Birds of a Feather Lodge Together?: Predicting Review Sentiment Using Social Categorization Theory

Completed Research Full Paper

Lauren Rhue
Wake Forest School of Business
rhuela@wfu.edu

Sherae Daniel
University of Cincinnati
sherae.daniel@uc.edu

Abstract

Reviews drive online sales. In a perfect world, reviews represent a way to make the market contain less friction. That is, they accurately estimate the value of products without bias from the person writing the review. However, some suggest that factors that do not represent the quality of the product alter the relationship between the review and the sale. We use Tajfel and Turner (1986)'s social categorization theory to develop a model that predicts the sentiment of the review left by a guest. We test our hypotheses using archival data about interactions on AirBnB. We find that when the host and the guest share the same ethnicity with positive stereotypes, the review is more positive. However, if the host and the guest share an ethnicity with negative stereotypes, the review is more negative. The discussion section reflects on these results with respect to theory and offers practical implications.

Keywords

Sharing economy, diversity, econometrics, similarity, social categorization

Introduction

A positive review on AirBnB holds great value for the host; it can lead to increased rentals and allow the host to charge a higher rate per stay. AirBnB is a platform that allows those willing to lend out their housing for a short term to a guest the opportunity to do so. The platform offers many benefits to guests and hosts. It allows guests the chance to save money while travelling, and it offers hosts an alternate opportunity to generate revenue.

While AirBnB offers many benefits, some suggest a dark side to the popular platform. Instead of bridging gaps between people, the platform may amplify the influence of human stereotypes and bias. In particular, instead of the review reflecting the quality of the AirBnB, it may reflect a bias based on the degree to which the guest is similar to the host. The in-group effects can hold huge financial impacts for future hosts. In this research, we explore the question "How do societal stereotypes impact the reviews for AirBnB guest-host dyads?"

To address this question we leverage the social categorization theory (Tajfel & Turner, 1979). We then develop and test two hypotheses with an archival dataset from AirBnB that includes 382,206 reviews. We find a nuanced story related to the impact of similarity between the host and guest on reviews. Our analysis reveals evidence that guest/host similarity is associated with a more positive review sentiment for those with European names. For those with names more commonly associated with Hispanic people, similarity leads to lower review sentiment. Our results hold implications for social categorization theory. It also offers some guidance to online platforms such as AirBnB. Finally, our results may enable guests and hosts to get more value from the platform.

In the next section we overview the context and give a review of the literature around temporary lodging online marketplaces, focusing on reviews. Next, we introduce the social categorization theory and offer hypotheses. Finally, we describe the empirical research and offer a discussion.

Background

Reviews in online marketplaces drive sales (Chevalier and Mayzlin, 2006) and provide economic value. These reviews hold even more significance when the products are lodging facilities that the consumer may not see before committing to use. AirBnB reviews lean towards the positive because those with negative experiences tend to skip writing reviews (Fradkin, Grewal, Holtz, & Pearson, 2015). While many of the reviews suggest positive experiences, a micro-analytic comparison of positive reviews reveals that guests use more nuanced or subtle cues to communicate less–than-positive experiences. Given that novices write reviews, opportunity exists for reviews to include bias.

Racial bias influences important outcomes for guests on sharing platforms. One signal that a guest can send to reveal their race is their name. Guests with distinctly African American names are 16 percent less likely to be accepted relative to identical guests with distinctly white sounding names (Edelman, Luca, & Svirsky, 2017). Requests from guests with African American-sounding names are 19.2 percentage points less likely to be accepted than those with white-sounding names. However, a positive review posted on a guest's page significantly reduces discrimination: when guest accounts receive a positive review, the acceptance rates of guest accounts with white-sounding and African American-sounding names are statistically indistinguishable (Cui, Li, & Zhang, 2016).

Racial bias also influences important outcomes on sharing platforms for hosts. On average, Asian and Hispanic hosts charge 8%-10% lower prices relative to their White counterparts on equivalent rental properties, after controlling for all guest-available information on rental unit characteristics, as well as additional information on neighborhood property values, area demographics, and occupancy rates (Kakar, Voelz, Wu, & Franco, 2018).

Prior work considers the influence of the guest's race or the host's race separately. Yet we believe some impacts relate to considering the guest and the host's race simultaneously. We look to social categorization theory to help us understand this simultaneous influence.

Social Categorization Theory

A fruitful starting point for understanding the factors that lead to more positive sentiment in reviews is the research that explores how other dyads evaluate each other. Theorists who study this have found social categorization theory to yield insight. A social identity is a categorization of the self into more inclusive social units (Tajfel & Turner, 1986). Individuals can categorize themselves according to gender, race, values, social roles, and etcetera. The extent to which an individual identifies with a group can influence their thoughts, feelings and behavior (Reed II, Forehand, Puntoni, & Warlop, 2012). We use social category theory because it describes a theory about how people view each other based on how they identify.

Tajfel and Turner (1986) suggest that individuals attribute positive characteristics to individuals they believe are in their same group because it helps them solidify a positive self-concept about themselves. People seek to develop, achieve, protect and maintain a positive self-concept. When they attribute positive attributes to those in their same category, they are helping to re-inforce their own self-concept. This can show up as a bias in favor of their in-group and against the out-group. This can result in behaviors, attributes, memories (Sahdra & Ross, 2007). For example, Sahdra and Ross show that dyads that identify highly with each other recall fewer incidents of in-group violence and hatred than did low identifiers. That is, people evaluate the groups to which they belong more favorably than groups to which they do not belong.

We contribute to social category theory by testing its presence in using real world data. Much of the examination of social category theory has been done using experiments. These were dyads or groups that were temporarily created.

Similarity has been considered in terms of gender and age. Sometimes similarity in a dyad leads to positive outcomes; however, it can lead to negative outcomes as well. For instance, several studies show that similarity between a supervisor and employee leads to better outcomes, while other studies show that similarity leads to worse outcomes. For example, Mathison (1986) found that female managers rated assertive women more negatively than did male managers. Finkelstein and Burke (1998) found older participants who identified strongly with their age group rated older (in-group) applicants as being economically less valuable to their company than did younger participants.

Another factor that may influence how one member of the dyad evaluates another member of the dyad is if a member of the dyad would prefer to have the characteristic of the other member of the dyad. Graves (1995) found that female recruiters provided more favorable evaluations to male applicants than female. Using social identity theory, the researchers suggested that the female interviewers, found to hold lower status positions within their respective organizations, attempted to distance themselves from their own group and psychologically identify with the higher-status group – in this case males.

In summary, the literature review suggests dyads with similar characteristics tend to lead to more positive outcomes. However, there are some conditions when similar dyads lead to worse outcome. These cases are often when the dyad shares a characteristic that is perceived to be less desirable (e.g. being a woman or of a minority race). Building on this theoretical foundation, we propose the following two hypotheses.

Hypotheses

A person feels more positive toward people demographically similar to themself (Chatman & Flynn, 2001; Wharton, Rotolo, & Bird, 2000). A person attributes positive characteristics to individuals they believe are in their same group because it helps them feel better about themselves (Tajfel & Turner, 1986). Individuals seek to create and keep a positive self-concept. When they attribute positive attributes to those in their same category, they are helping to re-inforce their own self-concept. This process can get stronger over time. So, as a person seeks to remember things they tend to remember the positive things about people who are similar to them (Tavani, Collange, Rateau, Rouquette, & Sanitioso, 2017). For these reasons we expect that:

> **H1:** For those dyads from an ethnicity with a positive stereotype, the degree of ethnic similarity between the host's name and the guest's name will positively impact the sentiment in the review.

The effects of similarity are not uniformly positive. Sackett & DuBois (1991) offer evidence that majority members (i.e. whites) rate whites higher while minority members (i.e. blacks) do not rate other blacks higher. Minority members are more likely than majority members to stereotype themselves (Swan & Jr., 1997).

People tend to notice characteristics that distinguish them from others in the social setting (McGuire, McGuire, Child, & Fujioka, 1978). The typical explanation for this finding is that minority membership, because of its rarity, or numerical distinctiveness, automatically attracts attention and thus becomes salient in the perception of self and others (Fiske & Taylor, 1991, pp. 247–266; McGuire & McGuire, 1988). For instance, if a woman is in a group of men, her feminine characteristics become salient in her mind. When a characteristic is salient, it may cause the person to focus on their membership in a group with other people with that characteristic. That is, if being a woman is salient the person becomes conscious of being categorized as a woman. Being a woman might be salient if everyone else in the room is a man.

When a guest is a member of a marginalized minority and encounters a host who is also a member of the marginalized minority, his attention to characteristics related to this group are heightened. He also may feel vulnerable if he perceives the marginalized minority is associated with negative characteristics. When a person feels vulnerable, they are more likely to use stereotypes and prejudices to self-enhance (Fein & Spencer, 1997). People actively pursue social identities that optimize effective psychological functioning by providing self-esteem.

Individuals carry various approaches for combating lower self-esteem associated with being a member of stigmatized group. One technique for dealing with this is to disassociate from the group or deny being a member of it. For these reasons, Hispanic guests may be hypersensitive to negative characteristics of Hispanic hosts. In order to distance themselves from these traits they may offer more reviews that are negative. This makes it clear to themselves that they do not approve of these negative attributes and are distant from it. It helps improve the guests own self-esteem. When they experience a member of their stigmatized group they rate that person lower to make themselves feel better. Brown (1950) found that Black social workers tended to be more punitive toward clients of their own race than toward White clients.

If a Hispanic guest notices something negative about a Hispanic host he may be more likely to focus on it. He may do this because he feels the negative thing reflects poorly on him as a member of the group. So that when majority guests visit that host, negative stereotypes about Hispanics will be reinforced. The Hispanic guest does not want negative stereotypes about guests reinforced. Because Hispanic guests may perceive the cost of a negative Hispanic host as high, they may be more negative when evaluating them. The negative Hispanic host could lower the Hispanic guest's self-esteem (Lewis & Sherman, 2003).

> **H2:** For those dyads from an ethnicity with a negative stereotype, the degree of ethnic similarity between the host's name and the guest's name will negatively impact the sentiment in the review.

Methodology

Data

We use archival data collected from the AirBnB platform by InsideAirbnb. There are two separate Airbnb datasets: a listings dataset and a reviews dataset. The listing dataset contains 44,137 listings from October 2017 across the five boroughs of New York City. These data contain several important variables such as a unique identifier for each listing and the first name of the host.

The review data set contains 801,734 reviews across all the hosts. The review dataset contains unique identifiers for the listing, the host, and the guest. Plus, the review dataset has the first name of the guest and the comments left by the guest. There is no score or star-rating associated with the comments, and it is possible to rate the listing without leaving a comment for the host.

Ethnic Similarity

We categorize host and guest names by ethnicity using a dataset from the Harvard Dataverse (Tzioumis, 2018). The data set uses publicly available home mortgage data in the U.S. to match names with ethnicities. Each name is associated with a number representing the frequency with which it was associated with an ethnicity (e.g. white, black, Asian-Pacific Islander (API), Hispanic, Asian, or mixed-race) on the mortgage applications. The frequency counts are normalized to 1, so each name is represented as a vector with a normalized frequency count for each ethnicity.

First, we pre-process the host and guest first names from the AirBnB data. We remove all special characters. In cases where there are two host names listed (e.g., Larry and Deb), only the first name (e.g., Larry) is included. After processing the AirBnB names, we match the hosts' and guests' first names from with the ethnicity vector from Harvard Dataverse. We assign a category of "unknown" for hosts and guests whose names are absent from the database. We are able to associate 73% of the host and guest names with an ethnicity vector.

Next, we process the hosts' and guests' ethnicity vectors from normalized frequency for each category into a vector of binary indicators for each category. We use the variable that represents the percentage of times each name is associated with each race. For each name, when the percentage for a particular ethnicity is above the median, we code the name as 1 for that ethnicity. Each name has a percentage for each ethnicity and so a name could be associated with multiple ethnicities. For example, the name Maria is 59% Hispanic, 34% white, 5% Asian Pacific Islander, and 1% other categories. Because the median for Hispanic is 2.2%, Maria is above the median for Hispanic but below the median for the other categories. In this example, Maria is above the median for Hispanic so the host is assigned a 1 for Hispanic.

The ethnic similarity for each dyad is determined by whether the host and guest name are both above median for the same ethnicity categories. If a host and guest are both above the median for any dimension, then the similarity is 1, and it is 0 otherwise. For example, the name Sofia is 49% Hispanic and above the median for Hispanic. If the host and guest names are Maria and Sofia respectively, then the host and guest dyad is classified as 1 for Hispanic and as 1 for the same race. Chad is above the median for white, so if the host and guest names are Maria and Chad respectively, then the guest's race is white and the dyad is classified as 0 for same race.

Review Sentiment

To categorize sentiment in the guest's review, we perform text analysis on the text of the reviews. First, we use the AFINN dictionary (Nielson, 2011) to analyze sentiment. The AFINN dictionary is designed for use with microblogging websites like Twitter, and the reviews on Airbnb are similarly short in overall length. We used a bag-of-words model for the text of each review. The words were stemmed and stop words (e.g., it) were removed. Individual words are scored for positivity and negativity on a scale from 5 to -5. The total score for each comment is the aggregation of the AFINN score for each word.

Summary Statistics

The final dataset contains 382,206 guest/host dyads in which the first names of the guest/host are in the dataset and the comments contain more than stop words. There are 19,386 hosts each with between 1 and 838 guests. The summary statistics for the analysis data are available in Table 1.

Variable	N	Mean	Median	Min	Max
Sentiment Score	382206	11.12	10	-87	208
Similarity (Hispanic)	382206	0.248	0	0	1
Similarity (White)	382206	0.252	0	0	1
Mean Review Length	382206	64.42	50	1	1317
Number of Reviews	382206	79.24	62	1	1134
Mean Listing Price	382206	143.81	117	0	999
Review Score (0-100)	382206	93.61	94	20	100
Neighborhood Demographics (% white)	382206	52.54	57.1	1.9	94.90
Neighborhood Demographics (% Hispanic)	382206	23.64	20.30	0	74.5
Neighborhood Demographics (% Asian)	382206	11.51	7.8	0.5	73.3

Table 1. Summary Statistics

The average sentiment score for the reviews is 11.12. We can also see that the average host has a high average score of 93 out of 100, and the average amount charged across all listings is $143. About 25% of the dyads are both from the same ethnicity that is associated with a positive stereotype (both white). Another 25% of the dyads are associated with an ethnicity that is associated with a negative stereotype (both Hispanic).

Empirical Analysis

Model

To understand how host and guest similarity shapes the guest sentiment from the guest review, we use the model specification below. We estimate a fixed effects model to capture the latent attributes across hosts.

$$Sentiment = \beta_1 * SimilarityPositiveStereotypes + \beta_2 * SimilarityNegativeStereotypes + \beta_3 * Covariates + \epsilon_i$$

In this model, the dependent variable *Sentiment* is the AFINN sentiment score from the text analysis. *Similarity* is the binary indicator value for the ethnic similarity of the host-guest dyad based on their names. *PositiveStereotypes* is the binary indicator for whether the group is one with positive stereotypes, in this case the white guest-host dyad, and *NegativeStereotypes* is a binary indicator for groups with negative stereotypes, in this case the Hispanic guest-host dyads.

This model includes several types of covariates. First, we include the gender of the hosts and the guests. We identify the gender to the hosts and guests using the U.S. social security database. We collect all the names and genders for babies born in the U.S. between 1920-2010. For each name, we calculate a ratio of males and females with that particular name. If the ratio of males to females exceeds unity, then the name was categorized as a male name. Otherwise, it was categorized as a female name. This process is used to classify the gender of the hosts and guests. With this method, we achieve 90% coverage of all the host and guest first names.

Second, we include control variables for the listing neighborhood. We use Census data to find the population demographics for a given zip code and match the listing zip code to the Census demographics. For each zip code, we have the percent of the population that is white, Hispanic, Asian, and black.

Third, we include covariates for the review characteristics. Mean Review Length describes the mean number of words in each review by host and guest. Mean Listing Price is the average price for each listing by host. Number of Reviews is the total number of reviews for each listing. Review Scores are the numeric scores for each listing. Lastly, there is a fixed effect for each host to control for differences in the host.

Results

Table 2 shows the results from the fixed effects regression estimation that support H1 and H2.

	Dependent variable = Sentiment				
Variable	**Estimate (1)**	**Estimate (2)**	**Estimate (3)**	**Estimate (4)**	**Estimate (5)**
Similarity* Positive Stereotypes		0.448 (0.0337)***	0.495 (0.034)***		0.448 (0.034)***
Similarity* Negative Stereotypes		-0.407 (0.034)***		-0.459 (0.034)***	-0.407 (0.034)***
Controls					
Guest Gender – Male	-0.611 (0.0350)***	-0.634 (0.0350)***	-0.618 (0.035)***	-0.629 (0.035)***	-0.634 (0.035)***
Guest Gender – Male* Host Gender – Male	0.0212 (0.048)	0.025 (0.048)	0.022 (0.048)	0.025 (0.048)	0.026 (0.048)
Mean Review Length	0.069 (0.0002)***	0.0687 (0.0002)***	0.069 (0.000)***	0.069 (0.000)***	0.069 (0.000)***
Review Scores	0.184 (0.010)***	0.183 (0.0098)***	0.184 (0.010)***	0.184 (0.010)***	0.183 (0.010)***
Number of Reviews	0.004 (0.0007)***	0.004 (0.0007)***	0.004 (0.001)***	0.004 (0.001)***	0.004 (0.001)***
Mean Listing Price	-0.0009 (0.0006)	-0.0009 (0.0006)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001).
Neighborhood Demographics					Yes
R^2	0.1836	0.1844	0.1841	0.1840	0.1844

Table 2. Fixed Effects Results

The first model includes only control variables, and the control variables are significant. The length of the review, the guest-star rating, and the number of guests for each listing are all positively associated with sentiment. In general, men use less positive language than women in their reviews, although there is no evidence that gender similarity is associated with positive sentiment because the interaction term for Male Host and Male guest is not significant.

The next model tests our hypotheses. The coefficient estimate on Similarity * Positive Stereotypes (0.448) is significant and positive. This estimate is consistent across all the specifications, suggesting that similarity between the host and guest is associated with an increase in positive sentiment when the pair are both from ethnicities associated with a positive stereotype. Our finding supports H1.

Second, the coefficient estimate on interaction between Similarity * Negative Stereotypes is significant and negative (-0.407). This estimate is consistent across all the specifications, indicating that the similarity between the host and guest is associated with more negative sentiment when the pair is associated with the same ethnicity that has a negative stereotype. This finding supports H2.

Robustness Checks

To test the robustness of our findings to alternate categorizations, we perform a similar analysis with Asian hosts and guests as the ethnicity with positive stereotypes and Black hosts and guests as the ethnicity with negative stereotypes. Table 3 shows the coefficient estimates for the revised model.

	Dependent variable = Sentiment			
Variable	**Estimate**	**Estimate**	**Estimate**	**Estimate**
Similarity* Positive Stereotypes	0.281 (0.032)***	0.277 (0.032)***		0.281 (0.032)***
Similarity* Negative Stereotypes	-0.083 (0.024)***		-0.076 (0.024)**	-0.082 (0.024)***
Controls				
Guest Gender – Male	-0.617 (0.035)***	-0.605 (0.035)***	-0.623 (0.035)***	-0.618 (0.035)***
Guest Gender – Male* Host Gender – Male	0.020 (0.048)	0.020 (0.048)	0.021 (0.048)	0.021 (0.048)
Average Number of Words	0.069 (0.000)***	0.069 (0.000)***	0.069 (0.000)***	0.069 (0.000)***
Review Scores	0.184 (0.010)***	0.184 (0.010)***	0.184 (0.010)***	0.184 (0.010)***
Number of Reviews	0.004 (0.001)***	0.004 (0.001)***	0.004 (0.001)***	0.004 (0.001)***
Mean Price	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001).
Neighborhood Demographics				Yes
R^2	0.1838	0.1838	0.1836	0.1838

Table 3. Fixed Effects Regression Estimates for Sentiment: Robustness Checks

Our findings continue to support H1 and H2. The coefficient estimates for Similarity*Positive Stereotypes is significant and positive (0.277), indicating additional support for H1. The coefficient estimates for Similarity*Negative Stereotypes is significant and negative (-0.076) as well, suggesting additional support for H2. These

We perform additional measures of robustness. We analyze the emotional content of the words using the NRC dictionary, a dictionary of emotional words. Because the distribution of emotional words is heavily skewed in the review data, we use binary variables that indicate whether particular emotions were included in the reviews. Again, we find support for our findings that similarity influences review sentiment.

We also considered different ways to operationalize similarity and to model the relationship. We ran a different model with a single measure for similarity and another measure for stereotype valiance (positive or negative). This different specification yields results that are consistent with our findings, and our specification allows us to discuss positive and negative stereotypes in the same model. Plus, we try continuous score for similarity, and we find support for our hypotheses.

Discussion

Theoretical Implications

The study presented herein extends and enriches the literature on the impact of societal stereotypes in a technology mediated environment and the antecedents of online reviews. Our work highlights how the way hosts and guests categorize themselves impacts the reviews that they leave. Furthermore, our work offers evidence that the category itself changes how similarity impacts review behavior in online platforms. Whether the dyad belongs to a category that is associated with positive or negative stereotypes holds importance.

We also advance the literature that seeks to understand in-group and out-group favoritism. Prior research shows that female and non-white leaders devalue low-status leaders during evaluation and promotion. We find a similar pattern in the online review context. There are two reasons a stereotyped minority might rate other stereotyped minorities lower than majority members. First, there may be benefits to ingratiating one's self to the majority because the majority members may have power to improve the ethnic minorities' status. Second, there may be a psychological benefit to distance from other minorities.

In a corporate setting, where this phenomenon has been studied, an ethnic minority may benefit from becoming a part of the majority ethnic group. However, tangible benefits are less likely in an online review context. Unlike in a corporate setting where individuals may know each other and be expected to interact closely in the future, a guest may not have intimate relationships with the host or other people who see the review. The findings presented here suggests the behavior is driven by a more psychological place that is associated with psychological benefits (instead of more tangible benefits).

Practical Implications

Our work holds implications for platforms like AirBnB, its guests, and its hosts. Given that our results suggest that the race of the host and guest impacts the sentiment of reviews, the reviews may not accurately represent the guest experience. If AirBnB would like to improve the review system by reducing bias in the system, they may prevent host and guest race from being visible. As suggested, in these works, it is possible that stereotyped ethnicities feel a psychological sense of belonging with other ethnicities by offering them higher reviews than they would give to a host of their same ethnicity.

Limitations and Future Research

In considering similarity, we examine coarse categories of ethnicity. More nuanced categories exist and could offer further insight into the reviews posted (Coleman, Wampold, & Casali, 1995). For example, Hispanic can be broken into Mexican, Puerto Rican and other categories. Our plans for future work includes using a different measure, a binary indicator for the largest ethnic percent, as a robustness measure. Furthermore, we faced some limitations around knowing the actual ethnicities and so we inferred them. Future research can consider a survey as an alternative method of inferring ethnicities although the influence of the ethnicity is based on perceptions of others and not the individuals' perceptions.

Some work suggests that the impact of similarity changes over time (Lankau, Riordan, & Thomas, 2005; Turban, Dougherty, & Lee, 2002) and so exploring the impact of similarity for guests who stay with the same host multiple times may be an interesting avenue for future research.

Finally, future work can control for alternative explanatory mechanisms. For instance, controlling for the quality of the home and the stay may be a useful control variable.

Conclusion

Our work begins the discussion around the value of reviews and the factors that drive them. Given the importance of reviews in our society, we hope that this work sparks a line of continued research.

REFERENCES

Chatman, J. A., & Flynn, F. J. 2001. "The influence of demographic heterogeneity on the emergence and consequences of cooperative norms in work teams," *Academy of Management Journal*, 44(5), 956-974.

Chevalier, J.A. and Mayzlin, D. 2006. "The effect of word of mouth on sales: online book reviews," *Journal of Marketing Research*, 43 (3), 345-354.

Coleman, H. L., Wampold, B. E., & Casali, S. L. 1995. "Ethnic minorities' ratings of ethnically similar and European American counselors: A meta-analysis," *Journal of Counseling Psychology*, 42(1), 55.

Cui, R., Li, J., & Zhang, D. 2016. "Discrimination with incomplete information in the sharing economy: Evidence from field experiments on Airbnb,"

Doosje, B., Branscombe, N. R., Spears, R., & Manstead, A. S. 1998. "Guilty by association: When one's group has a negative history," *Journal of Personality and Social Psychology*, 75(4), 872.

Dwyer, S., Orlando, R., & Shepherd, C. D. 1998. "An exploratory study of gender and age matching in the salesperson-prospective customer dyad: Testing similarity-performance predictions," *Journal of Personal Selling & Sales Management*, 18(4), 55-69.

Edelman, B., Luca, M., & Svirsky, D. 2017. "Racial discrimination in the sharing economy: Evidence from a field experiment," *American Economic Journal: Applied Economics*, 9(2), 1-22.

Fein, S., & Spencer, S. J. 1997. "Prejudice as self-image maintenance: Affirming the self through derogating others," *Journal of Personality and Social Psychology*, 73(1), 31.

Finkelstein, L. M., & Burke, M. J. 1998. "Age stereotyping at work: The role of rater and contextual factors on evaluations of job applicants," *The Journal of General Psychology*, 125(4), 317-345.

Fradkin, A., Grewal, E., Holtz, D., & Pearson, M. 2015. "Bias and reciprocity in online reviews: Evidence from field experiments on Airbnb," *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, 641-641.

Glomb, T. M., & Welsh, E. T. 2005. "Can opposites attract? Personality heterogeneity in supervisor-subordinate dyads as a predictor of subordinate outcomes," *Journal of Applied Psychology*, 90(4), 749.

Graves, L. M. G. N. P. 1995. "The Effect of Sex Similarity on Recruiters' Evaluations of Actual Applicants: A Test of the Similarity Attraction Paradigm," *Personnel Psychology*, 48(1), 85-98.

Henningsen, A. and Toomet, O. 2011. "maxLik: A package for maximum likelihood estimation in R," *Computational Statistics* 26(3), 443-458. DOI 10.1007/s00180-010-0217-1

Kakar, V., Voelz, J., Wu, J., & Franco, J. 2018. "The Visible Host: Does Race Guide Airbnb Rental Rates in San Francisco?" *Journal of Housing Economics*, 25-40.

Lankau, M. J., Riordan, C. M., & Thomas, C. H. 2005. "The effects of similarity and liking in formal relationships between mentors and protégés," *Journal of Vocational Behavior*, 67(2), 252-265.

Lewis, A. C., & Sherman, S. J. 2003. "Hiring you makes me look bad: Social-identity based reversals of the in-group favoritism effect. Organizational Behavior and Human Decision Processes, 90(2), 262-276.

Mathison, D. L. 1986. "Sex differences in the perception of assertiveness among female managers," *The Journal of Social Psychology*, 129(5), 599-606.

McFarland, L. A., Ryan, A. M., Sacco, J. M., & Kriska, S. D. 2004. "Examination of structured interview ratings across time: The effects of applicant race, rater race, and panel composition," *Journal of Management*, 30(4), 435-452.

McGuire, W. J., McGuire, C. V., Child, P., & Fujioka, T. 1978. "Salience of ethnicity in the spontaneous self-concept as a function of one's ethnic distinctiveness in the social environment," *Journal of Personality and Social Psychology*, 36(5), 511.

Nielsen, F.A. 2011. "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs," *Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages* 718 in CEUR Workshop Proceedings 93-98. 2011 May. http://arxiv.org/abs/1103.2903.

Reed II, A., Forehand, M. R., Puntoni, S., & Warlop, L. 2012. "Identity-based consumer behavior," *International Journal of Research in Marketing*, 29(4), 310-321.

Sackett, P. R., & DuBois, C. L. 1991. "Rater-ratee race effects on performance evaluation: Challenging meta-analytic conclusions," *Journal of Applied Psychology*, 76(6), 873.

Sahdra, B., & Ross, M. 2007. "Group identification and historical memory," *Personality and Social Psychology Bulletin*, 33(3), 384-395.

Swan, S., & Jr., R. S. W. 1997. "Gender stereotypes and social identity: How being in the minority affects judgments of self and others," *Personality and Social Psychology Bulletin*, 23(12), 1265-1276.

Tajfel, H., & Turner, J. 1986. "The social identity theory of intergroup behavior," In S. Worchel & W. Austin (Eds.), Psychology of Intergroup Relations (pp. 7-24). Chicago: Nelson-Hall.

Tajfel, H., & Turner, J. C. 1979. "An integrative theory of intergroup conflict," In W. G. Austin & S. Worchel (Eds.), The social psychology of intergroup relations (pp. 33-47). Monterey, CA: Brooks-Cole.

Tavani, J. L., Collange, J., Rateau, P., Rouquette, M.-L., & Sanitioso, B. R. 2017. "Tell me what you remember and I will know who you are: The link between collective memory and social categorization," *Group Processes & Intergroup Relations*, 21(1), 91-108.

Turban, D. B., Dougherty, T. W., & Lee, F. K. 2002. "Gender, race, and perceived similarity effects in developmental relationships: The moderating role of relationship duration," *Journal of Vocational Behavior*, 61(2), 240-262.

Tzioumis, K., 2018. "Data for: Demographic aspects of first names", https://doi.org/10.7910/DVN/TYJKEZ, Harvard Dataverse.

Wharton, A. S., Rotolo, T., & Bird, S. R. 2000. "Social context at work: A multilevel analysis of job satisfaction," (Vol. 15): Kluwer Academic Publishers-Plenum Publishers.