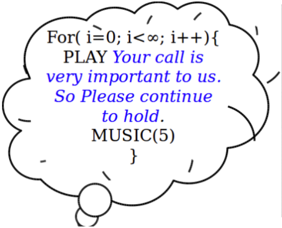
Racial Bias in Social Media Customer Service

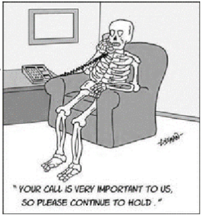
Priyanga Gunarathne, **Huaxia Rui**, Abraham Seidmann

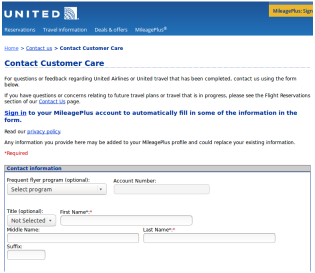
November 4, 2023

Agenda

IT Has Given Many Faces to Customer Service ...







UNITED

MileagePlus: Sign

Reservations Travel information Deals & offers MileagePlus®

Home > Contact us > **Contact Customer Care**

Contact Customer Care

For questions or feedback regarding United Airlines or United travel that has been completed, contact us using the form below.

If you have questions or concerns relating to future travel plans or travel that is in progress, please see the Flight Reservations section of our Contact Us page.

Sign in to your MileagePlus account to automatically fill in some of the information in the form.

Read our privacy policy.

Any information you provide here may be added to your MileagePlus profile and could replace your existing information.

*Required

Contact information

Frequent flyer program (optional): Select program

Account Number:

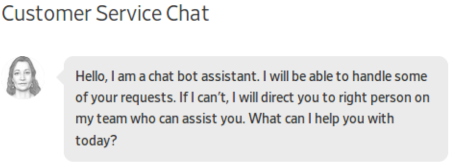
Title (optional): Not Selected

First Name*:*

Middle Name:

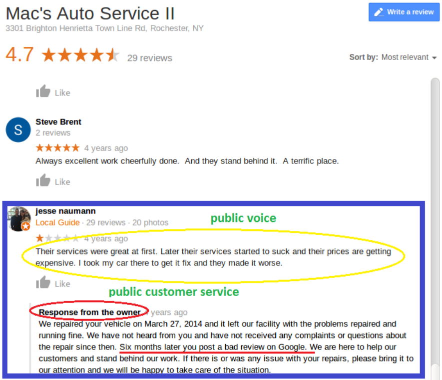
Last Name*:*

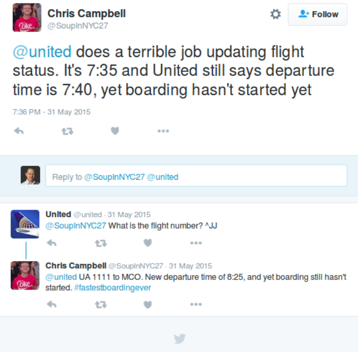
Suffix:



But One Feature Has Remained Unchanged ...

Until ...

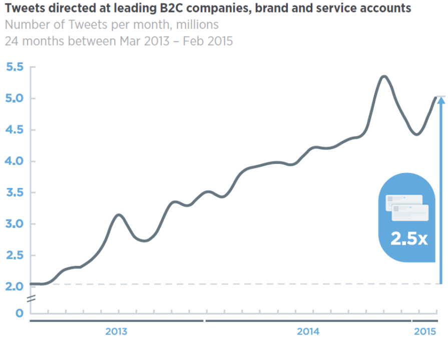




What make social media customer service (SMCS) so different from traditional channels?

Social Media Customer Service (SMCS) / Social Care

1. Public and connected
2. Extreme convenience



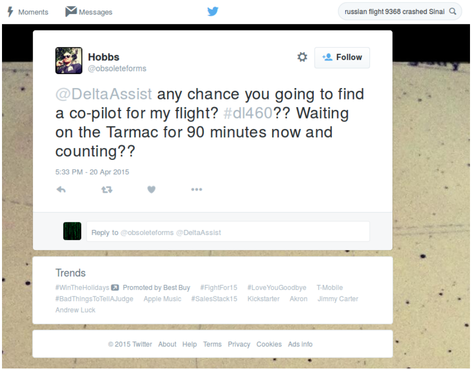
SMCS shifts power from firms to consumers.

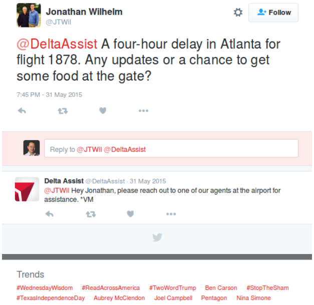
Response Rate and Delay to Complaining Tweets

In contrast with traditional call centers, on average, airlines respond to less than half of the tweets directed at them by complaining customers.

Airline	Response Rate (%)	Average Response Time (Minutes)
American Airlines	57.06	34.69
United Airlines	40.10	101.81
Southwest Airlines	35.99	221.03
Delta Airlines	46.70	17.52
JetBlue Airways	46.95	9.46
Virgin America	37.73	159.66
Alaska Airlines	52.42	81.48

Examples of Non-Response vs. Response





Research Question

Do airlines have (implicit) racial bias while delivering customer service on social media?

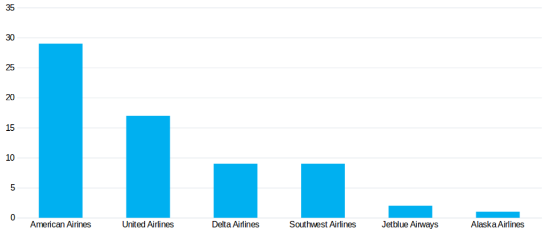


Figure: ATCR complaints of racial discrimination (2016-2017/8)

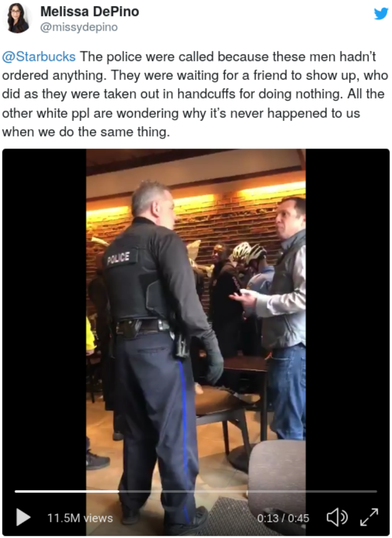
Discrimination, exclusion and unconscious biases are enormous problems that no one has mastered, and we would never suggest that we have it all figured out either ... What we know is we want to keep learning and we want to be even better. — Douglas Parker@AA

Agenda

Peer-to-Peer (**P2P**) Bias

- **Crowdfunding**: Loan listings with African-Americans in the attached picture are less likely to receive funding than similar whites. (Pope and Sydnor 2011).
- **Uber** and **Lyft**: longer waiting times and more frequent cancellations for African-American passengers (Ge et al. 2016).
- **Airbnb**: guests with distinctively African-American names are less likely to be accepted relative to identical guests with distinctively white names (Edelman et al. 2017).

Business-to-Consumer (**B2C**) Bias



A business entity is vicariously liable, under the *respondeat superior* doctrine, for negligent acts or omissions by their employees in the course of employment.

Table 1. Racial Bias Studies in Consumer Markets

	B2C	P2P
Physical /Offline	**Retail:** Schreer et al. (2009) **Housing:** Ondrich et al. (1999), Turner and Mikelsons (1992), Yinger (1993) **Car Sales:** Ayres (1991, 1995), Ayres and Siegelman (1995) **Health Care:** Blair et al. (2013), Penner et al. (2010), Sabin et al. (2008)	
Digital /Online	Current paper	**Airbnb.com:** Edelman and Luca (2014), Edelman et al. (2017) **Autobytel.com:** Morton et al. (2003) **Craigslist:** Ghoshal and Gaddis (2015), Doleac and Stein (2013) **eBay.com**: Ayres et al. (2015) **Prosper.com:** Pope and Sydnor (2011) **Kickstarter.com:** Younkin and Kuppuswamy (2018) **Uber and Lyft:** Ge et al. (2016)

Hypothesis

Complaints to social media customer service from a racial minority group, as cued by their social media profile pictures, are less likely to be responded compared to otherwise similar customers.

+

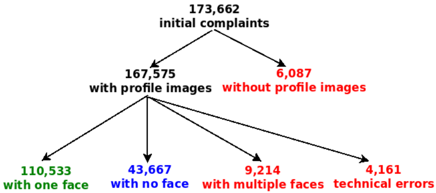
1. Some employees may dislike customers of racial minorities and be reluctant to serve them on social media.
2. Because of limited business-relevant information about social media users, some employees may assess customer value based on their racial identities.

−

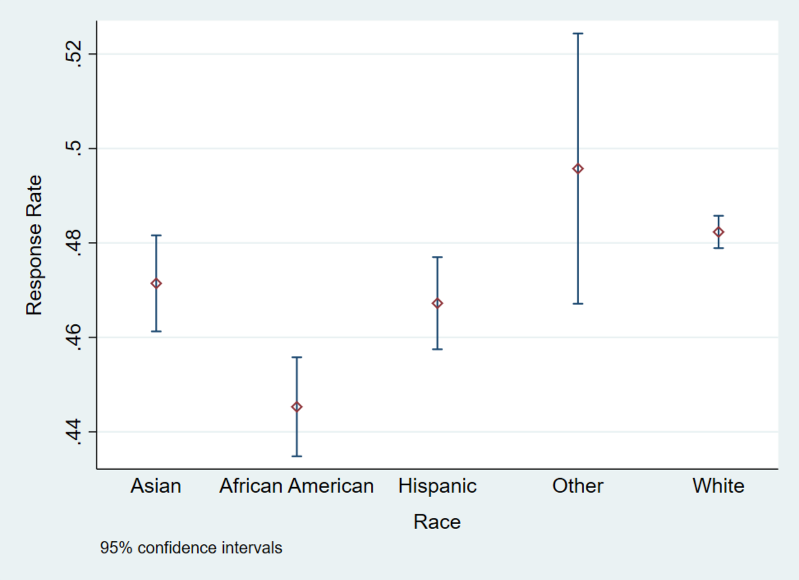
1. Compared to P2P bias, the stake of B2C bias is much higher for any company nowadays.
2. Unlike other forms of customer service where customer-brand interactions are private, social media customer service interactions are public.

Agenda

Data




```
{
    "images": [
        {
            "faces": [
                {
                    "attributes": {
                        "age": 36,
                        "asian": 0.00418,
                        "black": 0.95347,
                        "gender": {
                            "femaleConfidence": 0,
                            "maleConfidence": 1,
                            "type": "M"
                        },
                        "glasses": "None",
                        "hispanic": 0.00206,
                        "lips": "Apart",
                        "other": 0.00225,
                        "white": 0.03804
                    },
                    "chinTipX": 249,
                    "chinTipY": 403,
                    "confidence": 0.99992,
                    "eyeDistance": 126,
                    "face_id": 1,
                    "height": 323,
                    "leftEyeCenterX": 175,
```


Response Rate
.52
.5
.48
.46
.44
Asian
African American
Hispanic
Other
White
Race
95% confidence intervals

- **Dependent Variable**: 1 if responsed, 0 otherwise
- **Independent Variable**: race, 1 if African-American, 0 otherwise
- **Control Variables**:
  1. **cluster fixed effects** to control for content topic and style
  2. **traffic**: # of complaints received in the previous hour
  3. **linguistic features** including length, complexity, offensiveness, AFINN sentiment intensity, hashtag, URL
  4. # of retweets
  5. **user characteristics**: gender, # of followers, # of updates, profile description
  6. airline fixed effects and day-of-week fixed effects

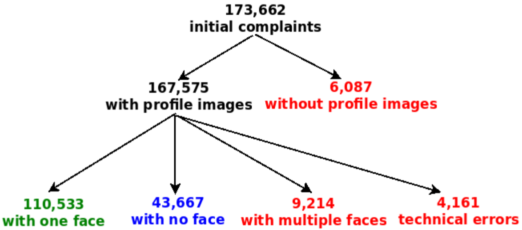
Agenda

Table 6. Estimation Results							
Variables	**Pooled**	**African American and White**		**Asian and White**		**Hispanic and White**	
	(1)	**PSM (2)**	**All (3)**	**PSM (4)**	**All (5)**	**PSM (6)**	**All (7)**
African American *(baseline: White)*	**-0.1231***	**-0.1181**	**-0.1258***				
	(0.0357)	**(0.0471)**	**(0.0358)**				
Asian *(baseline: White)*	0.0123			0.0052	0.0135		
	(0.0381)			(0.0517)	(0.0382)		
Hispanic *(baseline: White)*	0.0798					0.1899	0.0860
	(0.1040)					(0.1687)	(0.1045)
Other *(baseline: White)*	-0.5678*						
	(0.2952)						
Female	-0.0107	-0.0534	-0.0121	0.0241	-0.0024	0.2141	-0.0019
	(0.0187)	(0.0479)	(0.0194)	(0.0524)	(0.0196)	(0.1714)	(0.0203)
Observations	57,484	9,010	53,354	7,288	52,488	860	49,273
Log Likelihood	-34177.17	-5300.866	-31696.73	-4366.387	-31199.21	-435.4723	-29220.76
AIC	68504.34	10745.73	63537.47	8876.774	62542.42	1014.945	58585.53
BIC	69176.28	11257.37	64177.17	9373.141	63180.94	1357.444	59219.5

NOTE: For brevity, estimated results for a set of selected variables are reported. For all results, see Section A.8 of Appendix A.
*** p<0.01, ** p<0.05, * p<0.1 (Robust standard errors in parentheses)

Falsification Test

- The premise underlying our strategy to test the hypothesis is that visual cue is the most likely trigger of racial bias.
- So without a visual cue, we shouldn't be able to detect the bias. Otherwise, we have a spurious correlation.



Race Detection from Text

- 15,000 Twitter users in the U.S. who tweet in English, and have a valid profile image and a public timeline.
- Tag the race and gender using 849 AMT workers where each HIT (Human Intelligent Task) is labeled by three different workers from the U.S.
- Only users with at least two agreements and with specific race category are considered valid.
- 8,187 users with valid race labels; 8,971 users with valid gender labels.
- Fetch the latest 3,200 tweets for each valid user and use that as input to a deep learning-based race classifier.

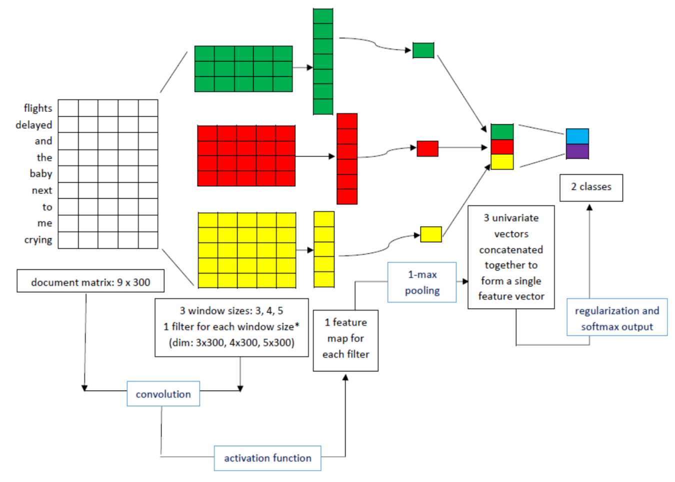


Figure B1. Deep-Learning Model Architecture
*Note: Only one filter for each window size is shown in the figure for demonstration purposes. In the actual implementation, 100 filters for each window size were used.**

Performance of the CNN

Table 3. Race Classifier - Prediction Performance								
	African American			Other			Average Accuracy	Average F1
System	Precision	Recall	F1	Precision	Recall	F1		
(1) Deep Learning	81.07	98.5	88.94	98.09	77	86.27	87.75	87.61
(2) SVM	78.54	86.00	82.01	84.53	76.5	80.31	81.25	81.16

Table 4. Gender Classifier - Prediction Performance								
	Female			Male			Average Accuracy	Average F1
System	Precision	Recall	F1	Precision	Recall	F1		
(1) Deep Learning	92.51	88.29	90.35	88.8	92.86	90.78	90.57	90.56
(2) SVM	91.25	89.43	90.33	89.64	91.43	90.52	90.4	90.43

Falsification Test Results

Table 7. Estimation Results — Falsification Test		
Variable	**(1) PSM**	**(2) All**
African American (*baseline: non-African American*)	0.0636 (0.1022)	0.0340 (0.0703)
Female	-0.0620 (0.1384)	0.0336 (0.0299)
Observations	2,212	43,048
Log Likelihood	-1161.827	-24756.58
AIC	2467.655	49657.16
BIC	2878.174	50281.4

For brevity, estimated results for a set of selected variables are reported. See Section A.11 in Appendix A for details.
*** p<0.01, ** p<0.05, * p<0.1 (Robust standard errors in parentheses)

Agenda

Why?

Sources of Discrimination

1. Animus-based (or taste-based) discrimination occurs when a certain group is treated differently because that group is disliked or hated.
2. Statistical discrimination theories predict that disparate treatment stem from decision-makers' use of observable characteristics as proxy for unobservable but outcome-relevant characteristics.

Conscious/Unconscious

1. Explicit bias is the traditional conceptualization of bias. With explicit bias, individuals are aware of their prejudices and attitudes toward certain groups.
2. Implicit bias refers to the attitudes or stereotypes that affect our understanding, actions, and decisions in an unconscious manner.

Why??

Unlike explicit bias, implicit bias is much harder to eradicate.

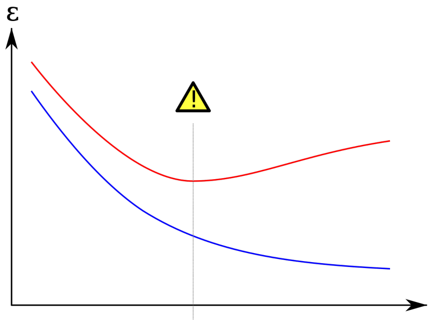


Figure: The challenge of generalization

Humans are **naturally** good at generalization.

- Maybe too good?
- What if implicit bias is the flip side of overfitting?

Humans are **naturally** good at generalization.

- Maybe too good?
- What if implicit bias is the flip side of overfitting?

How do we fight implicit bias? **Less data** ...

- For social media customer service, just mask customer profile pictures from agents.