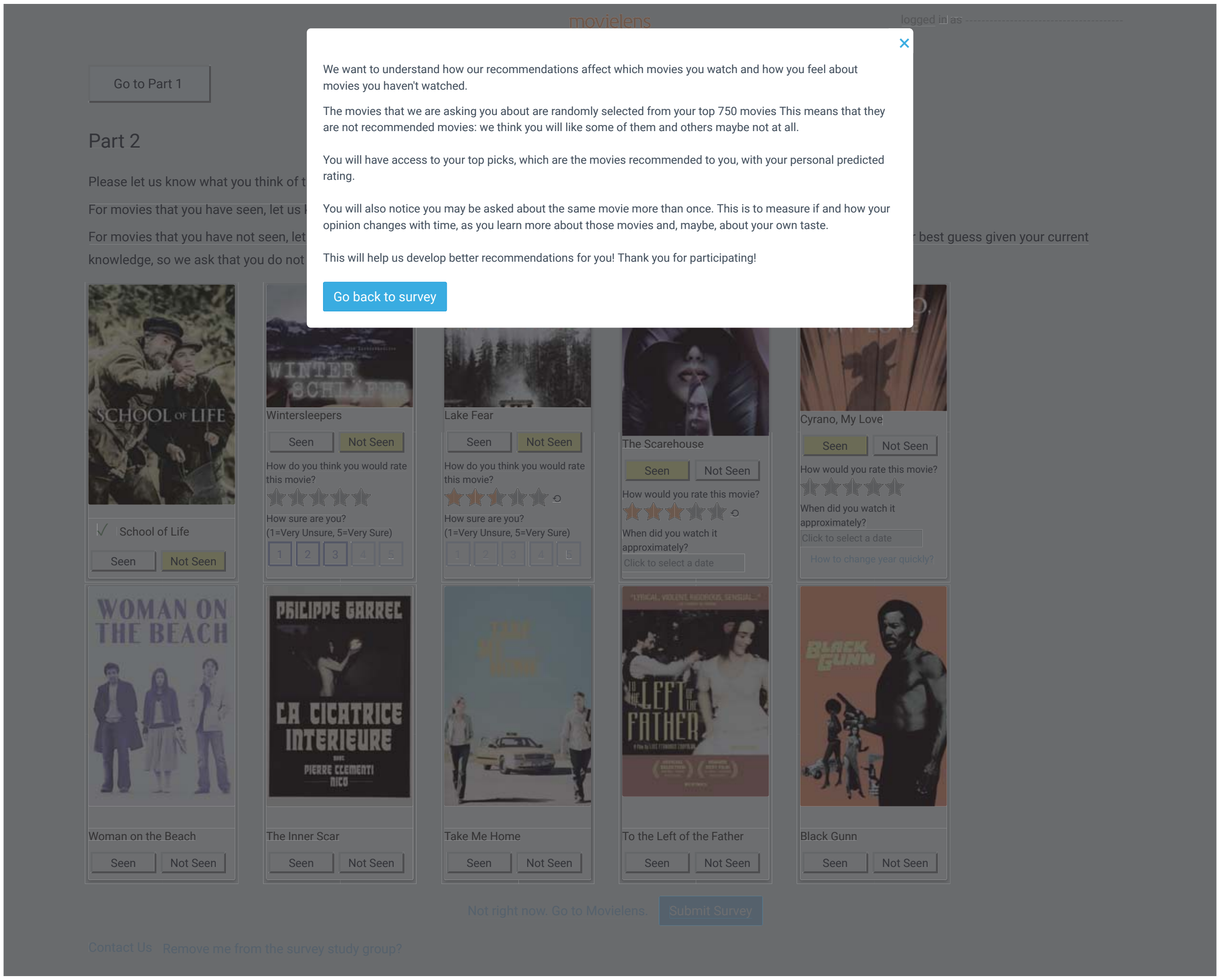


Figure 8: Information on the Sets of Elicited and Recommended Movies

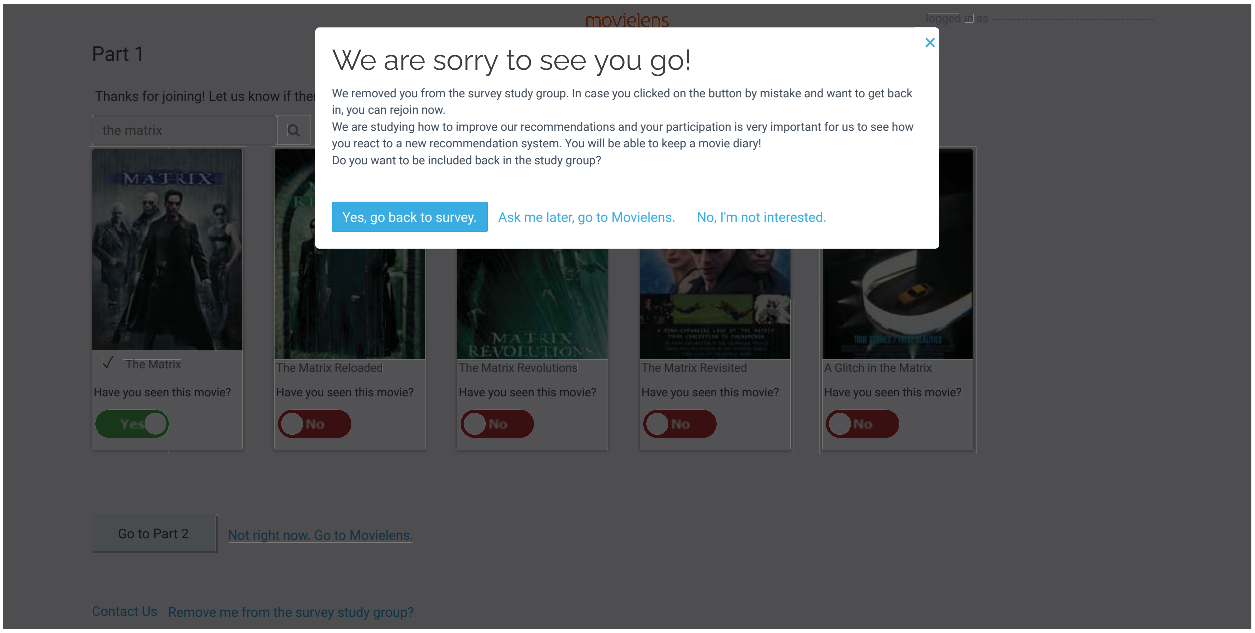


Figure 9: Opting Out