How Editorial, Temporal and Social Biases Affect Online Food Popularity and Appreciation

Markus Rokicki
L3S Research Center
Hannover, Germany
rokicki@L3S.de

Eelco Herder
L3S Research Center
Hannover, Germany
herder@L3S.de

Christoph Trattner
MODUL University Vienna
Vienna, Austria
christoph.trattner@modul.ac.at

Abstract

Measures of popularity and appreciation provide useful information for search and recommendation systems that facilitate access to growing amounts of user-generated content, such as online recipes. However, user rating and commenting behavior is not only influenced by the content itself, but also due to additional effects introduced by biases and contexts. Based on a large dataset of more than 400,000 online recipes, we investigate the nature of such biases and the impact on the number of ratings, comments and views. Our analysis shows that user feedback is significantly influenced by the recipe author's prior reputation, by friendship relations, similarity between user profiles, temporal and seasonal effects, and editorial choices. Furthermore, a regression analysis shows that for the number of ratings received by recipes in particular, an excellent fit can be obtained based on a combination of these biases. These results imply that the popularity of an item is heavily influenced by random bias introduced by various external factors that impact rating and commenting behavior in a relatively short time span after publication.

Introduction

User reputations in online communities are largely based on their contributions to these communities. This is particularly the case when these communities are centered around sharing items or knowledge that are directly of use to the community - such as advice and support on Stack Overflow, hotel reviews on TripAdvisor or recipes on Allrecipes. These reputations are built and measured in different ways, most notably through ratings and comments from other users.

In an ideal world, these ratings and comments objectively reflect the quality of an explanation, a review or a recipe. In reality, several types of *bias* have an impact on the reputation of users' contributions and therewith on the reputations of the users themselves. Particularly peer pressure and *herding* behavior are known sources of bias (Lee, Hosanagar, and Tan 2015). As an example, contributions from users with a good reputation or many friends usually obtain higher ratings and better comments than similar contributions from users who do not (yet) have a good track record.

In this paper, we investigate different types of bias that are expected to be found in food communities. On platforms like Allrecipes[1] or the German community Kochbar[2], users upload and share their recipes with the community - as well as with users who only search for recipes.

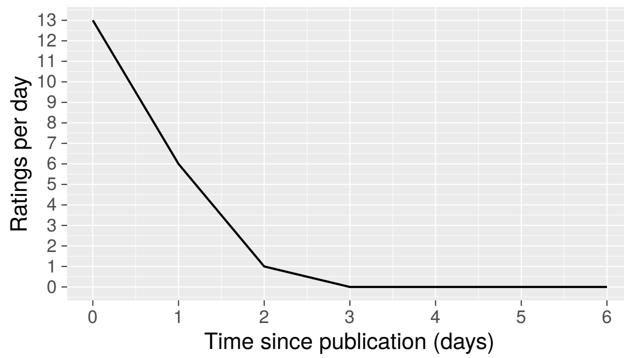


Figure 1: Median ratings per day over time for recipes in Kochbar.de since the publication.

Similar to books or movies, individual taste plays an important role in users' (subjective) perception of the 'quality' of a recipe. For this reason, search and recommendation are often based on *collaborative filtering* (e.g. (Freyne and Berkovsky 2010)): users will be recommended recipes that are liked by users who like similar recipes as themselves.

In contrast to books and movies, and particularly in contrast to news and social media posts, food recipes are expected to be relatively timeless: despite trends and seasonality, recipes of one or two, or even five, years old, are not yet 'stale' or old-fashioned. Nevertheless, we observe that time plays a relatively pronounced role: recipes receive most of the attention on recipe platforms shortly after being published (see Fig. 1). As we will show in this paper, the rating and commenting behavior in this relatively short time span is shaped by different types of biases, which have a direct impact on the longer-lasting popularity of a recipe and its creator.

Friendship is a known bias and its effect is a bit of a chicken-and-egg problem: people befriend people they like and people like their friends better than others. In addition, users' prior *popularity* is expected to have an impact on the

[1] `http://allrecipes.com/`
[2] `http://www.kochbar.de/`

ratings they will receive. Furthermore, *gender effects* play a role as well (Rokicki et al. 2016). *Time* aspects, such as day of week, public holidays and seasonality are known to have an impact on user activity and commenting behavior. Finally, *editorial choices*, such as featuring selected recipes, can boost its reputation as well. All these different choices and decisions are arguably only loosely related to the (perceived) quality of a recipe.

As ratings and comments are used as important selection criteria for search results and recommendations, the implication of these biases is that users may not necessarily be recommended the best, most popular or most relevant recipes. More general, as will be discussed at the end of this paper, it seems that popularity factors in social platforms with peer-reviewing mechanisms are at least to a certain extent random.

Contributions We investigate the influence of different types of biases on ratings, views and comment sentiment on online recipes. We explore the nature of social biases, such as the author's reputation and friendship relations, as well as temporal biases and editorial biases. The results indicate that these external factors are strongest during a short period after publication, with a significant, lasting impact on the popularity of a recipe. These results provide more insights on the usefulness and reliability of measures of popularity and appreciation in online communities that are characterized by users who alternatingly play the role of content creator and consumer.

Related Work

In this section we summarize relevant related work. We organize the content in three main sub-sections. First, we discuss several forms of bias in online ratings: how they are introduced, how they propagate and how they are perceived. Second, we summarize already existing work that has been performed in the context of understanding online food consumption and production behavior. Finally, we review research on recommender systems for online food (the systems we want to enhance given the insights from this paper).

Bias in Online Ratings

It is well-known that the design and functionality of a platform - for example the way ratings and comments are elicited, or the way items are recommended - shape user behavior. In addition, social norms within the platform are shaped by user communities as well as existing norms in a particular area, such as recipe exchange (Olteanu et al. 2016). Several types of bias have been studied in different contexts.

If one would have read a book, watched a movie or stayed in a hotel and would leave a rating or comment in a social platform, this comment or rating is most likely influenced by *prior ratings*: users tend to herd with the crowd when rating popular movies and to differentiate for niche movies; in all cases, they tend to follow their friends' opinions, particularly if the circle of friends is relatively small (Lee, Hosanagar, and Tan 2015). The main reason for this effect is that users do not want to stand out negatively among their peers. Similar results have been obtained in a controlled study in the domain of book ratings (Wang, Zhang, and Hann 2015). The authors also found that peer influence is stronger for older books, which already have an established reputation.

An additional factor that influences the direction in which an item's ratings and reputation is driving as a result of peer pressure, is that negative reviews have more impact than positive reviews (Chen and Lurie 2013). As a result, one negative review might cancel out a far larger number of positive reviews. However, the presence of temporal information (i.e. when the reviews have been given) reduces this effect.

It has been observed that ratings on recipe websites like Kochbar are overwhelmingly positive (Rokicki et al. 2016). The same effect has been observed on the accommodation marketplace Airbnb (Zervas, Proserpio, and Byers 2015), where over 95% of the ratings are 4.5 stars or higher - in contrast to TripAdvisor, where ratings follow a positively skewed normal-like distribution with an average of 3.8 stars. The authors hypothesize that one of the reasons may be that ratings are mutual (guests can rate hosts and hosts can rate guests).

Food communities like Kochbar have in common with Airbnb that ratings are mutual. Whereas in the case of books, movies or hotels, the authors, actors or hotel owners are not part of the community of readers, watchers or guests, in online recipe communities users can play alternatively the role of *creator* and *consumer* and maintain online 'friendships' in both roles. A major difference with Airbnb is that - similar to books and movies - recipes are created and published just once, whereas accommodations offered by Airbnb are constantly 'freshly available'. This leads to the unique situation that recipes in food communities have a short time span during which they receive the majority of ratings, which at the same time are influenced by peer pressure, mutual relationships and other types of bias.

Studies on Online Food Patterns

Studying and understanding online food patterns is a relatively new field of research and only a few studies exist. These have mostly been conducted on a smaller scale, through, e.g., online or telephone surveys (Tilman and Clark 2014). In contrast to these studies, we apply data-mining techniques to an online community platform, in order to study these patterns on a much larger scale.

One significant large-scale research effort in this context was a study conducted by (Ahn et al. 2011), who mined and analyzed three different large-scale online food community platforms from Europe, the US and China to unveil patterns on how recipes are created online in a global sense, how they vary and to find out which flavor components make, for instance, Indian food different from the rest of the world. A second noteworthy large-scale study was conducted by (West, White, and Horvitz 2013), who analyzed a large corpus of Microsoft Bing search logs to discover patterns in how people search and access information in the online food community website Allrecipes.com. The results indicate that queries for food on the Web follow weekly and yearly trends.

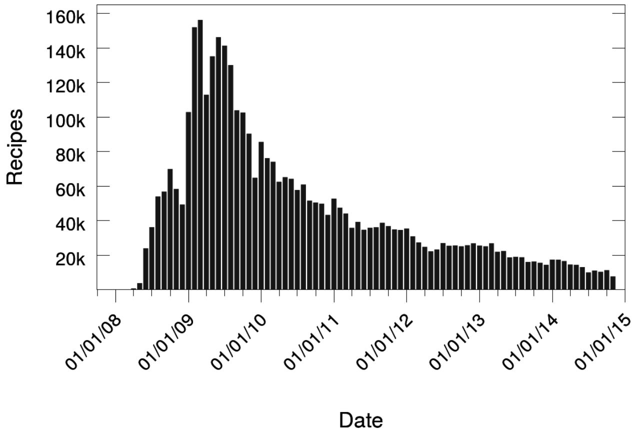


Figure 2: Overall number of recipes published per month over time.

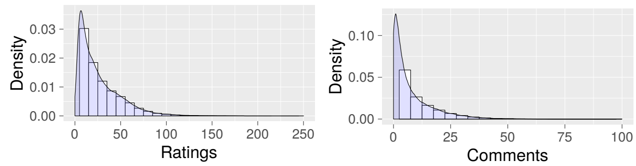
(Said and Bellogín 2014) analyzed a crawl of the online food community platform Allrecipes.com, showing significant correlations between diabetes in certain regions of the US and the types of recipes that users from these regions rated highest. (Abbar, Mejova, and Weber 2015) found a close relation between dietary choices and trends and preferences on Twitter; (Mejova et al. 2015) found similar patterns in Instagram, while (De Choudhury, Sharma, and Kiciman 2016) showed that Instagram data and bad food consumption patterns can be correlated to unhealthy regions in the US. These results may suggest that patterns found in online are representative for actual dietary choices and preferences.

(Wagner, Singer, and Strohmaier 2014) analyzed user preferences in the German online food community Kochbar by analyzing access log data for the recipes available in this platform. In line with West et al., they found that online food preferences follow temporal trends and vary in certain regions in Germany. In an earlier study, we observed and analyzed various temporal patterns in the Kochbar dataset (Kusmierczyk, Trattner, and Nørvåg 2015b; 2015a), the same dataset we use in this paper. In a recent paper, we investigated the impact of social influence and temporal patterns on food recommendation (Kusmierczyk, Trattner, and Nørvåg 2016). The work presented in this paper builds upon this research by investigating the nature of these patterns and how they introduce bias in user ratings and recipe popularity.

Studies on Online Food Recommender Systems

Although food recommendation has been infrequently studied, there is a small body of appropriate related work. Early attempts to design automated systems using case-based planning to recommend meals include CHEF (Hammond 1986) and JULIA (Hinrichs 1989). Other approaches include hybrid recommenders (Sobecki, Babiak, and Słanina 2006) and recommendations based on grouping of users (Svensson et al. 2000). More recent efforts try to understand a user's tastes, improving recommendations by breaking recipes down into individual ingredients (Freyne, Berkovsky, and Smith 2011; Freyne and Berkovsky 2010). (Teng, Lin, and Adamic 2012) make use of complementary and substitution networks and show which ingredients users add, remove, pair or substitute. This allows them to predict which variation of a recipe will receive the best ratings. More recent work in this context includes the studies of (Elsweiler and Harvey 2015; Elsweiler et al. 2015), (Ge et al. 2015), (Trattner and Elsweiler 2017), (Bianchini et al. 2016) and (Yang et al. 2016), proposing new methods for healthy online food recommendations.

Chef Watson, created by the team of IBM Watson, is to date the only "recommender" system that actually helps the user in creating recipes. Chef Watson uses machine learning techniques for automatically creating recipes that match user preferences, based on existing recipes from the Bon Appetit recipe website (available to the public at `https://www.ibmchefwatson.com/`).



(a) Ratings per recipe (b) Comments per recipe

Figure 3: Distributions of ratings and comments per recipe.

Dataset

For the purpose of our study, we rely on a large-scale crawl from Kochbar.de[3], a German online food community website to which users can upload and rate cooking recipes, obtained in (Kusmierczyk, Trattner, and Nørvåg 2015a). The dataset encompasses more than 400 thousand recipes published between 2008 and 2014. Figure 2 shows the long term development of the platform in terms of the number of recipes published per month. After an initial increase in recipe uploads, the platform entered a state of steady decline in activity. To minimize the impact of these long-term developments on our study, we restrict our analysis to recipes that have been uploaded in 2010 or later (195 thousand recipes remain). At this time, the number of known ingredients, as well as recipe innovation in terms of ingredient combinations were saturated (Kusmierczyk, Trattner, and Nørvåg 2015a).

Overall, almost 200 thousand users provided 2.7 million comments and 7.7 million ratings. The ratings are on a Likert scale, but overwhelmingly positive (99.1% gave a rating of 5). As can be seen in Figure 3, the distributions of the number of ratings per recipe and the number of comments per recipe are long-tailed. Gender and age information was given by 95 thousand and 57 thousand users, respectively. More than 18 thousand users have also actively contributed recipes to the platform. Furthermore, friendship relations exist between more than 15 thousand users, with 100 thousand friendship relationships in total.

[3] `https://www.kochbar.de`

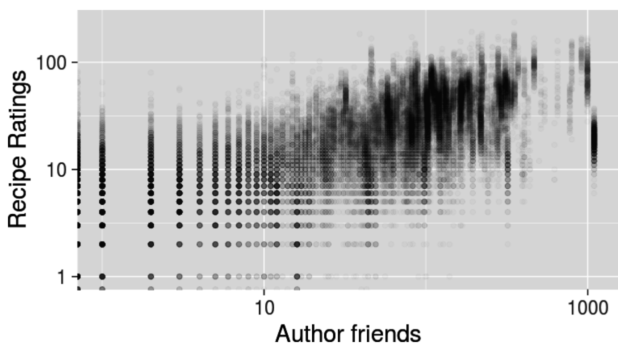


Figure 4: Scatter plot of the number of friends of recipe authors in relation to the number of ratings received by the recipes.

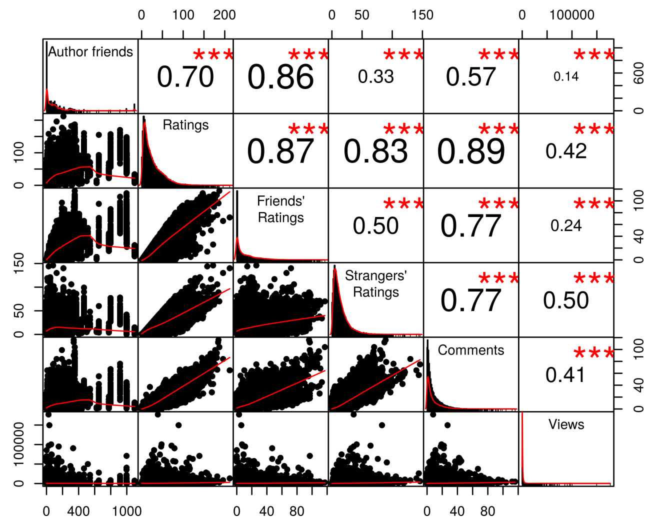


Figure 5: Spearman's rank correlations for recipe popularity in terms of number of ratings, number of comments, number of views and author popularity in terms of number of friends (*p < 0.05, **p < 0.01, ***p < 0.001).

Empirical Data Analysis

In this section we show to what extent the popularity of recipes – in terms of ratings, comments and views – depends on factors that are independent from the recipe content itself. We are interested in how well recipes are received on the platform in terms of popularity (number of ratings, comments, and views) as well as appreciation (comment sentiment). Whereas ratings are heavily skewed towards 5-star ratings, comment sentiment showed to be normally distributed on a scale from -4 (very negative sentiment) to +4 (very positive sentiment)[4].

Social Biases

Popular Recipe Authors Figure 4 shows the number of ratings received by recipes in relation to the number of friends of recipe authors. The plot reveals a likely correlation between author popularity in terms of number of friends and recipe popularity. This would be in line with our expectations, given the fact that a considerable fraction of ratings – especially for popular recipes – is given by friends of the recipe authors.

We further investigate this in Figure 5, which shows Spearman's rank correlations between recipe popularity (in terms of number of ratings, number of comments, and number of views) and author popularity (in terms of number of friends). We can observe a high correlation between author popularity and recipe popularity in terms of numbers of ratings ($\rho = 0.70$, $p < .001$) and comments ($\rho = 0.57$, $p < .001$). In contrast, the correlation between author popularity and recipe views – which, in contrast to ratings and comments are not restricted to registered platform users – is quite low (although still significant), with only a value of $\rho = 0.14$, $p < .001$. The correlation values between the number of views and the number of ratings ($\rho = 0.42$, $p < .001$) and comments ($\rho = 0.41$, $p < .001$) are relatively low as well.

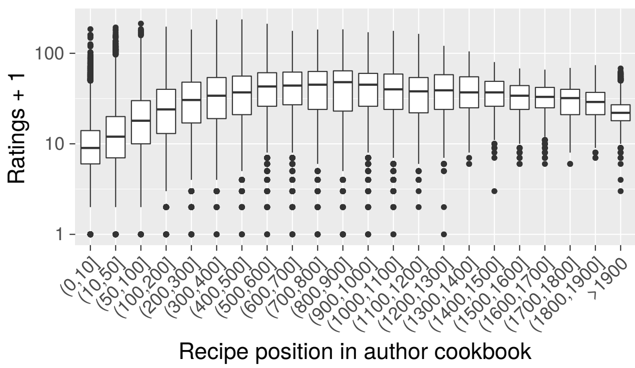
The effect of an author's established reputation on the number of ratings can be shown in Figure 6(a): the average number of ratings for a recipe grows with the number of previously published recipes, an effect that stabilizes after about 300 recipes – which is apparently the point that the recipe author has completed the reputation building phase. By contrast, the number of views remains quite stable 6(b). In both figures, the average number of ratings or views starts to decrease after about 1500 recipes – it is beyond the scope of this paper to speculate about the reasons.

The above results indicate that author popularity and friendship do result in a bias regarding ratings and comments, but not so much regarding the (more or less) independent and arguably more objective number of views.

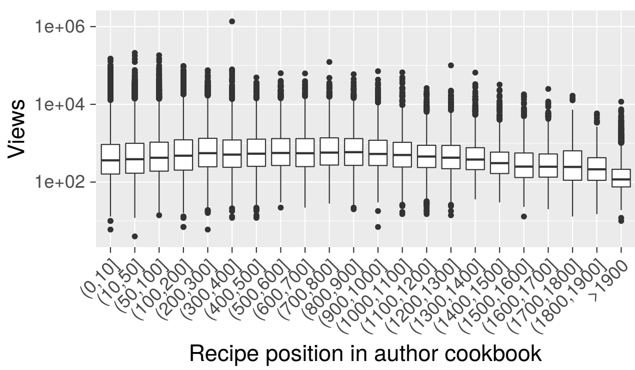
Friendship Overall, 49.9% of all comments and 45.5% of ratings are given to recipes from friends. Figure 7 shows the fractions of ratings received from friends of the recipe authors in relation to the overall number of ratings received. We can observe that, interestingly, the fraction of ratings/comments received from friends is higher for more popular recipes, starting with recipes with approximately more than 50 ratings and recipes with more than 20 comments. We bisect the recipe data and find that recipes with more than 50 ratings receive a significantly higher fraction of ratings by friends ($M = 0.54$) than recipes with less than 50 but more than 10 ratings ($M = 0.39$), $W = 2.147 \cdot 10^9$, $p < .001$, $r = .316$. This implies that the Kochbar community is dominated by a dense clique of users with close connections.

Comments on friends' recipes are longer ($M = 79.5$ characters) compared to comments on strangers' recipes ($M = 77.3$), $W = 2.603 \cdot 10^{11}$, $p < .001$, $r = .041$. However, if we distinguish commentator gender, a slightly different picture emerges. The difference between male comments on friends' recipes ($M = 75.4$) and male comments on strangers' recipes ($M = 73.3$) is not significant, $W =$

[4]Sentiment was computed using the German version of SentiStrength (`http://sentistrength.wlv.ac.uk`).



(a) Ratings



(b) Views

Figure 6: Ratings and views of recipes in relation to their position within the recipe author's cookbook.

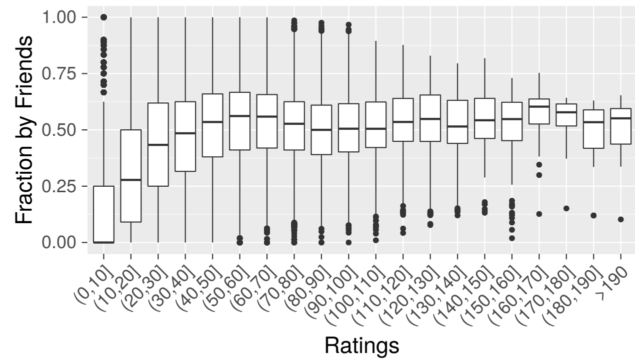


Figure 7: Fractions of ratings received from friends of recipe authors in relation to the popularity of recipes terms of ratings.

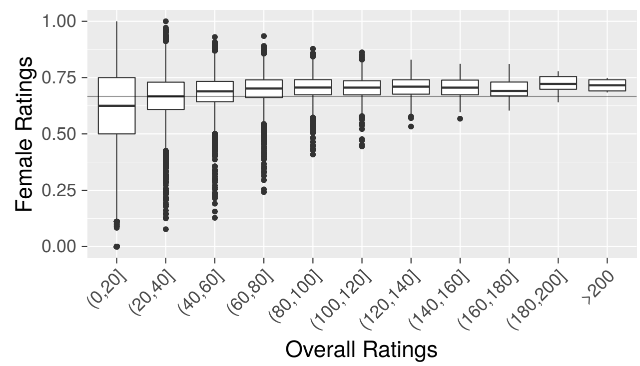


Figure 8: Fraction of ratings received from female users in relation to recipe popularity in terms of the overall number of ratings received. The horizontal line is the overall median.

$1.249 \cdot 10^{10}$, $p = 0.95$, while female comments on friends' recipes ($M = 80.7$) are significantly longer than female comments on strangers' recipes ($M = 76.7$), $W = 1.454 \cdot 10^{11}$, $p < .001$, $r = .058$.

In terms of sentiment, the results are similar. Overall, comments on friends' recipes have a more positive sentiment ($M = .272$) in comparison to comments on strangers' recipes ($M = .240$), $W = 2.556 \cdot 10^{11}$, $p < .001$, $r = .022$. This effect is significant both for male commentators ($M = .214$ vs $M = .147$, $W = 1.300 \cdot 10^{10}$, $p < .001$, $r = .041$) and for female commentators ($M = .297$ vs $M = .275$, $W = 1.400 \cdot 10^{11}$, $p < .001$, $r = 0.017$).

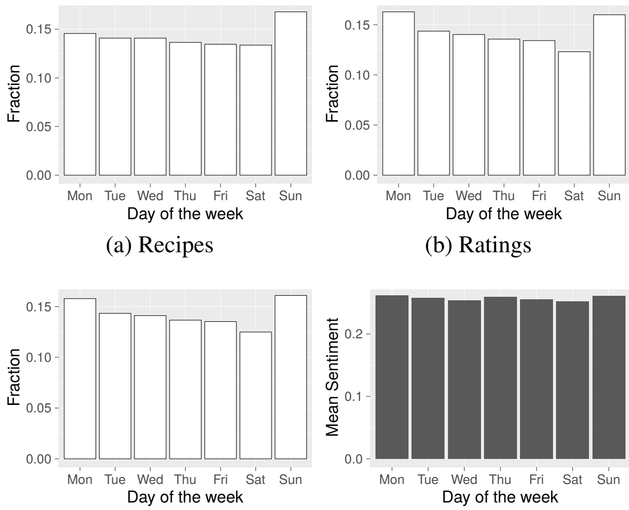
To summarize, recipes receive about as many ratings from 'friends' as from 'strangers'. Comments on recipes from friends are longer than on other recipes, particularly in the case of female users, and the sentiment is more positive.

Gender Bias

The results in the previous section highlight that gender influences user feedback on recipes and this warrants further study. We start by analyzing rating behavior. We expect a preference for rating recipes of users of the same gender and measure the fraction of ratings received from female users for recipes of male and of female recipe authors. However, contrary to this intuition, we find that male recipes receive a higher fraction of ratings from female users (M = 0.68) compared to female recipes (M = 0.67) instead, $W = 2.423 \cdot 10^9$, $p < .001$, $r = .098$. Figure 8 contrasts the fractions of ratings received from female users in relation with the total number of ratings received. The plot reveals a slight upwards trend for the fraction of female ratings with increasing recipe popularity.

Another aspect of user feedback is the nature of the comments on recipes. Comments from female users ($M = 78.9$ characters) are slightly, but significantly longer than comments from male users ($M = 74.3$), $W = 1.634 \cdot 10^{11}$, $p < .001$, $r = .024$. Female commentators write longer comments ($M = 81.2$) on male recipes than on female recipes ($M = 77.6$), $W = 8.448 \cdot 10^{10}$, $p < .001$, $r = .034$. Similarly, male commentators also write longer comments ($M = 76.5$) on male recipes than on female recipes ($M = 73.3$), $W = 7.541 \cdot 10^9$, $p < .001$, $r = .0304$.

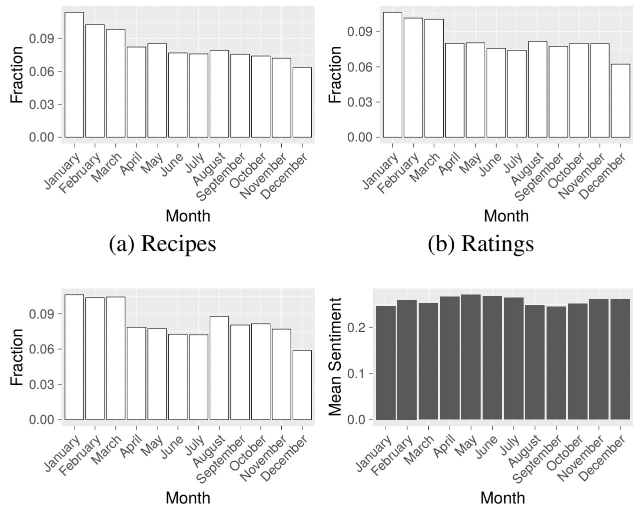
Sentiment in comments on recipes of authors with the same gender ($M = .285$) is slightly, but significantly higher than sentiment in comments on recipes of authors of the opposite gender ($M = .225$), $W = 2.003 \cdot 10^{11}$, $p < .001$, $r = .049$. However, as female users are more positive in their comments overall ($M = .287$) compared to male users ($M = .184$), $W = 1.530 \cdot 10^{11}$, $p < .001$, $r = .086$, we suspect that the observation for same-gender comments



(a) Recipes (b) Ratings

(c) Comments (d) Comment Sentiment

Figure 9: Recipes, ratings, comments, and comment sentiment by day of the week.



(a) Recipes (b) Ratings

(c) Comments (d) Comment Sentiment

Figure 10: Recipes, ratings, comments, and comment sentiment by month.

stem from the overall gender distribution. Indeed, female users are more positive in their comments on female recipes ($M = .293$) compared to male recipes ($M = .269$), $W = 8.035 \cdot 10^{10}$, $p < .001$, $r = 0.017$. As male users are also more positive in their comments on female recipes ($M = .187$) compared to comments on male recipes ($M = .175$), $W = 7.266 \cdot 10^9$, $p < .001$, $r = .007$, we can conclude – in line with (Rokicki et al. 2016) – that recipes from females receive more positive comments in general.

In summary, we can conclude that even though there are differences in commenting behavior between males and females, the tendencies are the same: recipes from women receive shorter, more positive comments.

Temporal Biases

User behavior also varies over time. Figure 9 shows the distributions of recipes, ratings, comments as well as the average comment sentiment on a day of the week level. The distribution of recipe uploads shown in Figure 9a is relatively constant across the week, with a single peak on Sunday (when 16.4% of all recipes are uploaded). Similarly, users are more active in terms of ratings and comments on Sunday as well. However, across the week we rather observe a continuous downwards trend between Sunday and Saturday. Comment sentiment varies slightly but significantly across the days of the week, $F(6, 1420215) = 4.48$, $p < .001$, $\eta_p^2 = 1.89 \cdot 10^{-5}$, with users being slightly more positive on Sundays and Mondays.

Figure 10 shows the number of recipes, ratings, comments and the sentiment per months. As the distribution of recipes over months seems to be influenced by the development of the platform (see the dataset description), we only concentrate on the sentiment. Comment sentiment varies significantly across months of the year, $F(11, 1420210) = 16.05$, $p < .001$, $\eta_p^2 = .00012$, and is most positive during the early Summer season as well as during the Christmas season.

In summary, appreciation of recipes slightly varies in seasonal, as well as weekly patterns.

Editorial Biases

On the platform, there are multiple mechanisms through which editorial biases may be introduced. Among others, the front page of the platform *features* "recommended recipes" and "recipes of the week", selected by editorial staff. A recipe of the week is selected every week on Mondays, starting in May 2008 and spanning the whole duration of our dataset. For recommended recipes, on the other hand, the time of being featured is not available – however, we expect that a number of recipes has not been picked up by the editorial staff immediately upon being published, as is the case for recipes of the week. We confirm this for the recipes of the week in Figure 11, which shows the delay in days between publishing of recipes and being featured as recipe of the week. The figure shows that, although a large number of recipes becomes recipe of the week within a week of being published (38.3%), the majority is picked up only later – after a median delay of 13 days.

Figure 12 shows a boxplot of ratings per week for recipes of the week, relative to the date of being featured. Most ratings are received within a week of being featured, however, there is also a large number of outliers, with high numbers of ratings even months earlier or later – as described in the previous paragraph, the majority of recipes is not featured within one week of being published and we expect initial spikes in the number of ratings recipes receive directly after they are uploaded.

To differentiate between these two effects, we compare recipes that were featured within 3 days of being uploaded to the platform to recipes that were not featured. Figure 13

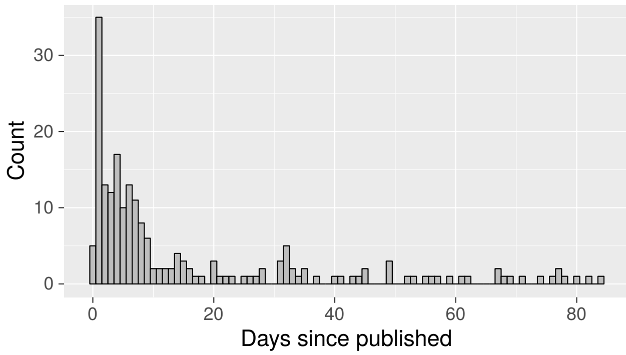


Figure 11: Recipe age when featured as recipe of the week.

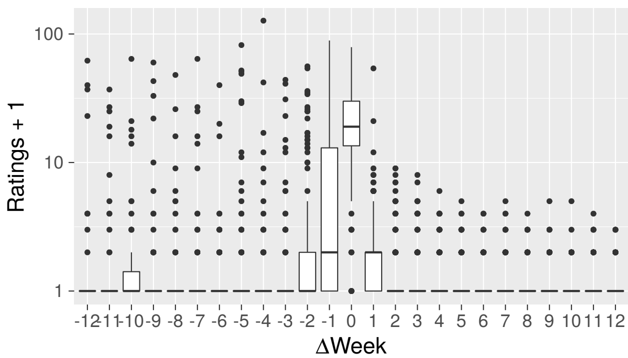


Figure 12: Ratings per week relative the date of being featured as recipe of the week.

shows mean and median ratings per day starting with the date of being published over time for both groups of recipes (We only consider recipes with at least 10 ratings in total). The plot shows that the initial number of ratings is slightly lower for recipes of the week (Median 11 vs Median 13). However, the ratings decay slower compared to recipes that where not featured, with a median of two ratings per day even after a week of being featured – in contrast to recipes that were not featured, where median number of ratings reaches zero only three days after being published.

In short, by selecting recipes to be featured, the platform introduced significant editorial bias, resulting to a boost in ratings, directly after the moment of being featured. Unfortunately, we cannot analyse in retrospective whether this leads to major differences in the long term, as the featured recipes are arguably all recipes that would have attracted more ratings and comments anyway.

Regression Analysis

Finally, Table 1 shows the results of three regression models using the mean comment sentiment, number of ratings and views as dependent variable and editorial, temporal and social feature sets as independent variables. We omit recipes with NULL values, resulting in a dataset of 128 thousand recipes (including NA rows, we have 194 thousand recipes). All models were optimized using a stepwise search procedure, using the R software package and the best models and

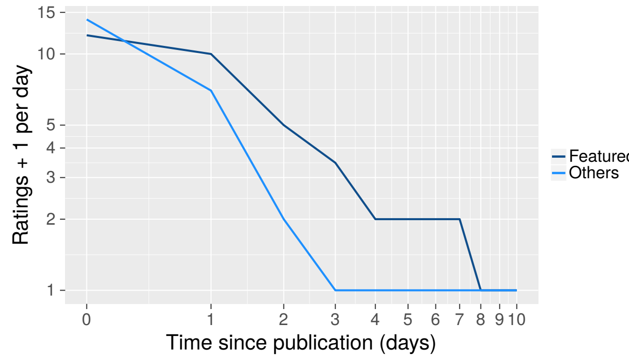


Figure 13: Median number of ratings per day for recipes that were featured as "recipe of the week" within 3 days of being uploaded in comparison recipes that were not featured in any way. Only recipes that received at least 10 ratings in total are considered.

according features are presented. As shown, most of the investigated features are significant.

In the Poisson models for the number of ratings and views, editorial biases uniformly have significant positive correlations, although one of the features is eliminated in each model due to redundancy. All temporal, author, social and gender features are significant as well. Features for author popularity, 'Number of friends' and 'Guest book entries', show positive correlations with recipe popularity, while author activity, in terms of 'Recipes published' in particular, has a negative correlation. 'Recipe position in cookbook' on the other hand, might serve as a proxy for the author's stage of reputation building at the time of publishing the recipe and has a positive correlation for the number of ratings but not for the number of views. Although 'Author gender male' has a positive correlation for ratings and views, receiving high shares of feedback from male users is negatively correlated. The same is the case for 'Comments by friends', suggesting that feedback from (impartial) strangers is a better predictor of overall popularity.

Similar patterns are observed when investigating the OLS model predicting the mean comment sentiment of a recipe. All feature sets investigated show a positive correlation, except the gender feature set and two features 'Guest book entries' and 'Recipes published' in the author feature set – which suggests that recipes geared toward males have a slight disadvantage in Kochbar and that being overly productive does not necessarily lead to more appreciation.

Conclusions and Discussion

In this paper, we have identified and described several types of bias that have a direct impact on the number of ratings and comments on recipes in the German food community Kochbar. Among others, the prior reputation and the number of friends is highly correlated with the ratings and comments, but not with the – arguably more objective – number of views. Further, friendship relations lead to more ratings and more positive comments. Also, a slight influence of temporal and seasonal features can be observed. Most im-

	Dependent variable:				
	Mean comment sentiment *OLS*	**Number of ratings** *Poisson*		**Number of views** *Poisson*	
	Coefficients (β)	Coefficients (β)	exp (β)	Coefficients (β)	exp (β)
Constant	0.026	1.276***	3.581***	3.959***	52.392***
Popularity and appreciation features					
Number of ratings (log)	0.015***				
Number of comments (log)	0.007***				
Comments #characters	0.003***	0.002***	1.002***	−0.005***	0.995***
Comments #words	−0.026***	−0.010***	0.990***	0.043***	1.044***
Comments sentences	0.027***			0.109***	1.115***
Comments sentiment		0.051***	1.052***	0.115***	1.122***
Comments sentimentality		0.092***	1.097***	0.245***	1.278***
Temporal features					
Recipe age (log)	0.029***	0.069***	1.071***	0.370***	1.448***
Weekend	0.004*	−0.014***	0.986***	−0.005***	0.995***
Spring	0.016***	−0.022***	0.978***	0.010***	1.010***
Summer	0.015***	−0.014***	0.986***	0.030***	1.030***
Autumn	0.008**	0.055***	1.057***	0.092***	1.096***
Editorial features					
Featured recipe	0.021**			1.235***	3.437***
Recommended recipe		0.379***	1.460***		
Recipe of the week		0.669***	1.953***	−0.070***	0.932***
Author features					
Number of friends (log)	0.005***	0.089***	1.093***	0.165***	1.180***
Guest book entries (log)	−0.004***	0.258***	1.294***	0.085***	1.089***
Recipes published (log)	−0.005**	−0.318***	0.728***	−0.182***	0.834***
Comments written (log)	0.001	0.052***	1.053***	−0.017***	0.983***
Recipe position in cookbook (log)	0.003*	0.147***	1.158***	−0.024***	0.976***
Social features					
Comments by friends	0.029***	−0.126***	0.881***	−0.615***	0.541***
Same age ratings		−0.271***	0.763***	−0.375***	0.687***
Same gender ratings	0.013	0.405***	1.500***	0.978***	2.660***
Gender features					
Author gender male	−0.013***	0.104***	1.109***	0.301***	1.351***
Male comments	−0.103***	0.029***	1.030***	0.142***	1.153***
Male ratings	−0.058***	−0.704***	0.495***	−0.845***	0.430***
Observations	128,893	128,893		128,893	
Adjusted R^2	0.026				
Adjusted McFadden R^2		0.503		0.114	
Log Likelihood	−56994.93	−677,376		−150,379,935	
Akaike Inf. Crit.	114036	1,354,797		300,759,917	
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 1: Recipe level regression of average comment sentiment, number of ratings and number of views.

portantly, editorial bias is – most likely deliberately – introduced by featuring recipes. As a result, as we have shown by means of a regression analysis, the popularity of a recipe can be successfully predicted by means of these external features. These results clearly show a very direct impact of external bias on the number of ratings and comments.

An important implication of these findings is that the popularity of an item is heavily influenced by bias introduced by various external factors that influence rating and commenting behavior in a relatively short time span after publication. As a result, information retrieval systems that sort items based on ratings and comments may not always recommend the best, most popular or most relevant ones. The number of views is probably a more objective measure. In a similar vein, in line with (Olteanu et al. 2016), the results imply as well that one should be careful in generalizing results that have been drawn from observed rating and commenting patterns.

One shortcoming of our paper is that our analysis was restricted a single – although large – online food community. As such, the specific external factors that we investigated are partially specific to online recipe platforms or even to the specific platform that we observed. However, the observation that ratings in platforms where users function both as creators and consumers (and therefore provide reviews as well as are being reviewed) are extremely positively skewed has been observed in other platforms as well, such as Airbnb (Zervas, Proserpio, and Byers 2015). The same yields for the effects of friendship and author popularity. In the future, we plan to replicate our study and confirm these observations for other online food sharing platforms as well.

Acknowledgements

This work was partially funded by the German Federal Ministry of Education and Research (BMBF) under project GlycoRec (16SV7172).

References

Abbar, S.; Mejova, Y.; and Weber, I. 2015. You tweet what you eat: Studying food consumption through twitter. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3197–3206. ACM.

Ahn, Y.-Y.; Ahnert, S. E.; Bagrow, J. P.; and Barabási, A.-L. 2011. Flavor network and the principles of food pairing. *Scientific reports* 1.

Bianchini, D.; De Antonellis, V.; De Franceschi, N.; and Melchiori, M. 2016. Prefer: a prescription-based food recommender system. *Computer Standards & Interfaces*.

Chen, Z., and Lurie, N. H. 2013. Temporal contiguity and negativity bias in the impact of online word of mouth. *Journal of Marketing Research* 50(4):463–476.

De Choudhury, M.; Sharma, S.; and Kiciman, E. 2016. Characterizing dietary choices, nutrition, and language in food deserts via social media. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, 1157–1170. ACM.

Elsweiler, D., and Harvey, M. 2015. Towards automatic meal plan recommendations for balanced nutrition. In *Proceedings of the 9th ACM Conference on Recommender Systems*, 313–316. ACM.

Elsweiler, D.; Harvey, M.; Ludwig, B.; and Said, A. 2015. Bringing the "healthy" into food recommenders. In *Proceedings of the 2nd International Workshop on Decision Making and Recommender Systems*, 33–36. CEUR-WS.org.

Freyne, J., and Berkovsky, S. 2010. Intelligent food planning: personalized recipe recommendation. In *Proceedings of the 15th international conference on Intelligent user interfaces*, 321–324. ACM.

Freyne, J.; Berkovsky, S.; and Smith, G. 2011. Recipe recommendation: accuracy and reasoning. In *International Conference on User Modeling, Adaptation, and Personalization*, 99–110. Springer.

Ge, M.; Elahi, M.; Fernaández-Tobías, I.; Ricci, F.; and Massimo, D. 2015. Using tags and latent factors in a food recommender system. In *Proceedings of the 5th International Conference on Digital Health 2015*, 105–112. ACM.

Hammond, K. J. 1986. Chef: A model of case-based planning. In *AAAI*, 267–271.

Hinrichs, T. R. 1989. Strategies for adaptation and recovery in a design problem solver. In *Proc. of the 2 nd Workshop on Case-Based Reasoning*, 115–118.

Kusmierczyk, T.; Trattner, C.; and Nørvåg, K. 2015a. Temporal patterns in online food innovation. In *Proc. of WWW'15 Companion*.

Kusmierczyk, T.; Trattner, C.; and Nørvåg, K. 2015b. Temporality in online food recipe consumption and production. In *Proc. of WWW'15*.

Kusmierczyk, T.; Trattner, C.; and Nørvåg, K. 2016. Understanding and predicting online food recipe production patterns. In *Proceedings of the 27th ACM Conference on Hypertext and Social Media*, HT '16, 243–248. New York, NY, USA: ACM.

Lee, Y.-J.; Hosanagar, K.; and Tan, Y. 2015. Do i follow my friends or the crowd? information cascades in online movie ratings. *Management Science* 61(9):2241–2258.

Mejova, Y.; Haddadi, H.; Noulas, A.; and Weber, I. 2015. # foodporn: Obesity patterns in culinary interactions. In *Proceedings of the 5th International Conference on Digital Health 2015*, 51–58. ACM.

Olteanu, A.; Castillo, C.; Diaz, F.; and Kiciman, E. 2016. Social data: Biases, methodological pitfalls, and ethical boundaries.

Rokicki, M.; Herder, E.; Kusmierczyk, T.; and Trattner, C. 2016. Plate and prejudice: Gender differences in online cooking. In Vassileva, J.; Blustein, J.; Aroyo, L.; and D'Mello, S. K., eds., *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization, UMAP 2016, Halifax, NS, Canada, July 13 - 17, 2016*, 207–215. ACM.

Said, A., and Bellogín, A. 2014. You are what you eat! tracking health through recipe interactions. In *RSWeb@ RecSys*.

Sobecki, J.; Babiak, E.; and Słanina, M. 2006. Application of hybrid recommendation in web-based cooking assistant. In *Knowledge-Based Intelligent Information and Engineering Systems*, 797–804. Springer.

Svensson, M.; Laaksolahti, J.; Höök, K.; and Waern, A. 2000. A recipe based on-line food store. In *Proceedings of the 5th international conference on Intelligent user interfaces*, 260–263. ACM.

Teng, C.-Y.; Lin, Y.-R.; and Adamic, L. A. 2012. Recipe recommendation using ingredient networks. In *Proceedings of the 4th Annual ACM Web Science Conference*, 298–307. ACM.

Tilman, D., and Clark, M. 2014. Global diets link environmental sustainability and human health. *Nature* 515(7528):518–522.

Trattner, C., and Elsweiler, D. 2017. Investigating the healthiness of internet-sourced recipes: Implications for meal planning and recommender systems. In *Proc. of WWW'17*.

Wagner, C.; Singer, P.; and Strohmaier, M. 2014. The nature and evolution of online food preferences. *EPJ Data Science* 3(1):1–22.

Wang, A.; Zhang, M.; and Hann, I.-H. 2015. Socially nudged: A quasi-experimental study of friends' social influence in online product ratings. *Information Systems Research, forthcoming*.

West, R.; White, R. W.; and Horvitz, E. 2013. From cookies to cooks: Insights on dietary patterns via analysis of web usage logs. In *Proceedings of the 22nd international conference on World Wide Web*, 1399–1410. ACM.

Yang, L.; Hsieh, C.-K.; Yang, H.; Dell, N.; Belongie, S.; and Estrin, D. 2016. Yum-me: Personalized healthy meal recommender system. *arXiv preprint arXiv:1605.07722*.

Zervas, G.; Proserpio, D.; and Byers, J. 2015. A first look at online reputation on airbnb, where every stay is above average. *Where Every Stay is Above Average (January 28, 2015)*.