Users and suppliers of complementary products on digital platforms routinely face several non-trivial challenges due to an overload of information and fierce competition for attention. As such, effective management and governance mechanisms facilitating network effects and value co-creation with suppliers are an important determinant of a platform’s competitive advantage. In the present study, we analyze a peer-based recommendation system on a digital music platform, arguing that it can facilitate the match between products’ horizontal characteristics and users’ idiosyncratic tastes or encourage users’ exploration of new varieties, leading to an increase in products’ performance. We derive predictions from a simple theoretical model of peer recommendations which we subsequently test at two levels of analysis. Our results provide strong evidence that peer recommendations indeed lead to increased performance of complementary products, and that they encourage users to extend the scope of the varieties they consume.
Peer Recommendations, Consumption Variety, and Product Performance: Evidence from a Digital Music Platform

January 2019

Abstract

Users and suppliers of complementary products on digital platforms routinely face several non-trivial challenges due to an overload of information and fierce competition for attention. As such, effective management and governance mechanisms facilitating network effects and value co-creation with suppliers are an important determinant of a platform’s competitive advantage. In the present study, we analyze a peer-based recommendation system on a digital music platform, arguing that it can facilitate the match between products’ horizontal characteristics and users’ idiosyncratic tastes or encourage users’ exploration of new varieties, leading to an increase in products’ performance. We derive predictions from a simple theoretical model of peer recommendations which we subsequently test at two levels of analysis. Our results provide strong evidence that peer recommendations indeed lead to increased performance of complementary products, and that they encourage users to extend the scope of the varieties they consume.

1 Introduction

Digitization has led to a substantial increase in the supply and consumption of cultural goods, such as music, books, or video games (Aguiar & Waldfogel 2016, 2018), as well as an increased prevalence of niche products (Brynjolfsson et al. 2011). In addition, the rise of digital platforms made it easier for consumers to find products that match their idiosyncratic tastes by effectively reducing search costs (Bakos 1997, Goldmanis et al. 2010). Still, actors on digital platforms face several non-trivial challenges. First, consumers are commonly confronted with an overload of information (Anderson & de Palma 2012, Dinerstein et al. 2018), complicating the process of product discovery.

*Due to space constraints we are not able to provide the appendix with this submission. It is available from the authors upon request.
Second, although even small suppliers of complementary products potentially have an audience, competition for attention is fierce on a digital platform (Anderson & de Palma 2012, Bordalo et al. 2016). In fact, popularity commonly follows a "long-tail" (Anderson 2004) distribution with the majority of sales being captured by few suppliers. Third, platform markets are often subject to winner-take-all dynamics (Katz & Shapiro 1994, Shapiro & Varian 1998) leading to fierce competition in the attraction and retention of both users and suppliers of complementary products (Eisenmann et al. 2006). As such, the management of a platform’s "ecosystem" has received a lot of attention in recent years (see McIntyre & Srinivasan (2017) for a comprehensive overview), highlighting its importance for a competitive advantage while asking the question what governance mechanisms exists to facilitate network effects and value co-creation with complementors, and how actors in the ecosystem are affected by them (Jacobides et al. 2018).

In the present study, we seek to advance the understanding of those issues by analyzing how peer-based recommendations might help users discover products on a digital music platform. Specifically, we ask two questions: (1) What is the impact of such recommendations on the performance of complementary products? (2) How do consumption patterns of users change in response to peer recommendations? We thereby understand consumption patterns in terms of the variety of the horizontal characteristics of complementary products they buy.

These questions carry important implications for the management and governance of platform ecosystems. First, the decision of suppliers to join a platform crucially depends on their ability to make positive profits there (Eisenmann et al. 2006). Second, attracting users to the platform not only depends on the sheer number of available products, but also their variety. This is because, on the one hand, a higher variety of products can satisfy a broader set of idiosyncratic tastes (Brynjolfsson et al. 2010, Crain & Tollison 2002), and on the other hand, variety as such offers value to consumers (Adler 1985, Chung & Cox 1994, Kim et al. 2002, Ratner et al. 1999). However,
the number and variety of complementary products only translate into value for users if they are able to find them in the first place. As such, we argue that peer recommendations help them in the process of product discovery through two mechanisms. First, it makes users aware of the existence of a product. Second, it helps them evaluate whether the recommended product is a match to their tastes, or if it offers an opportunity to expand their scope of consumption towards a previously unexplored variety. While we expect both mechanisms to translate into a higher probability of a user buying a product that has been recommended to them and – by extension – increased performance of products, the latter aspect is not as clear. If peer recommendations reveal products close to a user’s idiosyncratic tastes, this implies that users actually reduce the scope of varieties they consume as such product will be presented to them more prominently. In contrast, if they reveal products of high inherent quality users might be encouraged to deviate from their tastes as it reduces the uncertainty and costs associated with exploration (Bronnenberg 2015, Datta et al. 2018). As such, studying the effect of peer recommendation on consumption patterns advances the understanding of how they are perceived by users.

In contrast to the present study, prior literature on recommendation systems has largely focused on systems utilizing algorithms built on products’ popularity (e.g. ”Top-Ten” lists) or user’s prior behavior on the platform, and how they influence the macro-level distribution of sales (Kretschmer & Peukert 2017). Those studies do not paint a clear picture, with some finding an increase in concentration (Cai et al. 2009, Salganik et al. 2006, Sorensen 2007), while others find the opposite (Oestreicher-Singer & Sundararajan 2012a,b), suggesting heterogeneity in the effect. One possible driver of increased concentration is a potential bias towards hit products (Celma & Cano 2008, Fleder & Hosanagar 2009). More closely to this study, the theoretical analysis by Hervas-Drane (2015) suggests that peer recommendations can increase demand for niche products. In addition, studies on peer influence focus on changes in behavior at the individual level. In the context
of a digital platform, Wang et al. (2018) show that users are influenced by peers when given online ratings. Earlier studies already established the existence of peer influence (Cialdini & Goldstein 2004, Van den Bulte & Lilien 2001) in general. One argument is that consumers engage in observational learning, thereby updating their evaluation of a product (Liu et al. 2015, Moretti 2011, Zhang 2010), which in turn leads to reduced uncertainty when making a purchase decision (Zhang & Liu 2012).

We study the relationship between peer recommendations, consumption patterns, and product performance in the context of Bandcamp, a digital music platform where users buy and download albums – the complementary products – from independent artists. In addition, they have the possibility to receive peer recommendations by “following” other users and becoming informed about their purchases. Our data contain information about virtually all albums available on the platform, including their horizontal characteristics. In addition, they contain an exact timestamp for all purchases made by users, as well as them starting to follow others. In combination, this information lets us track albums’ performance and users’ consumption patterns over time, and identify which albums have been recommended to them prior to a transaction. To analyze our research questions, we first derive testable predictions from a simple theoretical model, which we subsequently test in a two-part empirical analysis, both at the album and the user level.

We find strong evidence that peer recommendations lead to increased album performance, both in terms of sales and revenues per week, even after controlling for unobserved time-varying album-specific effects. Specifically, a one percent increase in the number times an album is recommended leads to an increase in sales of up to 0.56 %, and an increase in revenues between 3.08 % and 0.84 %. In addition, our results suggest that the effect deteriorates over time and that it is stronger for albums sold at a higher price. Looking at users’ consumption patterns over time, we find strong evidence that they expand the scope of varieties they consume, utilizing an instrumental
variables approach as well as a broad set of fixed effect to account for endogeneity and unobserved effects. This effect, too, deteriorates over times. In addition, we find evidence that users become less reliant on peer recommendations in their exploration efforts as they make purchases at the platform.

Our study contains several contributions: First, we contribute to the literature around the management and governance of platform ecosystems by highlighting peer recommendations as a means to increase complementary products’ performance and to make product discovery easier for users. Second, we contribute to the literature on recommendation systems by analyzing a mechanism at the level at the individual user, and providing empirical evidence suggesting a decrease in sales concentration. This aspect also contains a contribution to the literature on peer influence on digital platforms. Lastly, we highlight important aspects connected to the consumption of cultural goods in the digital era.

The remainder of the study is structured as follows: In section 2 we lay out our simple model of peer recommendations and derive testable predictions. In section 3, we provide details about the empirical setting and our data set. We describe the construction of our key measurements, and lay our our empirical approach as well as our identification strategy in section 4. We present the results of our empirical analysis in section 5. Section 6 offers concluding remarks.

2 A Simple Model of Peer Recommendations

We develop a simple partial equilibrium model to predict how peer recommendations influence the behavior of users buying music albums on a digital platform. We do not explicitly model under what circumstances an album is recommended. We use this model to derive some predictions that we subsequently test empirically. Specifically, we aim to derive predictions about how peer recommendations affect (i) a user’s likelihood to buy a specific album and (ii) under what
circumstances peer recommendations might lead to a user buying albums that do not correspond to her tastes, or that encourages her to explore horizontal characteristics new to her.

2.1 Setup

Consider a digital platform where users buy music albums. Albums, indexed by $j$, are horizontally differentiated and consumers, indexed by $i$, are heterogeneous in their horizontal preferences for music characteristics. As such, both albums and users are located along a Hotelling line with locations $L_j$ and $L_i$, respectively. In addition, each album has an inherent quality, $\theta_j$. Whereas the aspect of horizontal differentiation captures matters of tastes for certain genres (e.g. Rock, Jazz, Hip-Hop) or musical characteristics (e.g. instrumental, experimental, progressive), the quality aspect captures album characteristics like an album’s recording quality or the singer’s ability to stay in tune.

We adopt a separable utility framework for users similar to Chiang & Knight (2011) and Petrova et al. (2018). As such, the utility user $i$ receives from buying album $j$ is described by

$$u_{ij} = \theta_j - \tau |L_i - L_j|,$$

where $\tau$ is a “transportation” cost parameter, capturing disutility from buying an album that deviates from user $i$’s preferences. Intuitively, this setup captures two distinct aspects of utility, namely an album’s inherent quality, $\theta_j$, as well as the match between album $j$’s horizontal characteristics and user $i$’s tastes via the term $\tau |L_i - L_j|$.

Next, given the experience good nature of music, users are not able to perfectly observe an

\footnote{Note that we ignore the price of an album in the present setup. We do this for several reasons: First, we are not interested in artist’s price setting behavior. Second, we do not observe price changes in our data. In our empirical analysis we control for (unobserved) aspects - most prominently album quality - by including (time-varying) fixed effects, which makes estimating a price coefficient impossible. In light of this, adding a price parameter in our model would add no further insights.}
Figure 1 Expected quality and utility

Notes: The graphic displays the interplay between album \( j \)'s expected quality, \( E(\theta_j) \), the match with user \( i \)'s tastes, and the expected utility, \( E(u_{ij}) \), she would receive from buying the album. Here, it is positive, such that the user would buy it.

album’s quality prior to purchasing it. As such, an album’s quality can be described by

\[
\theta_j = q_j + \epsilon_j,
\]

with \( q_j \) denoting quality aspects that can be observed by users ex ante, and \( \epsilon_j \) capturing quality aspects that are only revealed after buying the album. From a user’s perspective, we assume \( \epsilon_j \) to be a random variable with \( E(\epsilon_j) = 0 \). Users then decide to buy an album if \( E(u_{ij}) \geq 0 \), or

\[
E(\theta_j) \geq \tau |L_i - L_j|.
\]

A first intuitive insight of the model is that users are more willing to deviate from their tastes with purchases of albums that have a higher expected quality or, conversely, that they accept a lower expected quality for albums that are closer to their tastes.

2.2 Peer recommendations

The key question we explore here is about the nature of the informational content primarily connected to peer recommendations. As we explained earlier, we argue that such recommendations
can either provide information about the match between the recommended album’s horizontal characteristics and the focal user’s tastes, or if they provide information about the inherent quality of an album.\(^2\)

First, consider the case where peer recommendations primarily provide information about the horizontal match. Going forward, we will refer to this case as a match recommendation. In the context of the model, this implies that such albums are located within a specific distance around the focal user’s location on the Hotelling line. The implications with regard to the likelihood of a user buying such a recommended album and whether or not she will deviate from her established tastes are straightforward. As the term \(|L_i - L_j|\) is smaller for recommended albums by design, the necessary expected quality of the album to fulfill \(E(u_{ij}) \geq 0\) is smaller, as well, making a purchase more likely. This leads to the following propositions:

**Proposition 1.** In the case of a match recommendation, users are more likely to buy an album when it has been recommended to them.

**Proposition 2.** In the case of a match recommendation, users will not deviate from their tastes with the purchase of an album that has been recommended to them.

Next, consider the case where peer recommendations primarily provide information about an album’s inherent quality. Going forward, we will refer to this case as a quality recommendation. Here, predictions are not as straightforward. As we established earlier, users are not able to perfectly observe an album’s quality ex ante and have to base their decision whether or not to buy an album on their prior beliefs about the album’s quality. Given \(E(\epsilon_j) = 0\), absent any peer recommendation, an album’s expected quality is given by \(E(\theta_j) = q_j\).

However, if an album is recommended by a peer, and we consider the case of a quality

\(^2\)One could also argue that peer recommendations provide information about both at the same time. Our analysis then can be perceived as an effort to analyze which of the two is dominant in our setting.
Figure 2 Quality Recommendation

Notes: The graphic displays the effect of a recommendation on the expected quality of an album. The gray area indicates combinations of the expected quality and location of albums that generate positive utility for user \(i\), i.e. \(E(u_{ij}) \geq 0\). In this illustrative example, user \(i\)’s prior belief about album \(j\)’s quality does not lead to positive utility, i.e \(E(u_{ij}) < 0\). After receiving a peer recommendation, that belief is updated such that \(E(u_{ij}|r^{\theta}) \geq 0\). Thus, user \(i\) would buy album \(j\) only if it has been recommended to her.

recommendation, the user’s belief is updated such that she now expects a higher inherent quality. The posterior belief, after receiving a quality recommendation denoted by \(r^{\theta}\), is then given by

\[
E(\theta_j|r^{\theta}) = q_j + \gamma,
\]

with \(\gamma > 0\). As such, user \(i\) will buy album \(j\) that has been recommended to her if \(E(u_{ij}|r^{\theta}) \geq 0\), as illustrated in Figure 2. This leads to the following propositions:

**Proposition 3.** In the case of a quality recommendation, users are more likely to buy an album when it has been recommended to them.

**Proof.** A user is more likely to buy an album if it has a higher expected utility. Therefore, user \(i\) is more likely to buy album \(j\) if it has been recommended to her if \(E(u_{ij}|r^{\theta}) > E(u_{ij})\), or:

\[
q_j + \gamma > q_j - \tau |L_i - L_j| \Rightarrow \gamma > 0
\]

\[\square\]
Proposition 4. In the case of a quality recommendation, users are more willing to deviate from their established tastes if an album has been recommended to them.

Proof. A user purchases an album if \( E(u_{ij}) \geq 0 \Leftrightarrow E(\theta_j) \geq \tau|L_i - L_j| \). Absent a quality recommendation, this implies \(|L_i - L_j| \leq \frac{q_j}{\tau}\). For an album that has been recommended, this implies \(|L_i - L_j| \leq \frac{q_j + \gamma}{\tau}\). Therefore, a user is more willing to deviate from established tastes with the purchase of an album that has been recommended to her if:

\[
\frac{q_j + \gamma}{\tau} > \frac{q_j}{\tau} \iff \gamma > 0 \quad \square
\]

3 Data and Setting

3.1 Empirical Setting: Bandcamp.com

Bandcamp is a music distribution platform that has been launched in 2009. On the platform, users buy and download digital albums. Similar digital platforms are Spotify or Amazon Music. In contrast to those more well-known websites, Bandcamp has a distinct focus on independent music, thus features a large number of lesser-known and amateur artists not associated with a major label\(^3\) or, in most cases, any label at all. This is also reflected by the pricing behavior of those artists: From the 836,908 albums in our data set almost half (48.24 \%) are either offered for free or on a pay-what-you-want (PWYW) basis. This indicates that a considerable fraction of them appears to seek exposure rather than earning money through album sales. Another indication that artists on the website largely seem to struggle with setting an appropriate price is that Bandcamp

\(^3\)That is the Warner Music Group, Sony Music Entertainment, and the Universal Music Group
offers recommendations in this regard in their FAQ section\textsuperscript{4}. In addition to the possibility to offer albums on a PWYW basis, users are generally given the possibility to pay more than the asking price and Bandcamp reports that users actually do pay more in 40\% of all cases\textsuperscript{5}. This indicates that consumers of independent music on the site are not always concerned about the price they pay, but rather in the quality and discovery of music. While creating a presence on Bandcamp and uploading music is free, the platform claims a share of all revenues generated through digital sales\textsuperscript{6}, which is its primary source of revenue. This share amounts to 15\% of the selling price, but drops to 10\% as soon as 5,000 USD in revenue are reached. In addition, it allows artists to sell merchandise, claiming 10\% of the proceedings, and offers premium packages for artists and labels for a monthly fee. When offering an album on Bandcamp artists set up a customizable album page containing information about its price, the containing tracks, as well as additional information about its stylistic characteristics. The latter is primarily provided in the form of tags referring to genres, the location of the artist, references to similar artists, as well as additional characteristics. Aside from that, users usually have the possibility to stream some or all of the songs on the album at reduced quality before buying it. In addition, each album page shows all registered users who have already bought the album, as well as small review texts users can provide.

3.2 Peer Recommendations: Bandcamp for Fans

In early 2013, the website was considerably redesigned. The most prominent new aspect was the introduction of several new features under the label Bandcamp for Fans that sought to make it easier for users to traverse the platform and discover music\textsuperscript{7}.

Now, the website contains profile pages that prominently display the "collection" of albums a

\textsuperscript{5}Again, see https://get.bandcamp.help/hc/en-us/articles/360007802534-What-pricing-performs-best- [accessed: 09.01.2019]
\textsuperscript{6}See https://bandcamp.com/pricing [accessed: 09.01.2019]
\textsuperscript{7}See https://daily.bandcamp.com/2013/01/10/bandcamp-for-fans/ [accessed: 09.01.2019]
user has bought. In addition, users can provide information about their place of residency and their preferred music styles, and they are able to customize their profile page. Simultaneously, information about users who have bought a product are displayed on album pages. However, most relevant to the present study, is the possibility for users to connect with other users. As such, they are able to ”follow” other users and observe their purchases. Those links are akin to, for example, following users on Twitter. They are non-reciprocal, i.e. the follower becomes informed about the actions of the followed user and not vice versa. This happens in two ways: First, Bandcamp contains a feed that displays purchases made by users they follow. Second, as soon as a followee buys an album, the user receives an Email, informing her about that purchase. For the purpose of our study, we therefore consider a user to receive a peer recommendation as soon as one of their followees buys an album.

3.3 Dataset

Using web-scraping techniques, we acquired a snapshot of Bandcamp in March 2018. We obtained data on 836,908 virtually all albums available on the platform at that point in time. These data contain information about the asking price, the number of songs, the total length (in seconds), the region of origin, and the album’s release date. In addition, we collected the tags are associated with the album, providing information about, among other aspects, its music characteristics.

We were also able to obtain information about 637,979 registered users buying albums, as well as an exact timestamp when a purchase took place. Based on this information, we are able to

---

8Users are also able to follow artists and even tags. However, here we focus on the link to other users for two reasons: First, a user automatically follows the artist as soon as they buy an album from her, which does not make it a conscious choice. Second, becoming informed about new releases of an artist that a user has previously purchased from is more akin to a ”collaborative filtering” recommendation system, for which the question what kind of information is conveyed (match or quality) is irrelevant. In our empirical analysis we consider this by including a dummy indicating a user buying a second album from the same artist.

9In contrast, an example for a reciprocal link is two users becoming friends on Facebook and become informed about each other.

10Going forward, we refer to ”users they follow” as ”followees”.

11More precisely, we have information about which currency is asked from which we infer this information.
calculate each album’s sales at any given time between its release on Bandcamp and March 2018. In addition, this lets us track the exact order in which users purchased albums that exhibit certain characteristics, as well as how much time has passed between purchases. In total, we have this information for 5,922,346 purchases made by registered users\textsuperscript{12}.

Next, we were able to obtain information about the non-reciprocal links between users, as well as an exact timestamp when a user started following another one. This information serves as the basis to identify peer recommendations taking place. To reiterate: For a user to receive a peer recommendation, that user must first start following another user, and that user has to make a purchase. The combination of the exact timestamp for links being established and purchases being made provides us with exactly that information. In our data, a total of 138,963 users follow at least one other user, and we recorded a total of 1,053,928 non-reciprocal links.

4 Measurement and Empirical Approach

To test the predictions of our theoretical model we conduct a two-part analysis, at the product and at the user level.

First, a key insight from our theoretical model is that users are more likely to buy an album that has been recommended to them compared to albums that are not, \textit{regardless} of whether recommendations primarily contain information about the match with a user’s tastes (Proposition 1) or about an album’s inherent quality (Proposition 3). The reason is that the (expected) utility

\textsuperscript{12}Unfortunately, we were not able to obtain information about purchases made by unregistered users. As such, our calculations of total sales and revenues have to be understood as a lower bound. More importantly, if there is a systematic difference at the album level with respect to who buys an album, and a systematic difference between the behavior of registered and unregistered users, our results might contain a bias. One possibility of the nature of this bias is the following: If more popular artists are more likely to be bought by unregistered users, then sales figures for those artists are understated in our data. At the same time, unregistered users are not able to receive peer recommendations, which limits the issue to only a part of our analysis. Still, we are aware of this problem, and consider it a limitation of our study.
of buying the album is higher than for albums that are not recommended in either case. At the product level, this implies that an album that gets recommended more, will have an increased subsequent sales performance. Hence, we test this prediction at the product level and expect an album’s weekly sales to be higher, the more it has been recommended to users in the previous weeks, *ceteris paribus*.

Second, an additional insight of our theoretical model is that the degree to which a user is willing to deviate from her tastes crucially depend on the kind of information conveyed by a recommendation. In the case of a match recommendation, we expect users to not deviate from their tastes. In contrast, with quality recommendations, we expect users to be more willing to deviate from their tastes compared to albums that are not recommended. We test this at both the album and the user level. First, at the album level, we investigate to what degree the users who bought an album in a given week deviate from their tastes on average. Second, we also test this prediction at the user level. Here, we investigate directly, to what extent a user deviates from her tastes with the purchase of a specific album, conditional on whether or not that album had been recommended to her.

### 4.1 Music Characteristics and User Tastes

In our operationalization of a users’ tastes – or revealed preferences – at any given point in time, we use information about the tags assigned to albums she buys. The underlying idea is that these tags represent the musical characteristics she consumes over time. To analyze the extent to which a user deviates from her tastes over time, we construct a measure indicating how concentrated she is in those characteristics, and how this concentration changes with each album purchases. As such, if we observe a user to become less concentrated in the characteristics she consumes, this...
is indicative of her deviating from established tastes. In contrast, if we observe her to become more concentrated, this is indicative of her purchasing an album that exhibits the same (or very similar) music characteristics. In the following, we lay out in detail, how we calculate this measure of concentration.

4.1.1 Construction and Representation of User Tastes

We begin by describing the three-step process leading to our operationalization of user tastes:

(1) We reduce the set of tags in our analysis to the 200 most-assigned ones over the whole period between 2009 and March 2018. Some of those tags are references to the artist’s place of residence. As we are interested in music characteristics we removed all of those. After some additional manual cleaning of the data, we arrive at a set of 134 tags\(^{14}\) that indicate such characteristics, i.e. genres and other stylistic elements.

(2) For each album we constructed a vector indicating whether or not a specific tag is assigned to it. As such, this is a 134-dimensional vector of ones (i.e. a tag is assigned to the album) and zero otherwise.

\(^{14}\)A full list of tags is provided in the appendix.
and zeroes (i.e. a tag is not assigned to the album). This serves as our representation of music characteristics an album exhibits.

(3) We track each user’s purchases over time, as well as the music characteristics that are connected to those purchases. Specifically, for each album she buys, we add up the tag vectors associated with them. This in turn produces another user-specific 134-dimensional vector containing the number of times she has consumed each tag. Next, we represent this vector in shares. This last expression serves as the basis for the calculation of a user’s concentration in music characteristics. Figure 3 summarizes steps (2) and (3) using an illustrative, simplified example.

4.1.2 Taste Concentration

Using a representation of a user’s tastes that consists of shares of music characteristics serves as the basis for the calculation of her taste concentration at any given point in time. As a first step we construct a measure of distance between two tags to account for the possibility of pairs of tags representing the same or similar music characteristics. To that end, we use information about how often pairs are co-assigned to albums by artists. The conceptual distance between two tags, $k$ and $l$, is then given by:

$$d_{kl} = 1 - \frac{|k \cap l|}{k}.$$  

This measure is bound between 0 and 1, with higher values indicating a higher conceptual distance between two tags. By construction, it is asymmetric, i.e. $k$ might be closer to $l$ than $k$ is to $l$. A similar approach can be found in Leung (2014) and Kovacs & Hannan (2015).

In a next step we further adapt a measure found in Kovacs & Hannan (2015), namely a tag’s typicality. Intuitively, this construct measures how dominant a specific tag is in the user-specific tag vector at a specific point in time. Its typicality will be higher, the more often a user bought albums that have it assigned to them. Moreover, this measure incorporates the conceptual distance
of the focal tag $k$ to all other tags present in the tag vector at the same point in time. As such, tag $k$’s typicality in user $i$’s tag vector at time $t$ is given by:

$$\psi_{kit} = \frac{p_{kit}}{p_{kit} + \sum_{l \neq k} d_{kl} \cdot p_{lit}},$$

with $p_{kit}$ being the share of tag $k$ in user $i$’s tag vector at time $t$.

Finally, we arrive at a measure for the concentration in the user-specific tag vector by constructing a weighted average of all tags’ typicalities that appear in it at any given point in time. The intuition is that the higher each occurring tag’s typicality, the more concentrated are a user’s tastes. As such, the weighted average typicality of user $i$’s tag vector at time $t$ is given by:

$$\text{WAT}_{it} = \sum_k p_{kit} \cdot \psi_{kit} \sum_k \mathbb{1}(p_{kit} > 0).$$

This measure of concentration is also bound between 0 and 1, with higher values indicating a higher level of concentration.

Tracking the development of the WAT for each user over time lets us now measure whether or not, and to what a extent, a user deviates from established tastes with the purchase of an album. Specifically, a decrease in concentration indicates an increase in consumption variety. As such, when analyzing this question, we will evaluate the impact of peer recommendations on the change in WAT, or:

$$\Delta \text{WAT}_{it} = \text{WAT}_{it} - \text{WAT}_{it-1}.$$  

This measure is bound between -1 and 1\(^{15}\), with $\Delta \text{WAT}_{it} \geq 0$ indicating a purchase that corresponds to established tastes, and $\Delta \text{WAT}_{it} < 0$ indicating a deviation from these tastes.

\(^{15}\)In our empirical analysis, we scale the measure by a factor of 100 to improve the readability of our results. I.e. the measure is then bound between -100 and 100.
4.2 Empirical Framework

We test the predictions of our model in a two-part analysis, namely at the album level and at the user level. First, our predictions carry the implications that peer recommendations should increase an album’s sales performance. We analyze how album performance is affected by the number of times it has been recommended to users. Second, we test our predictions with regard to whether users deviate from their tastes more when purchasing an album that has been recommended to them. We conduct this analysis at both the album and the user level.

4.2.1 Album Level

We use a weekly panel of albums throughout the analysis and run three sets of regressions. First, we analyze the effect of peer recommendations on album performance. Here, we run two sets of regressions using different measures of performance as the dependent variable, weekly sales and weekly revenues. For the latter, a few remarks are in order: First, our model is silent about the role of prices. Still, we consider it an important and interesting aspect to what extent heterogeneous effects of peer recommendations with regard to pricing play a role. Second, As we captured a snapshot of the platform, we only observe the asking price at that point in time. As such, we do not observe price changes. Nevertheless, we are confident in our results as the price setting actors on the platform are independent and amateur artists. Therefore, we do not expect them to be particularly strategic in their behavior.\footnote{To confirm this assumption, we are currently in the process of scraping the prices of the albums in our data again. Once this process has concluded, we will be able to assess how many prices have in fact changed over the course of the previous year.} Given this constraint, we calculated weekly revenues by multiplying weekly sales with the asking price. Third, Bandcamp reports that users pay a higher amount than the asking price in 40% of all purchases. Unfortunately, we only observe the asking price. As such, our calculation of revenues are to be understood as a lower bound.
In the third set of regressions, we conduct our album-level analysis of the question whether or not and to what extent users increase their consumption variety. As we are able to identify which users bought a specific album in a given week, and how their concentration changes, we aggregate this information to the weekly album-level by taking a simple average of each buyers change in WAT. As such, our dependent variable here is the average $\Delta$WAT in a given week.

Next, as common in any media market, the distribution of album sales on the platform follows a long-tail distribution. As such, our data contain a high number of low-performing albums, and a low number of high-performing albums. We tease out additional heterogeneous affects along this line by running three separate regressions for dependent variable. First, we use the full sample of albums that sold at least one copy over the whole sample period. Second, we restrict the sample to albums that sold at least ten copies over the whole sample period. Last, we further restrict the sample to albums that sold at least 50 copies in total.

Our main independent variable of interest here is the number of times an album has been recommended. Given that our data contains the information at what exact point in time a user started following another user, and at what point in time users make purchases, we are able to calculate the number of times an album is recommended this way. To then analyze the impact on each of our three dependent variables, we run a regression on the one, twice, three times, and four times lagged number of weekly recommendations. This way, we are also able to tease out if and to what extent the effect deteriorates over time.

An empirical challenge in our context is that, in order for an recommendation to take place, another user has to have purchased the focal album. As such, the number of recommendations is a function of the number of sales in a given week. These sales, however, will also be affected by unobserved (to the econometrician) factors, like the album’s quality or unobserved popularity shocks. To account for those factors, it is important to include the once, twice, three times, and
four times lagged number of weekly sales in the regression as well. In addition, we include the cumulative prior sales of an album to further account for unobserved popularity. In addition, we include an album’s age and squared age in weeks as a control variable\textsuperscript{17}. To obtain elasticities, we use the natural logarithm for all our variables, with the exception of the average $\Delta$WAT and our age variables.

Lastly, a key construct of our theoretical model is an album’s inherent quality, which we do not observe. As such, to control for this aspect we include a time-varying album-month fixed effect. Not only will this control for unobserved album-specific characteristics, but also for possible price changes that might impact the results. Further, we include a week fixed effects to control for seasonal effects.

\textbf{4.2.2 User Level}

At the user level, we seek to further analyze whether or not and to what degree users deviate from established tastes when buying an album that has been recommended to them. Here, the unit of observation is a single transaction, i.e. a specific user buying a specific album. As such, we conduct a cross-sectional analysis at the user-album level. Here, we use the degree to which the weighted average typicality in a user’s tag vector changes with this specific purchase, $\Delta$WAT, as the dependent variable throughout the analysis.

Again utilizing information about the exact points in time a non-reciprocal link between users is established and when followees make purchases, we are able to assess whether or not a purchased album had been recommended to the focal user prior to the transaction. As such, the main independent variable of interest is a dummy indicating whether or not this is the case. In doing so, we vary the time window in which we allow this dummy to take the value 1. We run four separate

\textsuperscript{17}Age, however, is dropped in our regressions due to collinearity with the fixed effects.
regressions, considering time windows of one day (i.e. 24 hours), one week, two weeks, and one month prior the focal transaction. This way, we are also able to assess how immediate the effect is, and whether or not and to what extent it might deteriorate.

As changes in the taste of a user might be caused several reasons other than a peer recommendation, we include a number of control variables, both at the user and the album level. First, users vary in the number of other users they follow and the higher that number is, the more albums will be recommended to them. More importantly, a higher number of followees might be indicative of a user seeking to receive a diverse set of recommendations. Therefore, we include this number as a control variable throughout the analysis. Relatedly, a more popular album will also be recommended more often and this might be indicative of a higher quality as well. Therefore, to distinguish this effect from the effect of a peer recommendation, we include the focal album’s prior sales throughout the analysis. Further, if a user is not observed making a purchase on the platform, it is possible that her tastes change for reasons outside our observable spectrum. Therefore, we also control for the time since the last purchase a user has made on the platform. In addition, we control for the number of prior purchases made by the user. Next, we mentioned earlier that users automatically start following an artist as soon as she buys an album from that artist and will subsequently receive new releases as a recommendation. As this is likely to be a purchases corresponding to her tastes, this might bias our results towards an increase in concentration in her tag vector. Therefore, we include a dummy indicating whether or not the focal user has bought an album from the artist before. We also include a user’s taste concentration prior to the focal transaction as a control, as it might be indicative of a general affinity towards a diversity in music characteristics, or lack thereof. Aside from that, it is important to control for this for purely mechanical reasons, as a higher WAT score is more sensitive to the addition of new tags to the vector. Lastly, we control for a user’s number of prior purchases and an album’s age here as well.
Again, to further control for unobserved effects, we include three types of fixed effects. Namely, we include a user fixed effect, as well as – like in the album level analysis – an album-month fixed effect, and a week fixed effect.

4.2.3 Endogeneity Issues and Identification

Ideally, we would analyze our research questions in an experimental setup where links between users are randomly generated. However, in our study we have to rely on observational data. As such, the issue is that users choose how many users to follow, as well as whom to follow. As a result, simply using ordinary least squares regression would likely produce biased results caused by selection. Therefore, even though we include a variety of control variables and fixed effects, we are still concerned about endogeneity issues in our user-level analysis. Specifically, two variables raise concerns, namely the dummy indicating whether or not an album was recommended prior to focal transaction, and the number of followees a user has. For the former, the concern is that users might select into exactly the kind of informational content they want to receive. I.e. if they seek to discover albums that correspond to their tastes, they will follow users that provide this information. This would bias our results towards a positive effect of peer recommendations on $\Delta WAT$. Conversely, if they seek to discover music characteristics they haven’t explored yet, they will select followees accordingly, leading to a bias towards a negative effect of peer recommendations. Aside from the direction of the bias not being clear, it is also not clear how systematic this bias is. In addition, to this type of selection, we also consider how users most likely discover followees. Earlier we described that each album page contains the information which users have previously bought that album. If users indeed primarily discover followees this way, this would lead to selection caused by homophily and a systematic bias towards a positive effect of peer recommendations on $\Delta WAT$. Next, following a greater number of users might be indicative of seeking a diverse set of recommendations. Even though this is not our main variable of interest,
we are concerned that a bias here might entail a bias for the effect of peer recommendations as well. If users indeed seek more diverse recommendations by following a greater number of other users, we would expect a bias towards a negative effect on $\Delta WAT$ here.

We address these issues utilizing an instrumental variables approach. First, addressing the selection issue connected to the dummy indicating a peer recommendation we utilize the non-reciprocal nature of the links between users. The intuition is that a focal user’s followees will be affected by the recommendations they themselves receive. However, as the focal user does not observe these information flows, they should not affect her behavior directly. Therefore, we use these information flows to predict the purchases of the focal user’s direct followees. Specifically, we use the total number of recommendations all of the focal user’s followees have received prior to the focal transaction. While this should have an impact on a certain album being recommended, it
should be orthogonal to whether or not the focal user deviates from her established tastes\textsuperscript{18}. The intuition is illustrated in figure 4. Second, addressing the selection issue connected to the number of followees a user exhibits, we use information about the popularity of those followees prior to the links with the focal user being established. Specifically, we use the total number of other followers each of the focal user’s followees have at the time of the link establishment. The intuition here is that it is more likely that the focal user will follow a popular other user. This, however, should be orthogonal to her deviating from her established tastes or not.

5 Results

We organize the presentation of our results\textsuperscript{19} as follows: First, we provide descriptive statistics for the variables we use in the analyses at the album and the user level. Next, we discuss our results regarding album performance. Lastly, we discuss the results on the question whether or not and to what extent users extend their consumption variety.

5.1 Descriptive Statistics

5.1.1 Album Level

Table 1 contains summary statistics for our album level variables. Although we primarily use logarithmic values in our analysis, we also show absolute values to better provide insights about the scope and scale of our variables.

The top three rows of table 1 show summary statistics for the three dependent variable we use

\textsuperscript{18}We are aware that this might be an imperfect instrument. We expect a systematic bias of our results due to a homophily bias. Accordingly, if the focal user’s followees are subject to the same bias, we might not eliminate it fully and our estimates should be understood as a lower bound.

\textsuperscript{19}Even though we have data on peer recommendations from early 2013 through March 2018, at this stage we restricted out analysis to the year 2016. We did this to limit the computational burden of our regressions to be able to provide first results. Although we are confident that they already paint an accurate picture, they have to be understood as preliminary.
here. First, the mean number of weekly album sales (Sales$_t$) in the full sample is 0.27. This is due to the long tail distribution of sales. As such, our sample contains high number of albums that exhibit zero weekly sales most of the time. In fact, 131,072 albums have at least one sale, while 23,771 albums have at least 10, and a mere 4,308 exhibit at least 50 weekly sales over the whole sample period. Second, when using weekly revenues as the dependent variable, the sample is reduced from 3,880,737 observations to 2,636,940 as we consider albums with a positive price only. In addition, we do not report absolute values as our price data contains several different currencies, which makes them uninformative. This is one reason to obtain elasticities rather than marginal effects here as it allows us to obtain results for the whole universe of albums, regardless of currency. Third, the number of observations is again greatly reduced to 354,098 when using an album’s buyers’ average deviation from their established tastes (Avg. ∆WAT), as our sample only contains weekly observations in which albums receive at least on sale. The mean of this variable is -1.03, thus very slightly negative when considering its range. In addition, the minimum and maximum values in the sample are -99.14 and 99.48, respectively. However, 90 % of the observations are much smaller in scope and fall within a range of -4.26 and 0.68. This is due to the

<table>
<thead>
<tr>
<th>Variable</th>
<th>Obs.</th>
<th>Logarithms</th>
<th>Absolutes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Min</td>
</tr>
<tr>
<td>Sales$_t$</td>
<td>3,880,737</td>
<td>0.10</td>
<td>0.34</td>
</tr>
<tr>
<td>Revenues$_t$</td>
<td>2,636,940</td>
<td>0.23</td>
<td>0.72</td>
</tr>
<tr>
<td>Avg. ∆WAT$_t$</td>
<td>354,098</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Rec$_t-1$</td>
<td>3,880,737</td>
<td>0.12</td>
<td>0.58</td>
</tr>
<tr>
<td>Rec$_t-2$</td>
<td>3,880,737</td>
<td>0.12</td>
<td>0.59</td>
</tr>
<tr>
<td>Rec$_t-3$</td>
<td>3,880,737</td>
<td>0.13</td>
<td>0.60</td>
</tr>
<tr>
<td>Rec$_t-4$</td>
<td>3,880,737</td>
<td>0.13</td>
<td>0.62</td>
</tr>
<tr>
<td>Sales$_t-1$</td>
<td>3,880,737</td>
<td>0.10</td>
<td>0.34</td>
</tr>
<tr>
<td>Sales$_t-2$</td>
<td>3,880,737</td>
<td>0.10</td>
<td>0.35</td>
</tr>
<tr>
<td>Sales$_t-3$</td>
<td>3,880,737</td>
<td>0.11</td>
<td>0.35</td>
</tr>
<tr>
<td>Sales$_t-4$</td>
<td>3,880,737</td>
<td>0.11</td>
<td>0.36</td>
</tr>
<tr>
<td>Prior Sales</td>
<td>3,880,737</td>
<td>2.18</td>
<td>1.24</td>
</tr>
<tr>
<td>Age</td>
<td>3,880,737</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Age squared</td>
<td>3,880,737</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 2 User Level: Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>Mean</th>
<th>S.D.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>∆WAT</td>
<td>909,963</td>
<td>-0.204</td>
<td>4.730</td>
<td>-99.55</td>
<td>99.64</td>
</tr>
<tr>
<td>Recommended (1 Day)</td>
<td>909,963</td>
<td>0.011</td>
<td>0.102</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Recommended (7 Day)</td>
<td>909,963</td>
<td>0.020</td>
<td>0.141</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Recommended (14 Day)</td>
<td>909,963</td>
<td>0.024</td>
<td>0.154</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Recommended (30 Day)</td>
<td>909,963</td>
<td>0.030</td>
<td>0.170</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Followees</td>
<td>909,963</td>
<td>6.699</td>
<td>32.332</td>
<td>0</td>
<td>995</td>
</tr>
<tr>
<td>WAT_t−1</td>
<td>909,963</td>
<td>4.323</td>
<td>11.712</td>
<td>0.029</td>
<td>100</td>
</tr>
<tr>
<td>Days Since Last</td>
<td>909,963</td>
<td>10.507</td>
<td>47.364</td>
<td>0</td>
<td>2,308</td>
</tr>
<tr>
<td>Prior Purchases</td>
<td>909,963</td>
<td>65.678</td>
<td>84.737</td>
<td>2</td>
<td>1,443</td>
</tr>
<tr>
<td>Repeat Purchase</td>
<td>909,963</td>
<td>0.781</td>
<td>0.414</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Album Age</td>
<td>909,963</td>
<td>521.228</td>
<td>504.632</td>
<td>0</td>
<td>2,647</td>
</tr>
<tr>
<td>Prior Album Sales</td>
<td>909,963</td>
<td>691.841</td>
<td>904.944</td>
<td>0</td>
<td>7,154</td>
</tr>
</tbody>
</table>

The construction of our measure of concentration – weighted average typicality –, which theoretically ranges from 0 to 100 (after scaling), but takes values between 0.03 (smallest value) and 14.62 (ninety-fifth percentile) in 95 % of all observations in our user-level data. Our main variables of interest is the number of times an album was recommended to users in the prior one, two, three, and four weeks (Rec_t−T). On average, this is the case 1.32, 1.38, 1.46, and 1.60 times, respectively.

5.1.2 User Level

Table 2 contains summary statistics for the variables used in our user-level analysis. First, the dependent variable we use throughout the analysis here is the focal user’s change in the concentration in music characteristics she consumes, ∆WAT. The mean value is -0.204. Similar to our discussion of the album-level equivalent of this variable, the minimum and maximum of the variable are -99.55 and 99.64, respectively. Again, the majority of values found in the sample are much smaller in scope. Specifically, the first percentile is -11.90, the fifth -2.69, the ninety-fifth 1.88, and the ninety-ninth is 10.61.

Next, the main variable of interest is whether or not an album had been recommended to the focal user before the transaction took place. Here, we consider different time windows: one, seven, 14, and 30 days. This is the case for 1.1 %, 2 %, 2.4 %, and 3 % of all transactions, respectively.
Arguably, these figures are rather small, indicating that the feature is not extensively used by all users on the platform. In fact, out of the 48,388 users in our analysis only 18,271 exhibit at least one followee.

### 5.2 Album Performance

First, we analyze the impact of a higher number of peer recommendations in prior weeks on album performance in the focal week. Throughout the analysis, we use heteroscedasticity-robust standard-errors, which are clustered at the level of the album.

Table 3 contains the results of the fixed-effects regression at the album-level. The first performance measure we use is the natural logarithm of an album’s weekly sales (Columns 1–3). Column 1 shows the results for the full sample of albums with at least one sale, column 2 for albums with at least 10, and column three for albums with at least 50 sales. As we expected,
the impact of peer recommendations on subsequent sales performance is positive. For instance, an 1 % increase in recommendations in week directly prior to the focal one leads to an 0.56 % increase in weekly sales (column 1, row 1) for the full sample. In addition, the effect gets weaker the more time has passed. I.e., again for the full sample, the effect is slightly weaker, on average, for peer recommendations two weeks prior (0.49 %), and becomes even weaker for those three weeks prior (0.33 %) and four weeks prior (0.24 %). A very similar pattern is observed when we limit the sample to albums with at least 10 sales. Here, the effect deteriorates over time from 0.74 % in one week prior to 0.26 % in four weeks prior. When limiting the sample to albums with at least 50 sales, we still observe a positive effect for peer recommendations in one and two weeks prior. However, it is stronger for the latter (1.48 %) than for the former (0.75 %), before becoming statistically indistinguishable from zero.
Next, we use an album’s weekly revenues as an alternative performance measure Table 3, Columns 4–6). Here, we observe a very similar pattern: The impact of peer recommendations is generally positive but deteriorates over time. For instance, in the full sample, the effect of a 1% increase in peer recommendations lead to an increase in weekly revenues of 2.16%, 1.28%, 1.38%, and 0.87% in the one, two, three, and four weeks prior, respectively.

We illustrate the estimated coefficients in figure 5. Panel 5a shows those for our sales regressions, and Panel 5b for our revenues regression. While showing similar patterns, note that the effect is generally stronger for revenues than for sales. Considering that we calculate weekly revenues as the product of an album’s (time-invariant) price and its weekly sales, this suggests that peer recommendations are more effective for higher-priced albums.

To further explore heterogeneous effects of an album’s price, we split the sample into albums that are offered at a positive price and those that are offered for free. Table 4 shows the results
using weekly sales as dependent variable. While the estimated coefficients for non-free albums (columns 1, 3, and 5) show similar patterns as before, the effect is not as clear for free albums (columns 2, 4, and 6). For peer recommendations one week prior to the focal one, the effect is either statistically indistinguishable from zero, or even slightly negative when only considering albums with a minimum of 50 sales. Still, the majority of estimated coefficient is positive at varying statistical significance level. From these results, we take that peer recommendations follow a clear and beneficial pattern for non-free albums, but not as much for free albums.

### 5.3 Consumption Variety

We start by analyzing whether or not and the extent to which users deviate from their established tastes when receiving a peer recommendation at the album level. We again run a fixed-effects regression with heteroscedasticity-robust standard errors, clustered at the album level. Further, we use the same set of independent and control variables as before, but now use buyers’ average ∆WAT
in a given week as our dependent variable. We expect a negative influence of peer recommendations, if they primarily convey information about an album’s quality, and a positive influence, if they are based on the match with a user’s taste. Results are presented in table 5. Across the board, the estimated coefficient is negative, indicating that buyers, on average, deviate from their established tastes, the more often the focal album has been recommended in prior weeks. Similar to the pattern we observed in our performance regression, this effect is weaker the further in the past recommendations happened. Aside from this deterioration over time, we also find that the effect becomes smaller as we restrict the sample towards albums exhibiting more total sales. This suggests that peer recommendations are especially beneficial for emerging or lesser-known artists.

In a next step, we move from the album to the user level to provide further and more detailed evidence for the mechanism at hand. We run regressions with individual transactions (thus user-album) being the unit of observation. Again, we use heteroscedasticity-robust standard errors throughout the analysis, which are clustered at the user level here.

Table 6 contains the results utilizing our instrumental variables approach in a set of two-stage least squares regressions. In line with the results obtained at the album level we now find a strong, statistically significant negative effect, providing further evidence that users are encouraged to deviate more from their established tastes when buying an album after it had been recommended to them. Further in line with album-level results, we confirm the deterioration of this effect over time: The longer the time window we consider for a recommendation to having taken place, the weaker it becomes. Specifically, when considering a 24-hour-window it is -16.79, for a seven-day window it is -9.23, it is -7.35 for a 14-day window, and -6.07 for a 30-day window.

At this stage we have found robust evidence that users do deviate from their established tastes following a peer recommendation at two levels of analysis. we now investigate possible long-run effects. First, we are interested in a potential learning effect. The intuition is that users might
Table 6 Deviation From Tastes: 2SLS

<table>
<thead>
<tr>
<th></th>
<th>∆WAT 1 Day</th>
<th>∆WAT 7 Days</th>
<th>∆WAT 14 Days</th>
<th>∆WAT 30 Days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recommended</td>
<td>-16.7922***</td>
<td>-9.2252***</td>
<td>-7.3496***</td>
<td>-6.0711***</td>
</tr>
<tr>
<td></td>
<td>(4.7978)</td>
<td>(2.4614)</td>
<td>(1.9340)</td>
<td>(1.5877)</td>
</tr>
<tr>
<td>Followees</td>
<td>0.0228***</td>
<td>0.0218***</td>
<td>0.0224***</td>
<td>0.0234***</td>
</tr>
<tr>
<td></td>
<td>(0.0063)</td>
<td>(0.0053)</td>
<td>(0.0053)</td>
<td>(0.0054)</td>
</tr>
<tr>
<td>Controls</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>User FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Album X Month FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Week FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>909,963</td>
<td>909,963</td>
<td>909,963</td>
<td>909,963</td>
</tr>
<tr>
<td>Kleibergen-Paap F Statistic</td>
<td>15.22</td>
<td>35.95</td>
<td>48.71</td>
<td>60.20</td>
</tr>
<tr>
<td>N (Users)</td>
<td>48,388</td>
<td>48,388</td>
<td>48,388</td>
<td>48,388</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

rely more on recommendations when they are relatively new to the platform, but might develop
the ability to discover new music characteristics on their own over time. To test this, we run a set
of regressions that include an interaction term of the recommendation dummy with the number a
user’s prior purchases. If they indeed become less reliant on recommendations for music discovery,
we would expect – in light of the negative main effect – a positive effect of this interaction term.
Results are presented in table 7. Indeed, the estimated coefficient carries a positive sign across
all considered time windows, suggesting the existence of a learning effect. In line with previous
findings, this effect deteriorates over time, as well\(^{20}\).

Robustness Checks

Due to space constraints we omitted the discussion of our robustness checks with this submission.
However, we want to note that our findings are robust to (i) using the Herfindahl-Hirschman-

\(^{20}\)Here, the results for a one-day window have to be regarded with caution, as the first-stage F-statistic is only at
9.211. Therefore, the instruments we use might be weak, leading to a bias in the results. First-stage statistics for
the other three time windows, however, do not raise concerns.
Table 7 Learning Effects

<table>
<thead>
<tr>
<th></th>
<th>∆WAT</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) 1 Day</td>
<td>(2) 7 Days</td>
<td>(3) 14 Days</td>
<td>(4) 30 Days</td>
</tr>
<tr>
<td>Recommended</td>
<td>−23.1553***</td>
<td>−11.0970**</td>
<td>−8.7771***</td>
<td>−7.2670***</td>
</tr>
<tr>
<td></td>
<td>(5.2918)</td>
<td>(2.7064)</td>
<td>(2.1172)</td>
<td>(1.6928)</td>
</tr>
<tr>
<td>Recommended X Prior Purchases</td>
<td>0.1252***</td>
<td>0.0714***</td>
<td>0.0613***</td>
<td>0.0498***</td>
</tr>
<tr>
<td></td>
<td>(0.0247)</td>
<td>(0.0139)</td>
<td>(2.1172)</td>
<td>(0.0103)</td>
</tr>
<tr>
<td>Prior Purchases</td>
<td>−0.0065***</td>
<td>−0.0066***</td>
<td>−0.0065***</td>
<td>−0.0064***</td>
</tr>
<tr>
<td></td>
<td>(0.0007)</td>
<td>(0.0007)</td>
<td>(0.0007)</td>
<td>(0.0007)</td>
</tr>
<tr>
<td>Controls</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>User FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Album X Month FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Week FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>909,963</td>
<td>909,963</td>
<td>909,963</td>
<td>909,963</td>
</tr>
<tr>
<td>Kleibergen-Paap F Statistic</td>
<td>9.211</td>
<td>17.16</td>
<td>17.23</td>
<td>18.11</td>
</tr>
<tr>
<td>N (Users)</td>
<td>48,388</td>
<td>48,388</td>
<td>48,388</td>
<td>48,388</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Index (HHI) as an alternative measure of concentration, (ii) excluding all users that do not follow any other user from the analysis, and (iii) excluding all observations where a user made multiple purchases on the same day.

A detailed discussion as well as regression tables are available from the authors upon request.

6 Discussion and Conclusion

In this study, we empirically analyzed the impact of a peer-based recommendation system on the performance of complementary products and the consumption patterns of users on a digital platform. Using a simple theoretical model framework, we argued that they will increase product performance regardless of the information conveyed by peer recommendations, but that the change in consumption patterns depend on this information. Specifically, we argued that users of a digital platform will extend the scope of the varieties – or horizontal characteristics – they consume if peer recommendations reveal products of high quality, and that they will become more concentrated in
those varieties if products close to a user’s tastes are revealed. Our two-part analysis at the product and at the user level in the context of a digital music platform provides strong evidence that album performance indeed increases, both in terms of sales and revenues, and that users extend the scope of consumed varieties with purchases of products that have been recommended to them previously. We found further evidence suggesting that the performance effect is stronger for products that are sold at a higher price, and for more popular products. In addition, we provided evidence that users become less reliant on peer recommendations in their exploration efforts as they make purchases on the platform. Lastly, we found that the effect of peer recommendation deteriorates over time.

Our results highlight important aspects of the management and governance of platform ecosystems. First, a peer-based recommendation system makes product discovery easier for users, which translates into increased performance of complementary products. This carries important implications for the platform level. Facilitating sales of complementary products through the platform will have a direct effect on its profits if it claims a share of the proceedings. But even if that is not the case, a higher prospect of profitability for complementary products will increase their incentive to join the platform in the first place. On the one hand, this leads to increased platform revenues if an entry fee is charged. On the other hand, this will provide a competitive advantage over competing platforms. Second, the fact that peer-recommendations encourage users to extend the scope of the varieties they consume implies that users generally take a broader set of complementary products under consideration. This has the potential to further increase sales, and subsequently translate into increased platform profits and a competitive edge in the market. Third, extant literature treats what determines the strength of both direct and indirect network effects as a black box (McIntyre & Srinivasan 2017). We showed that peer recommendations can be an effective way to increase a single user’s benefit of the installed base of peers through increased information flows. In addition, a large number of variety of complementary products can only
translate into increased value for users if they are able to find them in the first place. As such, peer recommendations increase the number of such products each user can actually process, which should increase the value of the platform as such as well.

Further, we further contribute to the literature around recommendation systems. Prior studies have largely focused on algorithm-based recommendation systems and their impact on the macro-level distribution of sales. Specifically, we contribute in two ways. First, we provide insights about the effect of a recommendation system that builds on peer influence rather than information about products’ popularity or user’s past behavior. Second, we provide insights about a possible micro-level mechanism driving macro-level outcomes in terms of the distribution of sales. Specifically, we highlight that revealing high quality products to consumers may lead to a decrease of concentration in sales. Our study also contains a contribution to the literature on peer influence. So far, empirical evidence in the context of digital platforms had been scarce, especially regarding the influence on consumption patterns.

Our study contains several limitations as well. First, Bandcamp is an empirical setting that shows some particularities. Its focus on independent artists sets it apart from comparable, more popular digital music platforms such as Spotify or Amazon Music. As such, the external validity of our results might not as strong. In addition, the emergence of streaming platforms shifted the focus away from owning albums to listening to single songs (Datta et al. 2018). On Bandcamp, however, users still buy albums. Still, we believe that idiosyncratic tastes and exploration of new varieties are an important determinant of value to consumers, and that our results on the effect of peer recommendations carry general implications here. Second, due to the fact that we obtained a snapshot of the platform, we do not observe price changes. Although we are convinced that artists on Bandcamp are not strategic in their price setting behavior to the extent that it would bias our results, we are not able to completely rule out this possibility. Third, the instruments we
use in our identification strategy might be imperfect to the extent that a likely homophily travels across multiple nodes in the network of users, leading to a sustained selection bias in our user-level analysis. However, given how the estimated coefficients differ between our results obtained from OLS regressions and those obtained utilizing our instrumental variables approach, we are convinced that the conclusions drawn from our analysis are valid, at the very least qualitatively. We do acknowledge, however, that they might constitute a lower bound.

References


**URL:** https://www.ncbi.nlm.nih.gov/pubmed/14744228


58(5), 892–912.