International Journal of E-Business Research
Volume 18 • Issue 1

Peer Influence in the Adoption of Video Games

Yang Li, Montclair State University, USA*
Jiangen He, University of Tennessee, Knoxville, USA
https://orcid.org/0000-0002-3950-6098
Chuanren Liu, University of Tennessee, Knoxville, USA
Yanni Ping, St. John's University, USA

ABSTRACT

The authors investigate how peer influence affects customers' product adoption behaviors in emerging video game platforms. Understanding peer influence is critical to motivating users' willingness to purchase and improving game publishers' marketing performance. While similarities between socially linked users can be viewed as a consequence of social influence, homophily may also contribute to such phenomenon, causing identification difficulties in observational studies. Using data from the world's largest digital distribution platform for video games, the authors leverage state-of-the-art recommender system algorithms and propose an innovative framework to identify social influence in the adoption of video games when a confounding homophily effect is present. The results show that peer influence has a positive impact on platform users' adoption behaviors (i.e., a user tends to adopt a video game that has been purchased by his peers). This study also finds that peer influence would have been overestimated if homophily was not properly controlled.

KEYWORDS

Peer Influence, Product Adoption, Propensity Score Matching, Recommender System, Video Game

INTRODUCTION

Video gaming has been a large, steadily growing industry over the past few years. During the recent COVID-19 pandemic, people spent more time at home to prevent the spread of the virus. Accordingly, consumers are more likely to rely on video games to manage their mental health (Kim 2021), drastically increasing video game sales (Entertainment Software Association 2020). Unlike traditional brick-and-mortar stores, online video gaming platforms allow interactions between consumers. Specifically, an online game platform can be considered as a community of friends with shared interests. Users can view their friends' purchased games, gameplay statistics, and game screenshots. They can also send or receive messages and get personalized notifications about their friends' games when they are playing. These interactions make social influence more salient in online video gaming consumers' decision-making process.

DOI: 10.4018/IJEBR.309399
*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

This article aims to quantify the role of peer influence in the adoption of video games. Although the effect of social influence has been widely discussed, the identification of peer influence remains challenging for empirical studies using observational data (Manski, 1993; Bapna and Umyarov, 2015). The main difficulty is that product adoption can also be explained by homophily, in which dyadic similarities between users create correlated outcome patterns among friends. However, such patterns merely mimic viral contagions without direct causal influence (McPherson et al., 2001). Therefore, when a user purchases a new game, it is important to know whether peer influence or internal similarity between the user and their friends drives game adoption. The present study attempts to answer the following research questions: (1) Does peer influence exist in the purchase behavior of online video games? (2) Can researchers disentangle homophily from peer influence?

This study captures the similarity of users through consumer tastes, which are modeled as latent features based on past purchases and the recorded game genres. Game genres present a rich collection of information describing gameplay, such as action, strategy, or casual, and are closely associated with consumers' personality and psychological characteristics (Von Der Heiden et al. 2019; Potard et al. 2020). Accordingly, we utilize a model where the internal aspects of consumers will be incorporated into the 'genre-based' user similarity measures while we will reflect the external aspects through peer influence and other observable user features. The approach is similar to that taken in Sheu, Chu, and Wang (2017), where the consumers' internal personal cognition was evaluated through sensory, feel, think, act, and relate experiences in the games. If peer influence can be identified in the adoption of games, game publishers can generate more revenue by targeting influential users or building product features that facilitate peer interactions. Alternatively, if homophily dominates consumer purchase behavior similarity, game publishers should focus more on analyzing the shared taste of users who purchased similar games.

This study bridges two streams of research on the video game industry and product adoption. On one hand, there is scant empirical research investigating the video game industry. Additionally, it focuses more on various factors affecting game sales (Zhu and Zhang, 2010; Cox, 2014). On the other hand, previous literature has examined the effect of peer influence on product adoption in domains like technology adoption (Brown and Venkatesh, 2005), service adoption (Bapna and Umyarov, 2015), and application adoption (Aral and Walker, 2011; Davin et al., 2014). However, none of these explored the effect of social influence on video game platforms. To the best of our knowledge, this is the first study to investigate the impact of social influence on video game platforms.

It is worth mentioning that video game platforms differ from ordinary social network websites. Unlike social media platforms like Facebook or Twitter, online game platforms do not require pre-established interpersonal relationships between users. As a result, these platforms rarely consider social factors like relationships and trust among users. Therefore, the power of social influence remains to be exploited (Li et al., 2013) in video gaming platforms, which motivates this empirical study.

This paper uses data from the world's largest video game retailing platform to examine the effect of peer influence on the adoption of video games. The authors propose an innovative framework that integrates propensity score matching (PSM) (Rosenbaum & Rubin, 1983) and a recommender system to identify social influence in the adoption of video games when a confounding homophily effect is present. The basic idea is to reveal users' gaming preferences that drive the formation of friendships among similar users, as hypothesized by the homophily assumption. Homophily is then controlled by matching users into control/treatment groups with similar tastes measured by latent features. By examining the purchasing behavior of users with friends who adopted the game (treatment group) and those without friends who adopted the game (control group), the authors can estimate the direction and magnitude of social influence in video game adoption. Then, the authors propose similarity-enhanced matrix factorization (SEMF) and model users using low-rank latent features, instead of representing users with the games they have purchased. The authors choose the matrix factorization approach due to its various advantages. First, as a dimension reduction technique, matrix factorization can effectively alleviate the sparsity problem (Chen et al., 2011), which arises when each user only interacts with a small portion of available items. A user considers a few dozen games on video game platforms, while the platform contains thousands of games. This sparsity may lead to zero values in similarity calculations between users, making Boolean feature vectors of purchased games less effective in identifying users' tastes. Therefore, this article represents users by low-rank features derived from matrix factorization. Second, interactions between users and items may lead to missing values that are not-at-random (MNAR) (Little & Rubin, 2014). For example, users typically purchase and play games they like; they rarely purchase games they do not like. Therefore, ignoring the mechanism of missing values may introduce selection bias into the empirical study. This article addresses this issue by assigning a weighting matrix (based on features less affected by selection bias) in a formulation of matrix factorization. Third, side information describes games in detail. The authors used side information in this study to solve a regularized optimization problem to accurately measure users' tastes.

The remainder of this chapter is organized as follows. Section 2 reviews the related work. Section 3 shows the details of the proposed framework, integrating the recommender system and PSM for peer influence identification. Section 4 described the data. The performance of the recommender system and the effect of peer influence are presented in Section 5. Finally, the authors conclude and discuss their work in Section 6.

BACKGROUND

In this section, the authors review the main research works in peer influence identification and recommender systems. In his seminal paper, Manski (1993) analyzes the problem of identifying endogenous social effects based on observed behavior. Since his report, various approaches have been proposed to use observational data for peer influence identification. The basic idea is to control homophily, exogenous factors, or factors other than influence (Ma et al., 2015). Nair et al. (2010) quantify peer influence in physician prescription behavior using a fixed-effect model to control for possible homophily. Ma et al. (2015) separate homophily from influence in the context of caller ring-back tones purchases by considering the dynamic purchase pattern in a choice model.

This study is mostly related to the studies that identify peer influence by incorporating user-level covariates in regression and matching. For example, Aral et al. (2009) develop a dynamic matched sample estimation framework to distinguish influence and homophily effects. To accurately proxy individuals' tastes and preferences, they use a set of 46 covariates, including data on users' demographics, mobile device usage, and page views in matching. They find strong evidence of social influence in the adoption of a mobile application. Nitzan and Libai (2011) investigate the role of social influence in customer retention of a mobile service provider. They include several demographic variables and usage characteristics to account for similarities in customers' tastes. However, user-level demographics are usually difficult to obtain on online platforms due to privacy policies. In this regard, researchers must make the best use of the available data on user-item interaction in the analysis.

Collaborative filtering (CF) is the most widely used approach for mining user-item interactions (Goldberg et al., 1992). It predicts the ratings an individual would give to items by analyzing preference similarity across individuals. However, in the context of product adoption, explicit ratings on products are hard to collect; users' adoption decisions implicitly express user-item interactions. To address the implicit feedback, Hu et al. (2008) propose to use a weighting matrix that measures confidence in observing positive and negative examples. This study extends their work and improves the prediction performance by incorporating game genre information in matrix factorization. Since game genres present a rich collection of gameplay information (describing gameplay, such as action, strategy, or casual) and are closely associated with consumers' personality characteristics and psychological functioning (von der Heiden et al. 2019; Potard et al. 2020), users' tastes can be accurately measured by the low-rank user vectors, which reflect both external and internal user aspects. To the best of our knowledge, there is no prior work that incorporates user-item interactions when identifying peer influence.

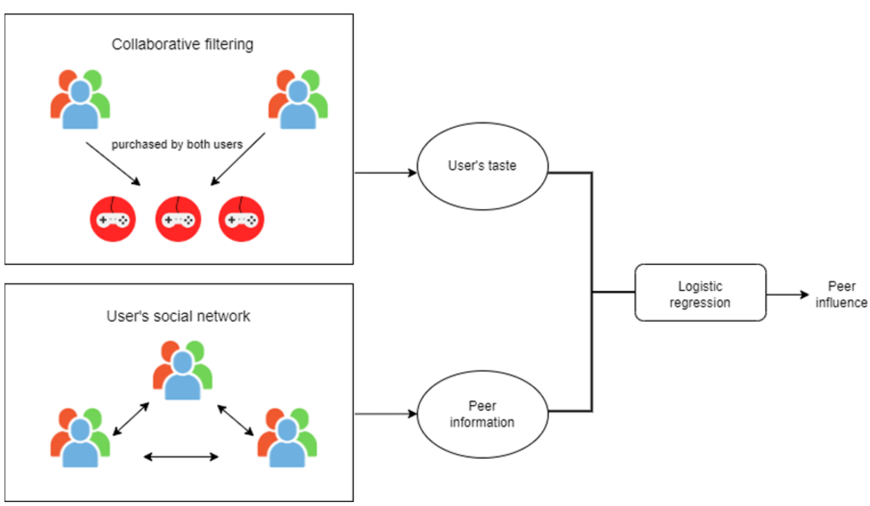
RESEARCH METHODOLOGY

The authors present a framework to identify peer influence in game user adoption behavior in this section. The confounding factor, homophily, is controlled by users' tastes. Specifically, a new collaborative filtering model, namely SEMF, is proposed to generate low-rank latent vectors that capture users' tastes. A logistic regression model is then used to estimate the effect of peer influence. The proposed framework is summarized in Figure 1.

Collaborative Filtering

This section begins with an analysis that captures users' tastes. Intuitively, a user's preference for video games can be measured by their purchased games. Thus, the authors can use a vector of zeros and ones indicating unpurchased and purchased games, respectively, to proxy a user's taste. This approach, however, is problematic in practice. First, there are thousands of games in the dataset, which will lead to long user vectors with large dimensions, making PSM based on these vectors less efficient (Hill et al., 2011). As a result, some dimension reduction techniques are necessary to derive low-rank user vectors. Second, the taste measurement might be biased if this study only looks at games purchased by a user. For example, not buying a game can stem from reasons beyond not liking it. The user might be unaware of the game's existence or unable to afford it. Besides, a user may have different preference levels for the purchased games. They may like some purchased games more than others. In other words, there is a lack of active rating on products compared to movie rating websites, resulting in the problem of implicit feedback. This problem requires a weighting scheme on different games to accurately measure users' tastes. Moreover, the simple vectorization of purchased games ignores the rich side information about video games. The authors argue that game genres provide evidence for users' gaming preferences. In the context of video gaming, the most reliable and content-rich descriptors of a game lie in its genres voted by users. Meanwhile, video game platforms recommend products to their users based on a collection of predefined genres that describe gameplay, including action, strategy, casual, etc. These factors show that the information embedded in game genres plays a vital role in customers' buying behaviors. Therefore, the measure's

Figure 1. The framework to identify peer influence



performance on users' tastes can be improved by incorporating this information. This study develops a collaborative filtering model for implicit feedback with side information to address concerns related to simple vectorization. The derived low-rank vectors are used to measure users' tastes. The following notations are used throughout this paper:

- $U = \{u_1, u_2, \ldots, u_m\}$: set of m users;
- $I = \{i_1, i_2, \ldots, i_n\}$: set of n items (games);
- x_u: k-dimensional low-rank vector representing a user's preference over latent factors;
- y_i: k-dimensional low-rank vector representing an item (game);
- $X = \left(x_1, \ldots, x_u, \ldots, x_m\right)^T$: $m \times k$ low-rank user matrix, where T is matrix transpose;
- $Y = \left(y_1, \ldots, y_i, \ldots, y_n\right)^T$: $n \times k$ low-rank item (game) matrix;
- $R = \{r_{ui}\}_{m \times n}$: user-item preference matrix, where $r_{ui} = 1$ if the user u purchased item i, and $r_{ui} = 0$ otherwise;
- $P = \left(p_1, \ldots, p_i, \ldots\right)^T$: game information matrix;
- $Q = \left(q_1, \ldots, q_u, \ldots\right)^T$: user information matrix;
- $S = \{s_{ij}\}_{n \times n}$: item-item similarity matrix, where s_{ij} is the similarity between item i and item j;
- $T = \{t_{ui}\}_{m \times n}$: user-item similarity matrix, where t_{ui} is the similarity between item u and item i;
- $W = \{w_{ui}\}_{m \times n}$: weight matrix;
- $\left\| \bullet \right\|_F$: Frobenius norm of a matrix.

Matrix factorization models have been widely used for recommendations (Koren et al., 2009). In cases of the explicit rating, where users actively evaluate items by giving both positive (likes) and negative (dislikes) examples as measured by numeric ratings in a score matrix R, the authors can approximate R by decomposing it into two low-rank matrices X and Y that represent users and items, respectively. However, with implicit user feedback, the dataset only consists of binary values reflecting a user's action or inaction. For example, in this study, the dataset only indicates if a user has purchased a game without explicit ratings. A naive approach to this problem is to simply treat purchased games as positive examples with value 1 in R. This study then treats unpurchased games as negative examples with value 0 (i.e., AMAN). A better strategy is to assign a weight matrix (i.e., wAMAN) to control the relative contribution of quadratic terms to the loss function:

$$J\left(X, Y\right) = \sum\nolimits_{u,i} w_{ui} \left(r_{ui} - x_u^T y_i\right)^2 + \lambda \left(\sum\nolimits_u \left\| x_u \right\|_F^2 + \sum\nolimits_i \left\| y_i \right\|_F^2\right) \tag{1}$$

where λ is a regularization parameter that prevents overfitting. The weight w_{ui} can be considered as the confidence level that represents the confidence in the value of r_{iu}. This model gives large values to w_{ui} if user u tends to like item i (and vice versa).

The authors explore side information embedded in games to improve matrix factorization performance and determine the weight matrix. In addition, this study uses genres as the source of information for a specific video game because they play an essential role in effective marketing (Quax et al., 2013; Hofacker et al., 2016).

Formulating Game Information Vectors

In the dataset, each user, regardless of whether they own the game, can vote for genres that they think are most relevant to the game from a collection of candidates. By taking advantage of the wisdom of the crowd, the assembling of user-generated genre tags will accurately characterize a video game in terms of gameplay. In this regard, a game can be represented by a collection of genre tags (just as a document can be represented by a bag of words). The model denotes by:

$$G_i = \left\{z_{i1}, z_{i2}, \ldots, z_{ig}, \ldots\right\} \tag{2}$$

the collection of genre tags of game i with z_{ig} as words. Tf-idf weightings can be used to build the item information matrix. Therefore, words that are descriptive and unique to a game have large weights. For a given word z_{ig}, let:

$$idf\left(z_{ig}\right) = log\left(\left|I\right| / \left|z_{ig} \in V_j\right|\right) \tag{3}$$

where $\left|I\right|$ is the number of games and $\left|z_{ig} \in V_j\right|$ is the number of games whose vocabularies V_j contain z_{ig}. Let:

$$p_{ig} = idf\left(z_{ig}\right) \times tf\left(z_{ig}\right) \tag{4}$$

where $tf\left(z_{ig}\right) = \left|z_{ig} \in G_i\right| / \left|G_i\right|$ is the term frequency of word z_{ig} in game i. By considering game genres, the model formally formulates the vector representation of games as:

$$p_i = \left(p_{i1}, p_{i2}, \ldots, p_{ig}, \ldots\right) \tag{5}$$

This study further develops a similarity measure between games using cosine similarity:

$$s_{ij} = cos\left(p_i, p_j\right) \tag{6}$$

where s_{ij} is the similarity between game i and game j.

Formulating User Information Vectors

Like game information vectors, user information vectors can be defined using a collection of genre tags of purchased games. Let:

$$q_u = \sum_{i \in I_u} p_i \tag{7}$$

where q_u is the aggregation of game information vectors owned by user u. This allows us to make a comparison between users and games. Thus, a similarity measure between users and games can be defined as follows:

$$t_{ui} = cos\left(q_u, p_i\right) \tag{8}$$

where t_{uj} is the cosine similarity between user u and game i. The weighting scheme can be determined in such a way that more weights will be given to cases where users are similar to games in terms of genres.

Similarity-Enhanced Matrix Factorization (SEMF)

This section now shows how to incorporate rich item information into the factorization model. Motivated by the assumption that similarity between items should be consistent regardless of the representation method, this study constrains latent features as factorized matrices of the item similarity matrix. In other words, if two items are similar in terms of genre information, they should also be similar in terms of latent features. Mathematically, this study aims to find a solution by minimizing the following loss function:

$$J\left(X,Y\right) = \sum_{u,i} w_{ui}\left(r_{ui} - x_u^T y_i\right)^2 + \sum_{j>i}\left(s_{ij} - y_i^T y_j\right)^2 + \lambda\left(\sum_u \left\|x_u\right\|_F^2 + \sum_i \left\|y_i\right\|_F^2\right) \tag{9}$$

This model assigns weights to negative examples by looking at the similarity between the user and the game. The more similar, the fewer weights should be assigned to that negative example. Hence, for negative examples:

$$w_{ui} = 1 - cos\left(q_u, p_i\right) = 1 - t_{ui} \tag{10}$$

For positive examples, weights are determined by considering the interaction between the user and the game. Let f_{ui} be the total amount of time user u has spent on game i. The more time a user has spent on a game, the more interested they are in the game. As a result, weights for positive examples are defined as:

$$w_{ui} = 1 + \alpha f_{ui} \tag{11}$$

where α is a positive number that controls the increase rate of confidence.

The low-ranked matrices can be solved by alternative least square (ALS). To solve X, the model first fixes Y and take derivatives of the loss function with respect to x_u:

$$\partial J / \partial x_u = -2\sum_i w_{ui} y_i\left(r_{ui} - x_u^T y_i\right) + 2\lambda x_u \tag{12}$$

Let the partial derivative $\partial J / \partial x_u = 0$, we get:

$$x_u = \left(\lambda I_d + \sum_i w_{ui} y_i y_i^T\right)^{-1}\left(\sum_i w_{ui} r_{ui} y_i\right) \tag{13}$$

where I_d is an identity matrix of rank d. Similarly, let $\partial J / \partial y_i = 0$, we get:

$$y_i = \left(\sum\nolimits_u w_{ui} x_u x_u^T + \sum\nolimits_j y_j y_j^T + \lambda I_d \right)^{-1} \left(\sum\nolimits_u w_{ui} r_{ui} x_u + \sum\nolimits_j s_{ij} y_i \right) \tag{14}$$

The update is repeated until convergence. The algorithm is summarized as follows:

Initialize x_u and y_i with random numbers
If $r_{ui} = 0$ then
    $w_{ui} = 1 - t_{ui}$
Else
    $w_{ui} = 1 + \alpha f_{ui}$
End if
Repeat
    $x_u = \left(\lambda I_d + \sum\nolimits_i w_{ui} y_i y_i^T \right)^{-1} \left(\sum\nolimits_i w_{ui} r_{ui} y_i \right)$
    $y_i = \left(\sum\nolimits_u w_{ui} x_u x_u^T + \sum\nolimits_j y_j y_j^T + \lambda I_d \right)^{-1} \left(\sum\nolimits_u w_{ui} r_{ui} x_u + \sum\nolimits_j s_{ij} y_i \right)$
Until x_u, y_i converge
Return x_u, y_i

Identifying Peer Influence

Once the model has derived low-rank latent features x_u that capture users' tastes in phase 1, the effect of peer influence (i.e., if a user has friends who purchased a specific game) can be estimated by matching users who have adopter friends to users who have no adopter friends based on their tastes. As discussed in previous literature (Aral et al., 2009), the main difficulty in delivering an unbiased estimate of the contagion effect is that studies using observational data would never observe the buying behavior of a user who has adopter friends had he not had adopter friends. This is also called the counterfactual problem. This article addresses this issue by assuming unconfoundedness (Rubin, 1978), which states that the outcome is conditionally independent of treatment after controlling for a rich set of observable characteristics. To this end, this study defines the treatment as having at least one friend who has purchased a game in Phase 1. Treated examples are then matched with users who were likely to be treated, conditional on latent features x_u, but who had not been treated. Logistic regression is used to estimate the propensity of being treated as follows:

$$P(treat_u = 1 \mid x_u) = exp\left(\alpha_u + \beta_u^T x_u + \varepsilon_u\right) / \left[1 + exp\left(\alpha_u + \beta_u^T x_u + \varepsilon_u\right)\right] \tag{15}$$

The adoption behavior of peers is measured by the treatment effect $treat_u$, a binary variable that equals one if the user has adopter friends of the game and zero otherwise. α_u and β_u are intercepts and coefficients. ε_u are idiosyncratic errors. Nearest neighbor matching is then employed to form the treatment and the control group. Therefore, for a single game, the effect of social influence can be estimated using the following specification:

$$logit\ D_u = \delta \times treat_u + C_u + \varepsilon_u \tag{16}$$

where D_u is a binary variable that indicates if user u eventually buys the game in phase 2. C_u are the control variables, including the number of games owned by the user, the number of friends of the user, and the age of the account. ε_u are idiosyncratic errors. The estimate of the treatment effect associated with the game is δ. A significant and positive value of δ indicates that having adopter friends will increase purchase likelihood, confirming the existence of peer influence.

DATA

This article uses data from STEAM, the world's most extensive digital distribution and gaming platform for video games owned and operated by the Valve Corporation. The digital rights management (DRM) system of STEAM restricts the unauthorized distribution of games. Therefore, users must log into the client portal to download and play the purchased games associated with their accounts. In addition to its digital store that sells various video games from different publishers, STEAM provides networking features like friends, groups, and messaging services. For example, after a user has launched a game, the STEAM Client runs alongside the current game and sends notifications on his friends' game-playing activities. The existence of these notifications promotes peer influence among friends.

This study uses both user-level data that contains users' purchase history and friends' information and game-level data with features like genre, price, publisher/developer, and release date. User-level data is obtained from a previous study (O'Neill et al., 2016). In their original research, O'Neill et al. (2016) crawled the STEAM website to collect data on 108.7 million user accounts, 196.37 million bidirectional friendships, and 384.3 million owned games with 2,814 distinct game titles (The full dataset is available at https://steam.internet.byu.edu/). This user-level data collection was conducted twice in 2013 as the initial crawl (phase 1) and once in 2014 as the follow-up crawl (phase 2). This dynamic network setting allows us to examine how various factors in phase 1 impact customers' purchasing behaviors in phase 2. In addition, this study investigates the peer influence in the adoption of the game Dota2, the most popular online video game released in 2013. The sample is tailored to include 50,000 randomly selected users and examine the effect of peer influence on their game adoption. For game-level data, the authors use the API provided by STEAMSpy (available at https://steamspy.com/), a third-party website that provides statistics of games on STEAM, to crawl the games' features. This article considers game genres as the source of rich information that describes games in terms of gameplay. With the emergence of contemporary video games that incorporate more technologies and aspects of gameplay, traditional game classification based on simple genre labels (e.g., Action, Strategy, Simulation, and Role-playing) fails to capture the novel complexity inherent in these games (Clarke et al., 2017). To solve this problem, STEAM leverages the wisdom of the crowd by allowing users to vote for genres of a game from a collection of 247 predefined sub-genre tags. This study, thus, collects the user-generated tags for the games in the sample.

RESULTS

The identification of peer influence depends on the ability of the proposed model to capture users' gaming preferences. Mean percentage ranking (MPR) is used to evaluate the performance of the proposed SEMF. For each user u in the test set, SEMF generates a ranked list of the games sorted by preference. Let $rank_u = \left(rank_{u1}, ..., rank_{ui}, ...\right)^T$ denote the $m \times 1$ ranked list comprised of percentile rankings of the games for user u. $rank_{ui} = 0\%$ means that the game is most preferred by the user while $rank_{ui} = 100\%$ means that it is least preferred. I_u is a $m \times 1$ vector comprised of binary variables of user u's game adoption behavior. The quality measure is then defined as the average percentage ranking for all users and items:

$$MPR = \frac{\sum_u rank_u^T I_u}{\sum_u I_u^T I_u} \tag{17}$$

Lower values of MPR are desirable, as they indicate the recommendation lists would capture users' preference by prioritizing purchased games. It is expected that a randomly recommended list would have an MPR around 50%. Figure 2 presents an intuitive example of calculating MPR in a simple scenario with two users and four games. $I_{u1} = \left(1001\right)^T$ shows user 1 has purchased game A and game D. $I_{u2} = \left(0100\right)^T$ shows user 2 has purchased game B. Note that a high percentile ranking 4/4, meaning dislike, is assigned to the adopted game B of user 2. Consequently, the calculated MPR is above 50%, indicating poor performance of the example recommender system.

Models with varying numbers of latent factors k, ranging from 5 to 50, are tested. Figure 3 shows the results. In general, MPR is improved as k increases. The figure shows a drastic decrease in the value of MPR when k increases from 5 to 20. The MPR can achieve 8%, which is much lower than 50% achieved by a random predictor. The improvement of the proposed model is limited as k increases. Thus, this study chooses $k = 20$. This section also compares the performance of the proposed SEMF to the widely used wAMAN. It can be shown that MPR gains 2% by incorporating the rich collection of user-defined tags. In sum, the proposed SEMF can capture users' tastes precisely.

Having derived the low-rank features x_u that accurately proxy users' gaming preference, the authors then conduct PSM and match users who do not have adopter friends (i.e., control units) to users who have adopter friends (i.e., treated units) based on x_u. The matched sample consists of 6,682 treated units and 26,728 control units. Figure 4 shows the histograms before and after matching. It can be shown that the propensity score distribution for the treated and control groups is quite

Figure 2. A numerical example of MPR

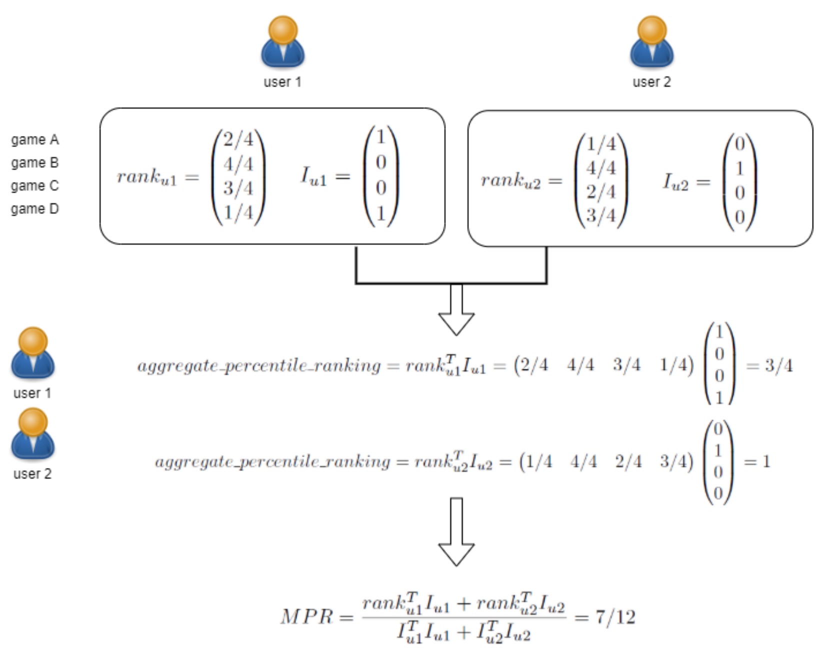


Figure 3. Impact of varying the number of latent factors

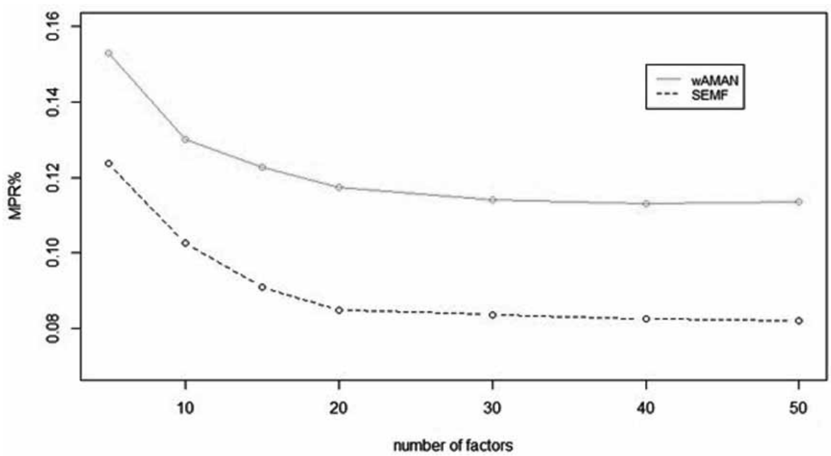
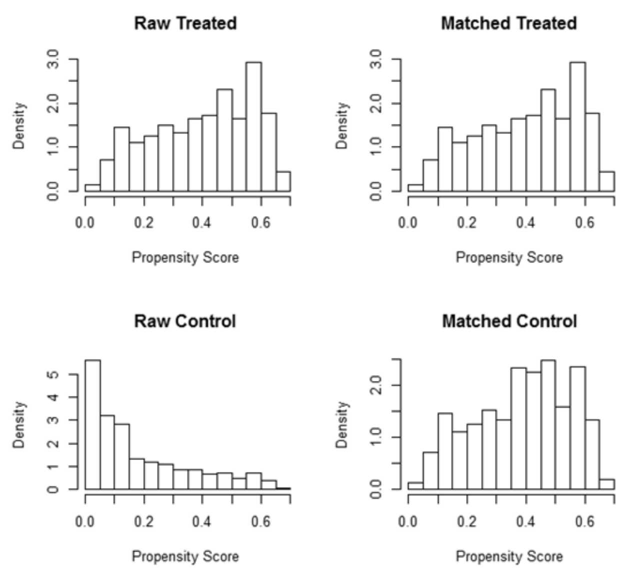


Figure 4. Histograms of propensity scores before and after matching



different before matching due to unbalanced covariates. After matching, propensity scores from both groups are similar, indicating that the matching is successful.

The main findings on the effect of peer influence are presented in Table 1. Column 1 shows that the impact of having adopter friends on users' purchase decisions is significantly positive, with value 0.21. That is, the odds of purchasing the game for users with adopter friends over the odds of purchasing the game for users without adopter friends is $exp\left(0.21\right) = 1.23$. Column 2 reports the logistic regression results on the raw data without matching. The estimated effect of peer influence is 0.26. As illustrated in previous studies (Tucker, 2008; Hartman, 2010), ignoring homophily would give rise to an upward bias of the effect of social influence. In the sample of this study, the bias results in a 24% inflation in the estimate.

Table 1. Main effects of peer influence

Logit Regression	(1) Matched Sample	(2) Raw Sample
Peer Influence	**0.214*** (0.034)**	**0.255*** (0.033)**
Number of purchased games	-0.077*** (0.024)	-0.078*** (0.023)
Number of friends	0.158*** (0.017)	0.153*** (0.016)
Age of account	-0.011 (0.011)	-0.012 (0.011)
Intercept	-1.762*** (0.115)	-1.753*** (0.114)
Number of observations	43,410	50,000
$P > \chi^2$	0.0000	0.0000

Standard errors in parentheses
*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

To further test the robustness of the proposed method, the authors define the treatment as having at least T adopter friends. Logistic regression is used on the matched sample with an array of threshold values, including $T = 1$, $T = 3$, and $T = 5$. As Table 2 shows, the estimates of the effect of peer influence are robust as indicated by similar amplitudes of significant coefficients.

In summary, the proposed SEMF model can accurately predict users' gaming preferences by recommending relevant games to users. Based on the latent features that proxy users' gaming preference, the proposed model controls for homophily by matching users with adopter friends to users without adopter friends. Logistic regression is then used to estimate the effect of peer influence.

Table 2. Effects of peer influence based on various thresholds

Logit regression	(1) Matched sample T=1	(2) Matched sample T=3	(3) Matched sample T=5
Peer Influence	**0.214*** (0.034)**	**0.193*** (0.031)**	**0.159*** (0.033)**
Number of purchased games	-0.077*** (0.024)	-0.072*** (0.022)	-0.075*** (0.021)
Number of friends	0.158*** (0.017)	0.159*** (0.017)	0.159*** (0.016)
Age of account	-0.011 (0.011)	-0.011 (0.011)	-0.012 (0.011)
Intercept	-1.762*** (0.115)	-1.761*** (0.114)	-1.764*** (0.112)
Number of observations	43,410	43,410	43,410
$P > \chi^2$	0.0000	0.0000	0.0000

Standard errors in parentheses
*** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$

The findings indicate that peer influence is consistently and positively associated with purchasing behavior. Moreover, the impact of peer influence will be overestimated if homophily is ignored.

CONCLUSION

This article studies peer influence on the adoption of video games. Utilizing the state-of-the-art recommender system algorithm, the authors can control for homophily and identify the potential causal effect of peer influence on video game adoption among users of emerging online game platforms. The proposed SEMF achieved an excellent performance in capturing users' preference for games, as indicated by low MPR. Peer influence is then identified by matching users with similar tastes as measured by low-rank features. This study found a positive effect of peer influence. The results also confirm that the magnitude of peer influence would be overestimated if homophily is not considered.

The theoretical contribution of this paper is twofold. First, this article adds to the growing body of research on the video game industry by introducing the framework that identifies peer influence on game adoption behaviors. As opposed to previous studies focusing on factors that influence game sales at an aggregated game level, this study examines consumer purchase behavior at a more fine-grained user level. The results help researchers better understand customers' decision-making processes on online game platforms. Second, this study contributes to the literature on peer influence in product adoption by proposing an innovative yet simple framework that disentangles homophily from social influence between linked users. The proposed method can be applied to different online social network contexts that exhibit user-item interactions. Another merit is the use of observational data in the identification strategy, which does not require experimental manipulation and can be applied to more scenarios where experimental data is difficult to obtain.

The findings in this paper also have practical implications for the game industry. After controlling user similarities based on consumers' tastes, a positive effect of peer influence is recognized in consumers' game adoption. As younger people are the main consumers of online games, most customers respect opinions suggested by their mates, and they are more easily affected by peer behavior (Sheu, Chu, & Wang, 2017). Our results provide empirical evidence to confirm the peer influence on game adoption. Therefore, stakeholders of video games, such as game developers, game publishers, and game distribution platforms, could generate more revenue from games by employing strategies that promote peer influence among friends. For example, they may identify influential users and motivate them to recommend the products or develop product features that can effectively facilitate user interactions and enhance peer influence. In addition, the model presented in this work measures peer influence quantitatively and demonstrates that the magnitude of peer influence would be overestimated if homophily is not considered. This shed light on the importance of knowing users' diversified tastes and creating user segmentation. Game developers and game platforms can collaborate and design games with customized features to target specific user groups, which will strengthen the joint impact of homophily and peer influence.

The potential caveat of this research may be due to the use of a focal game DOTA2 in the analysis of peer influence, which limits the scale of this study to investigate the role of item-related features in the consumer decision-making process. In particular, the purchase price is a non-negligible factor in consumer decision-making because users are more inclined to adopt a game if the cost is low. Therefore, it can be expected that the effect of social influence is heterogeneous with respect to price. To this end, the authors will explore the moderating effect of price on social influence in future studies. Another limitation is the small sample size. Among 108 million users on STEAM, this study randomly selects 50,000 users due to the algorithm's time complexity in solving the optimization problem. The authors will develop a parallel algorithm for large-scale datasets in follow-up studies to fully explore the effect of peer influence.

The current study can be extended in several directions for future research. First, it is worthwhile to explore the moderating effect of user-specific characteristics on peer influence. Because the user-level social network data is available, researchers can further study if users' network properties (e.g., centrality, structural hole) would affect peer influence in the adoption of games. Second, this study only examines the peer influence in adopting the most popular game in 2013. Researchers can include more games to study item-related heterogeneity (e.g., game publisher and game genre) in peer influence.

CONFLICT OF INTEREST

The authors certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

FUNDING AGENCY

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

Aral, S., Muchnik, L., & Sundararajan, A. (2009). Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(51), 21544–21549. doi:10.1073/pnas.0908800106 PMID:20007780

Aral, S., & Walker, D. (2011). Creating social contagion through viral product design: A randomized trial of peer influence in networks. *Management Science*, *57*(9), 1623–1639. doi:10.1287/mnsc.1110.1421

Bapna, R., & Umyarov, A. (2015). Do your online friends make you pay? A randomized field experiment on peer influence in online social networks. *Management Science*, *61*(8), 1902–1920. doi:10.1287/mnsc.2014.2081

Brown, S. A., & Venkatesh, V. (2005). Model of adoption of technology in households: A baseline model test and extension incorporating household life cycle. *Management Information Systems Quarterly*, *29*(3), 399–426. doi:10.2307/25148690

Chen, Y., Wu, C., Xie, M., & Guo, X. (2011). Solving the sparsity problem in recommender systems using association retrieval. *Journal of Computers*, *6*(9), 1896–1902. doi:10.4304/jcp.6.9.1896-1902

Clarke, R. I., Lee, J. H., & Clark, N. (2017). Why video game genres fail: A classificatory analysis. *Games and Culture*, *12*(5), 445–465. doi:10.1177/1555412015591900

Cox, J. (2014). What makes a blockbuster video game? An empirical analysis of US sales data. *Managerial and Decision Economics*, *35*(3), 189–198. doi:10.1002/mde.2608

Davin, J. P., Gupta, S., & Piskorski, M. J. (2014). *Separating homophily and peer influence with latent space* (Working Paper 14-053). Harvard Business School. 10.2139/ssrn.2373273

Entertainment Software Association. (2020). *Essential facts about the video game industry*. https://www.theesa.com/wp-content/uploads/2021/03/Final-Edited-2020-ESA_Essential_facts.pdf

Goldberg, D., Nichols, D., Oki, B., & Terry, D. (1992). Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, *35*(12), 61–70. doi:10.1145/138859.138867

Hartmann, W. R. (2010). Demand estimation with social interactions and the implications for targeted marketing. *Marketing Science*, *29*(4), 585–601. doi:10.1287/mksc.1100.0559

Hill, J., Weiss, C., & Zhai, F. (2011). Challenges with propensity score strategies in a high-dimensional setting and a potential alternative. *Multivariate Behavioral Research*, *46*(3), 477–513. doi:10.1080/00273171.2011.570161 PMID:26735884

Hofacker, C. F., De Ruyter, K., Lurie, N. H., Manchanda, P., & Donaldson, J. (2016). Gamification and mobile marketing effectiveness. *Journal of Interactive Marketing*, *34*, 25–36. doi:10.1016/j.intmar.2016.03.001

Hu, Y., Koren, Y., & Volinsky, C. (2008). Collaborative filtering for implicit feedback datasets. *2008 Eighth IEEE International Conference on Data Mining* (263-272). doi:10.1109/ICDM.2008.22

Kim, M. (2021). Does playing a video game really result in improvements in psychological well-being in the era of COVID-19? *Journal of Retailing and Consumer Services*, *61*, 102577. doi:10.1016/j.jretconser.2021.102577

Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer*, *42*(8), 30–37. doi:10.1109/MC.2009.263

Li, Y. M., Wu, C. T., & Lai, C. Y. (2013). A social recommender mechanism for e-commerce: Combining similarity, trust, and relationship. *Decision Support Systems*, *55*(3), 740–752. doi:10.1016/j.dss.2013.02.009

Little, R. J., & Rubin, D. B. (2014). *Statistical analysis with missing data* (2nd ed.). Wiley., doi:10.1002/9781119013563

Ma, L., Krishnan, R., & Montgomery, A. L. (2014). Latent homophily or social influence? An empirical analysis of purchase within a social network. *Management Science*, *61*(2), 454–473. doi:10.1287/mnsc.2014.1928

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, *60*(3), 531–542. doi:10.2307/2298123

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415–444. doi:10.1146/annurev.soc.27.1.415

Nair, H. S., Manchanda, P., & Bhatia, T. (2010). Asymmetric social interactions in physician prescription behavior: The role of opinion leaders. *JMR, Journal of Marketing Research*, *47*(5), 883–895. doi:10.1509/jmkr.47.5.883

Nitzan, I., & Libai, B. (2011). Social effects on customer retention. *Journal of Marketing*, *75*(6), 24–38. doi:10.1509/jm.10.0209

O'Neill, M., Vaziripour, E., Wu, J., & Zappala, D. (2016). Condensing STEAM: Distilling the diversity of gamer behavior. *Proceedings of the 2016 Internet Measurement Conference*, 81–95. doi:10.1145/2987443.2987489

Potard, C., Henry, A., Boudoukha, A.-H., Courtois, R., Laurent, A., & Lignier, B. (2020). Video game players' personality traits: An exploratory cluster approach to identifying gaming preferences. *Psychology of Popular Media*, *9*(4), 499–512. doi:10.1037/ppm0000245

Quax, P., Beznosyk, A., Vanmontfort, W., Marx, R., & Lamotte, W. (2013), An evaluation of the impact of game genre on user experience in cloud gaming. *2013 IEEE International Games Innovation Conference (IGIC)*, 216–22. doi:10.1109/IGIC.2013.6659141

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. doi:10.1093/biomet/70.1.41

Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Annals of Statistics*, *6*(1), 34–58. doi:10.1214/aos/1176344064

Sheu, J.-J., Chu, K.-T., & Wang, S.-M. (2017). The associate impact of individual internal experiences and reference groups on buying behavior: A case study of animations, comics, and games consumers. *Telematics and Informatics*, *34*(4), 314–325. doi:10.1016/j.tele.2016.08.013

Tucker, C. (2008). Identifying formal and informal influence in technology adoption with network externalities. *Management Science*, *54*(12), 2024–2038. doi:10.1287/mnsc.1080.0897

von der Heiden, J. M., Braun, B., Müller, K. W., & Egloff, B. (2019). The association between video gaming and psychological functioning. *Frontiers in Psychology*, *10*, 1731. doi:10.3389/fpsyg.2019.01731 PMID:31402891

Zhu, F., & Zhang, X. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing*, *74*(2), 133–148. doi:10.1509/jm.74.2.133

Yang Li is an Assistant Professor in the Department of Information Management and Business Analytics at Montclair State University. He received his Ph.D. in Decision Sciences from Drexel University. His main areas of interest are data mining, machine learning, and causal inference in empirical studies. Dr. Li's research focuses on user behavior on social media platforms, online advertising, and informetrics. His teaching interests include data/text mining, statistics, cloud computing, and optimization.

Jiangen He is an Assistant Professor of Information Sciences in the School of Information Sciences at the University of Tennessee, Knoxville. Dr. He's research is in Science of Science, Information Visualization, and Human-AI Interaction. His work focuses on visual analytics, data-driven predictions, and metrics in science of science and studies of how socialness in artificial intelligence affects our information behavior. He received a B.S. in Business Management and a B.S. in Computer Science (Nanjing Tech University, China), an M.S. in Information Science (Wuhan University, China), and a Ph.D. in Information Science (Drexel University, USA).

Chuanren Liu is an Assistant Professor in the Business Analytics and Statistics Department at the University of Tennessee, Knoxville. His research interests include data mining and knowledge discovery, and their applications in business analytics. He has published papers in refereed journals and conference proceedings, such as IEEE Transactions on Knowledge and Data Engineering, INFORMS Journal on Computing, European Journal of Operational Research, Annals of Operations Research, Information Sciences, Knowledge and Information Systems and SIGKDD, ICDM, SDM, AAAI, IJCAI, UbiComp, IEEE BigData, DSAA, etc.

Yanni Ping is Assistant Professor in the Business Analytics and Information System Department at St. John's University, New York, U.S.A. Her main area of research is predictive analytics for E-commerce marketing.