Proceedings of the Tenth International AAAI Conference on
Web and Social Media (ICWSM 2016)

Twitter's Glass Ceiling: The Effect of Perceived Gender on Online Visibility

Shirin Nilizadeh
Computer Science Department
University of California at Santa Barbara, CA, USA
nilizadeh@cs.ucsb.edu

Anne Groggel and **Peter Lista**
Department of Sociology
Indiana University Bloomington, IN, USA
agroggel, plista@indiana.edu

Srijita Das, Yong-Yeol Ahn, and **Apu Kapadia**
School of Informatics and Computing
Indiana University Bloomington, IN, USA
sridas, yyahn, kapadia@indiana.edu

Fabio Rojas
Department of Sociology
Indiana University Bloomington, IN, USA
frojas@indiana.edu

Abstract

Social media is a new public sphere where people can, in principle, communicate with each other regardless of their status. However, social categories like gender may still bias online communication, replicating offline disparities. Examining over 94,000 Twitter users, we investigate the association between *perceived gender* and measures of *online visibility*: how often Twitter users are followed, assigned to lists, and retweeted. Our analysis shows that users perceived as female experience a 'glass ceiling,' similar to the barrier women face in attaining higher positions in companies. For users in lower quartiles of visibility, being perceived as *female* is associated with more visibility; however, this tendency flips among the most visible users where being perceived as male is strongly associated with more visibility. Our results suggest that gender presented in social media profiles likely frame interactions as well as perpetuates old inequalities online.

Introduction

One of the most important developments in early 21st century computing is social media. As a new, enhanced 'public sphere'[1] online platforms such as Twitter, Facebook, and Reddit allow users, in principle, to share and transmit ideas to potentially large audiences. Despite the promise of social media, existing studies show that online attention is highly skewed with a relatively small number of users attracting a disproportionate amount of attention (Kwak et al. 2010; Wu et al. 2011). Thus, an important theoretical agenda for the study of social media and online communities is to understand why some users attract more attention while other users go unnoticed.

Empirical research shows that many factors influence how much attention one receives on social media (Suh et al. 2010). User behaviors affect attention: for instance, users who post more or post about certain topics may attract more attention online (Romero et al. 2011; Wu et al. 2011; Weng et al. 2010). Individuals or groups with substantial 'real world' reputations, such as politicians, can readily gain online visibility (Morales et al. 2014). It is relatively less understood, however, *how social status, such as gender, interacts with online attention*. Differences in visibility can manifest in two ways: first, an individual's social status affects online behavior, which then can affect attention. Second, an individual's demographic information contained in user profiles may impact the amount of attention they receive. A female user, for example, might achieve less visibility because of gender bias. The "action-identity" theory that we offer is that the decision to follow someone on social media is affected by the user's behavior (actions) and the information they offer about themselves (identity) through their profile. In other words, actions and identities influence visibility. This theory speaks to previous scholarship on how individuals communicate online identities through their behavior (Kim, Zheng, and Gupta 2011) and ethnographic research on computer-mediated communication that examines interactions within an online social context (Carter 2005). While some social media users may intentionally misrepresent their gender, such deception has little bearing on the reality of potential bias against their 'perceived' identity.

We implement two gender inference methods of Twitter profile images and user names to assess how user actions and attributes contribute to online visibility. The correlation of demographic characteristics and user behavior is not the focus of this paper. Instead, we focus on: *how identity affects a user's attention and visibility in social media.*

Motivated by the "action-identity" theory of visibility, we test whether the gender inferred by user profiles affects users' number of followers, how many lists they are added to, and how often they are retweeted on Twitter. In other words, does presenting oneself as female or male confer an advantage or disadvantage in developing visibility controlling for other online behaviors? Thus, we test the following hypotheses:

- H1 Perceived female Twitter users have fewer followers than perceived male and gender ambiguous users.
- H2 Perceived female Twitter users appear on fewer lists than perceived male and gender ambiguous users.
- H3 Perceived female users are less likely to be retweeted than perceived male and gender ambiguous users.

[1] Jürgen Habermas used the term "public sphere" to describe the social space where people gather freely to discuss matters of interest on roughly equal footing (Habermas 1991).

Motivated by the 'glass ceiling effect' literature which posits that women face an invisible barrier at the highest levels of an organization (Cotter et al. 2001), we propose a more fine-tuned hypothesis:

H4 The relationship between perceived gender and measures of visibility will operate differently in lower and higher levels of visibility: perceived female users are penalized more at higher tiers of visibility.

We test the hypotheses with a random sample of over 94,000 Twitter accounts collected in 2015. Our analysis confirms that there is a significant association between gender and visibility, especially at the highest tier of visibility. For instance, among users with fewer followers, those with female names or profile images have slightly more followers than others. But among those with many followers, those with male names have more followers. We posit that social status, and gender in particular, may still play an important role on social media. Our findings suggest that there exists a 'glass ceiling' on social media, where disadvantages against female users are found among the users with high visibility. These differences exist even after we control for account age (years), number of lifetime tweets, the 20 most common topics that appear in the Twitter who-is-who database, offline popularity through verified accounts, and five other factors. This finding mirrors social science literature, which documents greater gender disparity as women reach the 'highest' levels of their occupation (Williams 2005).

Related Work

Gender Inequality and Glass Ceiling

Our work speaks to a growing body of literature which investigates how similarities or differences between users' identities and social statuses provides a mechanism for evaluation (Greenwald, McGhee, and Schwartz 1998; Eagly and Karau 2002; Anderson et al. 2012). In interactions, individuals often categorize others based on their education, gender, ethnicity, or occupation (Willis and Todorov 2006). These social categories often come embedded with a set of assumptions and stereotypes. For instance, men are generally assumed to be more competent than women (Fiske et al. 2002). These beliefs can affect how much individuals participate, how their input is evaluated by others, or the level of influence or authority they achieve (Ridgeway 2001).

Gender inequalities — including differences in social status, political influence, and control over material resources — continue to persist offline and are thus an important area of study (Padavic and Reskin 2002; Tilly 1999). Multiple disciplines report occupational gender segregation (Sugimoto, Ni, and Larivière 2015) and gender disparities in pay and employment, especially for mothers (Correll and Benard 2007).

Research has shown a gender bias, at least implicitly, in professional settings (Oakley 2000; Tomaskovic-Devey et al. 2006). Women struggle to break through an occupational barrier, or glass ceiling, whether it be a scholar's citations and output (Ward, Gast, and Grant 1992; Sugimoto et al. 2013), promotion to supervisory positions (Reskin and Ross 1992), or wage inequalities (Merluzzi and Dobrev 2015). The question of whether similar barriers exist online for users perceived as female remains unanswered.

Online Social Media and Gender

Gender can frame online interactions and impact how people use social media. Research shows that women initially lagged behind men in access to and usage of the Internet (Bimber 2000). While men and women use Twitter in roughly equal proportions today, (Duggan and Smith 2013) research shows that on many social media platforms men and women continue to differ in their interpersonal communication styles (Kivran-Swaine et al. 2012), what they present or conceal about themselves (Quercia et al. 2012), and their online behavior (Kim, Sin, and He 2013; Thelwall 2008; Ottoni et al. 2013). There is growing interest in identifying the demographics of social media users and their characteristics (Mislove et al. 2011; Chang et al. 2010). Studies show gender disparities on Wikipedia in content, contributors, and revisions for the top tier of Wikipedians (Wagner et al. 2015; Reagle and Rhue 2011; Lam et al. 2011; Antin et al. 2011; Lam et al. 2011). As more researchers investigate the influence or popularity of social media users and groups, examining whether offline status hierarchies exist online becomes a rich new avenue of study (Aral and Walker 2012; Magno and Weber 2014; Garcia, Weber, and Garimella 2014).

Visibility and Influence Online

Given the vast potential audience in global social media, many users naturally strive for more attention. Larger visibility, attention, and audience can be translated into monetary benefits and influence on public discourse (Jürgens, Jungherr, and Schoen 2011; Vergeer and Hermans 2013). Thus, many studies quantify user's 'visibility' and 'influence' using a wide array of techniques and metrics (Brown and Feng 2011; Aral and Walker 2012). We underline the distinction between influence and visibility. Influence focuses on the *changes* that one can make, while we use the term 'visibility' to refer the size of the audience or the *reach* one has, which does not necessarily translate into *effects*. Celebrities tend to be highly visible, but are not necessarily influential, while domain experts may have less visibility but stronger influence. We use basic metrics of online visibility as a first step towards measuring online gender disparities.

Data and Methods

Twitter User Data Collection

We collected data from Twitter's 'Gardenhose,' which includes about a 10% sample of Twitter's entire public stream (McKelvey and Menczer 2013). All data collection was done according to regulations outlined by our organization's human subjects ethics board. Because the Gardenhose is an activity-based sample, sampling users by selecting random tweets or by examining a small time window will preferentially sample more active users. To minimize this bias, we first obtained a set of 34,470,349 users who appeared in the Gardenhose stream in the month of February 2015, and then randomly sampled 104,179 users.

To minimize the effect of bots on our analysis, we removed the user accounts whose followers-to-friends (followers-to-followee) ratio is less than 0.1, following a previously suggested spam-filtering procedure (Thomas et al. 2011). Through this filter, we removed about 9% of the accounts, which is close to the previously reported estimates on the number of social bots on Twitter (Seward 2014). The filtering left 94,645 users in our final sample. We cross-checked the effectiveness of this approach by using the 'bot-or-not' API, which implements a machine learning-based method with many features (Ferrara et al. 2014). We sampled 200 user accounts and assessed them through the bot-or-not application. Among our sample, about 84% of those could be evaluated through bot-or-not; the users that could not be accessed by bot-or-not included those who had changed their screen name, made their profiles protected, or left Twitter. Out of the accounts examined through bot-or-not, only three (1.7%) were identified as a bot, according to a threshold of 80% in the bot-or-not score. 137 users out of 168 users (about 80%) from our random sample had a probability of less than 50%, suggesting that it is unlikely that our results are heavily skewed by the presence of bots. Also note that even having bots in our sample does not invalidate our approach because we examine the relationship between users' self-representations and their visibility.

Finally, we collected the first tweet of each user during February 2015 and obtained all available metadata about the tweet and the user. Our sample of 94,645 includes information about users' followers, friends, account age, number of tweets, listed count, profile image, location, profile description, URL, and whether the account is verified.

Retweeted Counts. For each user in our dataset, we counted all the retweets of the user's tweets that appear in the Gardenhose in February 2015. This count does not show the total number of retweets for a user because we only had access to a 10% sample of Twitter data. However, this data still tells who is retweeted relatively more or less than others.

Topics of Expertise. In an effort to control for potential confounding factors, we extract the information about topics of expertise or interests for active users. Twitter has a 'List' feature where users can organize others by creating topical feeds-lists of users. For instance, one may create a "Data Scientists" list and organize Twitter users into this list to keep track of tweets from those in the list. It has been shown that one can accurately identify domain experts by examining the lists that people are on (Sharma et al. 2012). That study also led to a "Twitter who-is-who" service, which we use to get the 'topic' labels (topics of expertise and interests) of users in our sample. We extracted the topic labels of 3,414 (out of 94,654) users and associated strength — number of times listed — of each topic label using the service. We manually cleaned several top topic labels. For instance, we merged 'businesses,' and 'biz' to 'business' and removed labels such as 'best,' 'bro,' 'new,' and 'uk.' Then, we extracted the 20 most popular topics: 'art,' 'artists,' 'bloggers,' 'business,' 'companies,' 'design,' 'info,' 'journalists,' 'life,' 'marketing,' 'media,' 'music,' 'news,' 'organizations,' 'politics,' 'sports,' 'technology,' 'tv,' 'world,' and 'writers.'

Gender Inference

We combine two gender inference methods to accurately assess the perceived, or presented, gender of Twitter profiles.

Name based Gender Inference. All Twitter users have a screen name of their choice (e.g., "@POTUS"). In addition, users usually populate a 'username' field in their profile. Users either disclose their real name (e.g., "Barack Obama") or put any text string of their choice. Even if users do not disclose their real name, any name put in the user name field may trigger others to view their account as a male or female one. Regardless of true gender and real name, if a user puts "Anne" in the profile it is likely that the other Twitter users perceive them as female.

To determine the gender of a Twitter user's name, we employ the U.S. Census-based method (Mislove et al. 2011). The U.S. Census documented 92,626 unique first names and their gender profiles in the U.S. from 1900 to 2013, representing one of the most comprehensive data sets on gender and names available (Center 2013). Although there exist many non-English names in the dataset, we restrict our analysis to the users who selected English as their language in order to more accurately estimate culturally relevant schemas for the gender of names.

We first extract the first word from the 'username' field of each user profile. We consider a 'username' a valid first name if it appears in the U.S. Census database. To assess the gender of users' first names, we consult the U.S. Census database, which documents the fraction of males and females with each first name. Although a majority of names is gendered (e.g., almost every "Anne" is female), there are gender-neutral names, such as "Pat," which exhibits a 40:60 female-to-male ratio. We classify first names into two classes: 'strongly gendered' and 'weakly gendered' names. Strongly gendered names exhibit a female-to-male ratio larger than 0.95 or smaller than 0.05; the other names are considered to have weak gender associations. We label users with strongly gendered names based on the dominant gender of the names while labeling users with weakly gendered names as 'gender ambiguous.'

Because we are interested in the association between visibility and perceived gender, our discussion and analysis focuses on strongly gendered users, those clearly perceived as male or female by others. However, we include comparisons with 'gender ambiguous' users and 'unknown' users, who did not have a valid first name in their user profiles. These additional categories are relevant because they reveal new avenues of study where disparities may exist between users clearly perceived as male or female, users with purposeful or accidental gender ambiguity, and other accounts such as those belonging to organizations. To assess the accuracy of this method, three coders manually inspected a random sample of 200 Twitter users. Accuracy was computed by comparing the gender estimated by coders with the values from the gender-name association method. Using majority vote, our gender assessment had 81.5% agreement with coders.

Table 1: Gender detection by first name and profile image

Gender	by name	by image
Total	54,195	15,463
Female	23,287 (43%)	5,721 (37%)
Male	23,352 (43%)	6,861 (44%)
Ambiguous	7,556 (14%)	2,881 (19%)

Table 2: Confusion matrix for users with name and image

		By Name		
		Female	Male	Ambiguous
By image	Female	3,243	329	459
	Male	325	3,840	544
	Ambiguous	746	934	217

Profile Image based Gender Inference. To improve the accuracy of gender inference as well as to cross-evaluate the name based inference method, we use an automatic facial feature recognition service 'Face++' (Fan et al. 2014). Recent developments in deep learning have significantly improved the performance of computer vision methods. Particularly in face recognition tasks, computer vision techniques have surpassed human performance (Taigman et al. 2014). Online services such as http://how-old.net by Microsoft (Team 2015) and Face++ have begun to provide APIs for detecting gender, race, age, etc. from a profile image and achieves high levels of recognition accuracy (Zhou, Cao, and Yin 2015). In August 2015, we used the Face++ online API and collected gender label of 15,463 users (19,072 users were assessed and the service failed to identify a person in 3,609 cases). We labeled users as ambiguous if the confidence level reported by Face++ was below 95%.

Combining Results of Name based and Image based Gender Detection. Table 1 compares the gender identification results by name and image. Table 2 shows the conflict between the two gender detection methods. The image based method identified the gender of 15,463 users and its intersection with the name based results (total 54,195 users) contains 10,637 users. We measured the agreement between the two methods and found a total of 654 conflicts. These cases were manually assessed to categorize 285 users as female, 236 users as male and 133 (113 cases of name-image conflicts, 11 organizational accounts, and 9 inactive accounts) as ambiguous. In sum, incorporating profile images in our gender identification allowed us to detect the gender of 3,844 additional users who had either not provided a name or were classified as gender ambiguous by the name. Furthermore, these two methods displayed good agreement where both identified 3,243 and 3,840 users as female and male and 217 as ambiguous. The Cohen's Kappa statistic for agreement between these two methods is 0.49.

In sum, we categorize users into four classes:

Female: users who have strongly gendered female names, profile images, or both.

Male: users who have strongly gendered male names, profile images, or both.

Ambiguous: users whose names or images do not strongly signal either gender. Users may have gender-neutral names (e.g. "Pat"), gender-ambiguous profile images, or a conflict in name and image that cannot resolved even with manual inspection.

Unknown: users whose first name did not appear in the US Census database and their images could not be assessed.

In our sample of 94,645 users, we identified 25,953 (27.4%) 'male' users and 25,394 (26.8%) 'female' users. 7,676 (8.1%) users are 'gender ambiguous' and 35,622 (37.6%) users fall into the 'unknown' category.

Variables and Statistical Models

We use multivariate regression to study the associations between gender and visibility. Since age of the accounts, description, number of tweets, and other user activity variables can affect Twitter visibility, we use such factors as control variables. Below, we explain our dependent, independent, and control variables in detail.

Dependent Variables

As the key proxies of *online visibility* on Twitter, we select the following three dependent variables: the 'number of followers, the 'retweeted counts,' and the 'listed counts.' The number of followers is the most basic measure of visibility, as it reflects the size of one's audience (Hong, Dan, and Davison 2011). Yet, the number of followers cannot paint the full picture for several reasons: first, because reciprocity plays a role — people often follow back if someone follows them — some users have many followers simply because they follow many users (Kwak et al. 2010). Second, one can create or buy followers (Messias et al. 2013). Third, the types of followers can be important. For instance, someone followed by a small number of celebrities and experts may be able to make a larger impact on public discourse than others with more, but less important followers (Cha et al. 2010).

Retweets and lists provide complementary information about visibility. Because being retweeted implies that someone considers the tweet worthy, retweets have been considered as another basic proxy of visibility and influence (Zaman et al. 2010). As explained above, lists strongly signal expertise (Sharma et al. 2012) and capture more domain-specific visibility (Velichety and Ram 2013).

Control Variables

Because many factors can affect the dependent variables (e.g., following many usually guarantees many followers) and our focus is the effect of perceived gender, we incorporate the following control variables in our models.

Basic Activity Variables. Activity is a key driving force for visibility. Older accounts have had more time to accumulate followers, users who tweet more have more chances to be noticed, and users who follow many others are more likely to have more followers by sheer reciprocity. Thus we consider: 'the number of tweets,' 'the number of friends (followees),' and 'the account age' (years) as controls.

Profile Characteristics. The way users present themselves in their profiles may affect how others perceive them, and evaluate their posts (Ridgeway 2001). To account for self-presentation, we consider several variables: whether the user has provided *a description*, *an external URL*, and *a location*. The differences in visibility can be attributed to offline status (e.g. celebrities). Thus, we examine the *verified* labels.[2] We also consider the *length of the user's description* (in characters) to account for the amount of self-exposure in the profile description (Otterbacher 2010).

Topics. The gender differences in topical expertise and interests on Twitter may explain the differences in visibility (Java et al. 2007). For instance, imagine that Twitter contains a huge topical group about politics and most members of this cluster are male. If Twitter users tend to follow other users who are interested in similar topics based on homophily, male users will display more visibility simply because of the size distribution of the topical clusters. We take this effect into account by extracting the 20 most common topics from the Twitter who-is-who database (see "Topics of expertise" in Data and Methods) (Sharma et al. 2012) and including them as Binary variables.

In sum, we examine the following control variables: account age (years), number of tweets, number of friends, has_URL, has_location, has_description, description length, verified, art, artist, bloggers, business, companies, design, info, journalists, life, marketing, media, music, news, organizations, politics, technology, TV, world, and writers.

Statistical Models

We explore the following three topics. First, we study associations between perceived gender and control variables (user activities and characteristics). Second, we investigate overall associations between users' perceived gender and Twitter visibility. Third, we focus on how the relationship between users' perceived gender and visibility operate differently across visibility quartiles.

Perceived Gender and Behavioral Characteristics. We examine how each gender presents themselves differently. We assess the statistical significance of the difference between genders using the appropriate tests such as the t-test, chi-square test, and Mann-Whitney *U* test.

Perceived Gender and Visibility. We apply multiple multivariate regression models and present the results from our Poisson regression model. Linear and (zero-inflated) negative binomial regression models show qualitatively consistent results, although some did not converge.

Twitter Glass Ceiling: Perceived Gender and Tiers of Visibility. Foundational studies of glass-ceiling effects examined data based on quartiles (Cotter et al. 2001) as the effect is about the top portion of the hierarchy. Moreover, since dependent variables often exhibit a skewed distribution, examining the whole population may not capture more nuanced patterns (Yu, Lu, and Stander 2003). Thus, we adopt the quartile regression technique to analyze our dataset in detail. We divide the data into quartiles based on each dependent variable and apply multivariate regression models. We report the results of Poisson regression but the results are robust across multiple regression models.

Results

Descriptive Statistics

In our sample, about 62% of the accounts provided first names or profile images. Among those about 43% were categorized as female, 43% as male, and 14% as gender ambiguous. Figure 1 shows the number of users by gender in each quartile for each dependent variable. Since men and women are roughly balanced in each quartile our results cannot be attributed to the possibility that the ratio of male to female users is substantially different in the top quartile than in other quartiles.

Figure 1: Number of users per quartile

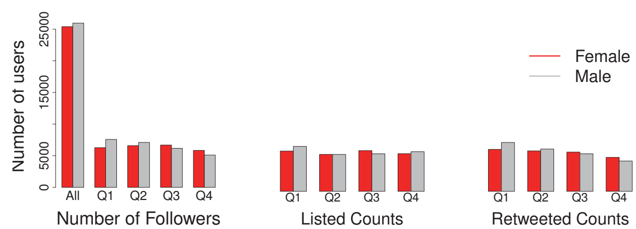


Table 3 shows the descriptive statistics of our variables. On average, users in our sample have roughly 1,400 followers, follow 502 other accounts, write over 7,000 tweets, and have accounts that are about 3 years old. Only 31% of the sample has been retweeted at least once and less than half (49%) have been listed. Most accounts upload profile images (about 98%) and provide a description (about 96%) but fewer use URLs (about 27%) and very few (about 0.3%) are verified.

In our sample, 3% of users are included in one of the 20 most common lists. Of these, news and media are the most common lists with 1.3% and 1.1% of users respectively.

Twitter Behavior by Gender

We first examine differences in each activity variable by gender, which provides insights into not only our models, but also general gender differences online.

Self Presentation in Profile. We investigate how individuals present themselves through their profiles, i.e., the profile description, account age, and whether they provide an external URL or location. The t-tests for account age and description length suggest that, on average, male users have older accounts than female users (M=3.37, SD=1.86 vs. M=3.16, SD=1.82) (t=-12.8, $p < 0.001$) and male users have longer descriptions than female users (M=63.84, SD=87.74 vs. M=60.67, SD=82.28) (t=-4.2, $p < 0.001$).

Table 4 shows the results of chi-square tests for the binary variables. Perceived male users are more likely to have

[2]Twitter 'verifies' accounts of some famous people and organizations. When accounts are verified a badge appears next to the user's name on their profile (Marwick and others 2011).

Table 3: Descriptive Statistics. We excluded those who have not been retweeted or listed when computing the median for the retweeted and listed variables.

	Type	N	Min	Max	Median
		Dependent variables			
Follower	Count	94,645 (100%)	1	42,391,843	204
Retweeted	Count	29,789 (31%)	1	46,388	2
Listed	Count	46,848 (49%)	1	76,290	2
		Independent variables			
Male	Boolean	25,953 (27.4%)			
Female	Boolean	25,394 (26.8%)			
Ambiguous	Boolean	7,676 (8.1%)			
		Control variables			
Account age	Years	94,645 (100%)	0	9	3
Desc. length	Count	94,645 (100%)	0	1,820	45
Friends	Count	94,645 (100%)	1	550,397	239
Tweets	Count	94,645 (100%)	0	1,040,973	1,776
Has desc.	Bool.	90,856 (96%)			
Has URL	Bool.	25,597 (27%)			
Verified	Bool.	244 (0.3%)			
Topic variables	Bool.	2,818 (3%)			

Table 4: Chi square test

	X^2	Mean(F)	Mean(M)	df	p-value
Has URL	112.73	0.23	0.27	1	***
Has image	30.89	0.984	0.976	1	***
Has location	55.64	0.90	0.92	1	***
Has description	0.95	0.96	0.95	1	0.33
Verified	26.41	0.0017	0.0043	1	***
Note:			*p<0.05; **p<0.01; ***p<0.001		

URLs, images, location in their profiles, and be verified users. The difference in the probability of having a description is not significant.

Twitter Activity. Figure 2 shows CCDF plots for variables that exhibit heavy-tailed distributions, including followers, listed, retweeted, tweets, and friends counts. Although the differences between male and female users is small, the CCDFs of male users tend to be above those of female users, indicating that, at least before controlling other variables, perceived males have a better presence among the 'elite' users with high visibility in Twitter.

Table 5: Mann-Whitney *U* test

	U	Med(F)	Med(M)	p-value
Followers	354,197,701	200	163	***
Listed (all)	331,010,000	0	0	0.34
Listed (exclude 0s)	75,005,000	3	2	***
Retweeted (all)	338,390,000	0	0	***
Retweeted (exclude 0s)	24,803,000	2	2	0.24
Note:			*$p < 0.05$ **$p < 0.01$ ***$p < 0.001$	

Since the distributions of these variables are highly skewed and exhibit heavy tails, we employ the Mann-Whitney *U* test. Although perceived male users tend to have older accounts, the median number of tweets for female users is significantly higher than for male users. The median number of friends for female users is significantly higher than for male users (241 vs. 220). Most importantly, the results show that perceived female users have significantly *more* followers than male users (200 vs. 163), and are listed more (when excluding users who are not listed at all). We find that female and male users have a statistically significant difference for retweets even when both have a median of zero. However, this difference goes away if we exclude those users with zero retweets.

Topics. Figure 3 shows the gender distribution for the 20 topics that we extracted, with log-ratio of perceived male to female. It shows that the gender differences between topics varies greatly although there are more male users with topic labels in all cases.

Chi-square tests show that male and female users are tagged to different topics ($p < 0.001$). Topics that reflect a male-dominated context such as sports, companies, and technology exhibit larger differences.

Summary. In sum, we find statistically significant differences between male and female users, in terms of self-presentation, activity, as well as how much they are followed, retweeted, and listed. Female users tweet more, follow more, have younger accounts, are more likely to have profile images, and to be retweeted and listed. At the same time, female users are less likely to be verified and to provide a location or URL in their profiles.

Quartile Regression and Glass Ceiling

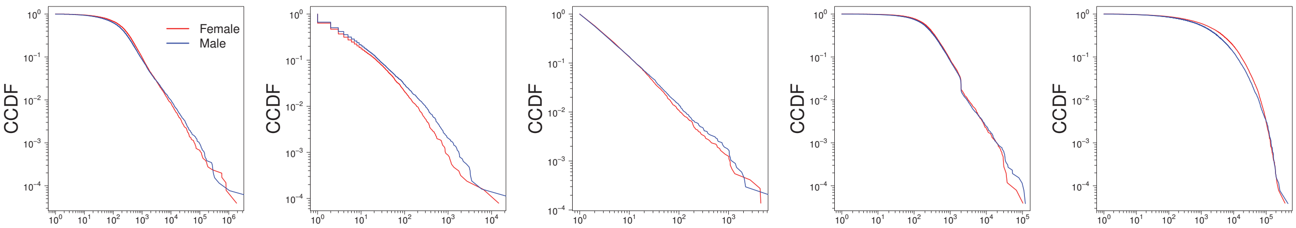
Our exploratory analysis showed that, contrary to our initial hypothesis, perceived female users have a higher median number of followers and are added to more lists (when excluding users who are not listed). Yet, this provides an incomplete picture because (i) we did not include control variables and (ii) our analysis was done for the whole dataset and overlooks the different effects across levels of visibility, from newcomers with few followers to celebrities with millions of followers.

Applying multivariate statistical models on the whole population as well as quartiles for each dependent variable better isolates the association between gender and visibility. Although we include control variables in all models, for brevity, we omit them from the result tables but full tables are available upon request. We add followers as a control for the retweet and list count models because more followers may result in being retweeted and listed more. In this paper we report Incident Rate Ratios (IRR), which are the exponentiated coefficients of Poisson regressions. This allows us to compare the rates of followers, retweets, and lists between perceived male users and perceived female users.

Table 6 shows the results of Poisson regression for the follower counts and retweeted counts comparing female vs. male. The first column shows the result for the entire sample such that male users have significantly more followers and retweets compared with female users ($p < 0.001$). This finding holds even when other gender categories are included in the analyses ($p < 0.001$).[3] Therefore, we find support for both Hypothesis 1 and Hypothesis 3, that perceived female

[3]For brevity, we do not present the tables for female vs. others.

Figure 2: CCDF plots for variables with power-law distribution



(a) Followers (b) Listed (c) Retweeted (d) Friends (e) Tweets

Figure 3: Gender distribution for top 20 popular topics

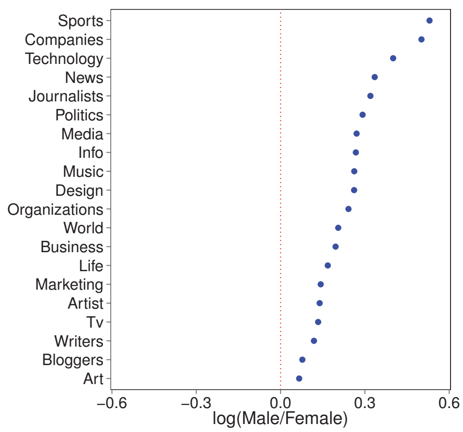


Table 6: Female vs. Male Poisson Regressions

	Poisson	0.25 Qnt.	0.5 Qnt.	0.75 Qnt.	1.00 Qnt.
	Dependent variable: Number of Followers				
Male	0.69***	−0.07***	−0.01***	−0.01***	0.81***
Male (IRRs)	2.00***	0.93***	0.99***	0.98***	2.24***
Observations	51,347	13,850	13,693	12,854	10,950
	Dependent variable: Retweeted count				
Male (Coeffs)	0.03***	NULL	NULL	−0.021	0.15***
Male (IRRs)	1.03***	NULL	NULL	0.98	1.16***
Observations	51,347	14822	13503	12,533	10,489
Note:					*p<0.05; **p<0.01; ***p<0.001

Twitter users are less likely to be followed or retweeted compared to male users or gender ambiguous users.

Quartile regression reveals that the effect of perceived gender changes as one moves from the least visible to most visible users. Since being retweeted is uncommon, the first two quartiles consist of accounts not retweeted at all. Thus, we show results for the last two quartiles of retweeted counts. While being perceived female is a weak indicator of followers and retweets in the first three quartiles, in the last quartile the effect changes direction. For followers, users perceived as male are expected to have an incidence rate 2.24 times that of female users (a 124% increase) and an incidence rate for retweets 1.16 times that of female users (a 16% increase) ($p < 0.001$).

We find similar results when we compare female users with male, gender ambiguous, and unknown users. Across the first three quartiles gender is a weak indicator of followers and retweets. However, in the fourth quartile male users, gender ambiguous users, and unknown users have statistically significant higher followers and retweet counts than female users ($p < 0.001$). This finding suggests that female users face a steeper climb to gaining a large following than male users, gender ambiguous users, and even users that do not provide a profile image or name. However, gender ambiguous or unknown users may be organizational accounts that tend to have more followers than individuals (McCorriston, Jurgens, and Ruths 2015).

Analysis of the listed count measure show that male users are listed fewer times than their female counterparts. This result holds both for the entire population and for the quartile regression. Thus, for this measure of visibility, our second and fourth hypotheses are rejected. However, listed count is the *only exception* to our hypothesized relationship between perceived gender and measures of visibility across tiers. When examining users' visibility across tiers, we find support for our hypotheses overall, with perceived female users being penalized at the highest tier for both number of followers and retweets.

Robustness of Results

Our results remain hold whether we employ either image based or name based gender inference methods and regardless of whether topic controls are added. Our results hold when we use OLS and negative binomial models, although some of these models do not converge. The results from these models were omitted for brevity but they can be provided upon request. Negative binomial regression on quartiles provide similar results, however the models did not converge when run on the entire population. A zero-inflated negative binomial regression model can be fitted for the whole population but is inappropriate for quartiles because the first two quartiles include nearly all zero values and the last two almost no zero values. We did not apply corrections for multiple testing (e.g. Bonferroni) because it would be inappropriate given the robust and consistent results across the different regression models (Perneger 1998).

Discussion

Our investigation has the following limitations. First, our methods may not capture all the ways that gender is 'perceived' or inferred by others on Twitter. For instance, we do not account for tweet content which could strongly signal gender (Cunha et al. 2012). Future studies may consider how differences in linguistic patterns such as hedging or the use of pronouns may impact users' inferred gender and visibility (Hemphill and Otterbacher 2012). Also, the U.S. Census-based method might miss names from other cultures and thus produce less accurate results for certain names. Although, the name based method displayed good agreement with the image based method, our results are more likely to be valid for the U.S. and other English-speaking countries. Second, our study is observational and cannot establish causal relationships. Third, although we try to control for 'offline fame' with the verified variable, it is difficult to distinguish the effect of judgemental biases from population bias. Our results may be explained by gender framing online interaction or by the replication of existing *offline* gender disparities. Fourth, the generality of our results, which are from Twitter, across different social media platforms is yet to be seen (Tufekcioglu 2014). Finally, our study focuses on 'visibility' and does not investigate other facets of online interactions. Our visibility measures do not speak to the content of communication, the types of tweets that are retweeted by female and male users, or the future implications of visibility. Investigation of these topics may further reveal how perceived gender impacts online social interaction.

Studies focusing explicitly on individual decisions can provide greater insight into the role that various social categories play in online communication and inequality. Future studies should examine the extent to which users misrepresent themselves, adapt behaviors to remain gender ambiguous, or conform to normative gender expectations. This paper takes a first, but important, step towards connecting "action-identity" theory to users online visibility.

By doing so, we highlight why scholarship must consider the situational meanings attached to users' self-presentation. While users' online identity might not be real, the consequences of their perceived gender are (Thomas and Thomas 1928). Gender stereotypes may be activated when social media users are perceived as male or female. Thus, users perceived as female may be met with offline gender stereotypes that often work in men's favor, with men assumed to be more competent and worthier of status (Glick et al. 2004). Gender inferred from user profiles may frame interaction and if these frames become consensual in blogging communities they have the power to perpetuate existing offline inequalities online (Johnson, Dowd, and Ridgeway 2006).

Our results suggest that the gender disparity in Twitter visibility is present at the top, where there are many offline celebrities and already-famous people. Therefore, one explanation may be that the most visible users 'bring' the gender disparity from offline to Twitter. While this disparity may be driven by a small number of 'elite' users who account for much of the total attention on Twitter (Wu et al. 2011), the fact remains that, at the highest levels of visibility, users perceived to be male come out on top.

Conclusions

This paper explores the relationship between Twitter users' perceived gender and the attention they attract. By analyzing a sample of more than 94,000 Twitter accounts with multivariate regression models, we find that the relationship between users' perceived gender and visibility displays both (i) an overall weak disadvantage towards perceived female users and (ii) a 'glass ceiling' effect, where perceived female users have a strong disadvantage in visibility only in the highest quartile. Our results suggest that gender or the social categories inferred from users' self-presentations may frame communication and even allow gender inequality to persist online.

Acknowledgments

This material is based upon work supported by the National Science Foundation under grant CNS-1252697. This work was also partially funded by the Office of the Vice Provost for Research at Indiana University Bloomington through the Faculty Research Support Program, and the Department of Sociology at Indiana University. Yong-Yeol Ahn was supported by a Microsoft Research Faculty Fellowship.

References

Anderson, A.; Huttenlocher, D.; Kleinberg, J.; and Leskovec, J. 2012. Effects of User Similarity in Social Media. In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining*, 703–712. ACM.

Antin, J.; Yee, R.; Cheshire, C.; and Nov, O. 2011. Gender Differences in Wikipedia Editing. In *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, 11–14. ACM.

Aral, S., and Walker, D. 2012. Identifying Influential and Susceptible Members of Social Networks. *Science* 337(6092):337–341.

Bimber, B. 2000. Measuring the Gender Gap on the Internet. *Social Science Quarterly* 868–876.

Brown, P. E., and Feng, J. 2011. Measuring User Influence on Twitter Using Modified K-Shell Decomposition. In *Fifth International AAAI Conference on Weblogs and Social Media*.

Carter, D. 2005. Living in Virtual Communities: An Ethnography of Human Relationships in Cyberspace. *Information, Community & Society* 8(2):148–167.

Center, T. W. M. 2013. Social Security Administration.

Cha, M.; Haddadi, H.; Benevenuto, F.; and Gummadi, P. K. 2010. Measuring user influence in Twitter: The million follower fallacy. *ICWSM* 10(10-17):30.

Chang, J.; Rosenn, I.; Backstrom, L.; and Marlow, C. 2010. ePluribus: Ethnicity on Social Networks. *ICWSM* 10:18–25.

Correll, S. J., and Benard, S. 2007. Getting a Job: Is there a Motherhood Penalty? *American Journal of Sociology* 112(5):1297–1339.

Cotter, D. A.; Hermsen, J. M.; Ovadia, S.; and Vanneman, R. 2001. The Glass Ceiling Effect. *Social Forces* 80(2):655–681.

Cunha, E.; Magno, G.; Almeida, V.; Gonçalves, M. A.; and Benevenuto, F. 2012. A Gender Based Study of Tagging Behavior in Twitter. In *Proceedings of the 23rd ACM Conference on Hypertext and Social Media*, 323–324. ACM.

Duggan, M., and Smith, A. 2013. Social Media Update 2013. *Pew Internet and American Life Project*.

Eagly, A. H., and Karau, S. J. 2002. Role Congruity Theory of Prejudice toward Female Leaders. *Psychological Review* 109(3):573.

Fan, H.; Cao, Z.; Jiang, Y.; Yin, Q.; and Doudou, C. 2014. Learning Deep Face Representation. *arXiv Preprint arXiv:1403.2802*.

Ferrara, E.; Varol, O.; Davis, C.; Menczer, F.; and Flammini, A. 2014. The Rise of Social Bots. *arXiv Preprint arXiv:1407.5225*.

Fiske, S. T.; Cuddy, A. J.; Glick, P.; and Xu, J. 2002. A Model of (often mixed) Stereotype Content: Competence and Warmth Respectively Follow from Perceived Status and Competition. *Journal of Personality and Social psychology* 82(6):878.

Garcia, D.; Weber, I.; and Garimella, V. R. K. 2014. Gender Asymmetries in Reality and Fiction: The Bechdel Test of Social Media. *arXiv Preprint arXiv:1404.0163*.

Glick, P.; Lameiras, M.; Fiske, S. T.; Eckes, T.; Masser, B.; Volpato, C.; Manganelli, A. M.; Pek, J. C.; Huang, L.-l.; Sakalli-Uğurlu, N.; et al. 2004. Bad but Bold: Ambivalent Attitudes toward Men Predict Gender Inequality in 16 Nations. *Journal of Personality and Social Psychology* 86(5):713.

Greenwald, A. G.; McGhee, D. E.; and Schwartz, J. L. 1998. Measuring Individual Differences in Implicit Cognition: the Implicit Association Test. *Journal of Personality and Social Psychology* 74(6):1464.

Habermas, J. 1991. *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*. MIT Press.

Hemphill, L., and Otterbacher, J. 2012. Learning the Lingo?: Gender, Prestige and Linguistic Adaptation in Review Communities. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, 305–314. ACM.

Hong, L.; Dan, O.; and Davison, B. D. 2011. Predicting popular messages in Twitter. In *Proceedings of the 20th International Conference Companion on World Wide Web*, 57–58. ACM.

Java, A.; Song, X.; Finin, T.; and Tseng, B. 2007. Why we Twitter: Understanding Microblogging Usage and Communities. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*, 56–65. ACM.

Johnson, C.; Dowd, T. J.; and Ridgeway, C. L. 2006. Legitimacy as a Social Process. *Annual Review of Sociology* 53–78.

Jürgens, P.; Jungherr, A.; and Schoen, H. 2011. Small Worlds with a Difference: New Gatekeepers and the Filtering of Political Information on Twitter. In *Proceedings of the 3rd International Web Science Conference*, WebSci '11. ACM.

Kim, K.-S.; Sin, S.-C. J.; and He, Y. 2013. Information Seeking Through Social Media: Impact of User Characteristics on Social Media Use. *Proceedings of the American Society for Information Science and Technology* 50(1):1–4.

Kim, H.-W.; Zheng, J. R.; and Gupta, S. 2011. Examining Knowledge Contribution from the Perspective of an Online Identity in Blogging Communities. *Computers in Human Behavior* 27(5):1760–1770.

Kivran-Swaine, F.; Brody, S.; Diakopoulos, N.; and Naaman, M. 2012. Of Joy and Gender: Emotional Expression in Online Social Networks. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work Companion*, 139–142. ACM.

Kwak, H.; Lee, C.; Park, H.; and Moon, S. 2010. What is Twitter, a Social Network or a News Media? In *Proceedings of the 19th International Conference on World Wide Web*, 591–600. ACM.

Lam, S. T. K.; Uduwage, A.; Dong, Z.; Sen, S.; Musicant, D. R.; Terveen, L.; and Riedl, J. 2011. WP: Clubhouse?: An Exploration of Wikipedia's Gender Imbalance. In *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, 1–10. ACM.

Magno, G., and Weber, I. 2014. International Gender Differences and Gaps in Online Social Networks. In *Social Informatics*. Springer. 121–138.

Marwick, A., et al. 2011. To See and be Seen: Celebrity Practice on Twitter. *Convergence: the international journal of research into new media technologies* 17(2):139–158.

McCorriston, J.; Jurgens, D.; and Ruths, D. 2015. Organizations Are Users Too: Characterizing and Detecting the Presence of Organizations on Twitter. In *Ninth International AAAI Conference on Web and Social Media*.

McKelvey, K. R., and Menczer, F. 2013. Truthy: Enabling the Study of Online Social Networks. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work Companion*, CSCW '13, 23–26. ACM.

Merluzzi, J., and Dobrev, S. D. 2015. Unequal on Top: Gender Profiling and the Income Gap among High Earner Male and Female Professionals. *Social Science Research* 53:45–58.

Messias, J.; Schmidt, L.; Oliveira, R.; and Benevenuto, F. 2013. You Followed My Bot! Transforming Robots into Influential Users in Twitter. *First Monday* 18(7).

Mislove, A.; Lehmann, S.; Ahn, Y.-Y.; Onnela, J.-P.; and Rosenquist, J. N. 2011. Understanding the Demographics of Twitter Users. *ICWSM* 11:5th.

Morales, A.; Borondo, J.; Losada, J.; and Benito, R. 2014. Efficiency of Human Activity on Information Spreading on Twitter. *Social Networks* 39:1–11.

Oakley, J. G. 2000. Gender-based Barriers to Senior Management Positions: Understanding the Scarcity of Female CEOs. *Journal of Business Ethics* 27(4):321–334.

Otterbacher, J. 2010. Inferring Gender of Movie Reviewers: Exploiting Writing Style, Content and Metadata. In *Pro-*

ceedings of the 19th ACM international Conference on Information and Knowledge Management, 369–378. ACM.

Ottoni, R.; Pesce, J. P.; Las Casas, D. B.; Franciscani Jr, G.; Meira Jr, W.; Kumaraguru, P.; and Almeida, V. 2013. Ladies First: Analyzing Gender Roles and Behaviors in Pinterest. In *ICWSM*.

Padavic, I., and Reskin, B. F. 2002. *Women and Men at Work*. Pine Forge Press.

Perneger, T. V. 1998. Whats Wrong with Bonferroni Adjustments. *BMJ: British Medical Journal* 316(7139):1236.

Quercia, D.; Las Casas, D. B.; Pesce, J. P.; Stillwell, D.; Kosinski, M.; Almeida, V.; and Crowcroft, J. 2012. Facebook and Privacy: The Balancing Act of Personality, Gender, and Relationship Currency. In *ICWSM*.

Reagle, J., and Rhue, L. 2011. Gender Bias in Wikipedia and Britannica. *International Journal of Communication* 5:21.

Reskin, B. F., and Ross, C. E. 1992. Jobs, Authority, and Earnings among Managers the Continuing Significance of Sex. *Work and Occupations* 19(4):342–365.

Ridgeway, C. L. 2001. Gender, Status, and Leadership. *Journal of Social Issues* 57(4):637–655.

Romero, D. M.; Galuba, W.; Asur, S.; and Huberman, B. A. 2011. Influence and Passivity in Social Media. In *Machine Learning and Knowledge Discovery in Databases*. Springer. 18–33.

Seward, Z. M. 2014. Twitter Admits that as Many as 23 Million of its Active Users are Automated.

Sharma, N. K.; Ghosh, S.; Benevenuto, F.; Ganguly, N.; and Gummadi, K. 2012. Inferring Who-is-who in the Twitter Social Network. *ACM SIGCOMM Computer Communication Review* 42(4):533–538.

Sugimoto, C. R.; Lariviere, V.; Ni, C.; Gingras, Y.; and Cronin, B. 2013. Global Gender Disparities in Science. *Nature* 504(7479):211–213.

Sugimoto, C. R.; Ni, C.; and Larivière, V. 2015. On the Relationship between Gender Disparities in Scholarly Communication and Country-level Development Indicators. *Science and Public Policy* 42(6):789–810.

Suh, B.; Hong, L.; Pirolli, P.; and Chi, E. H. 2010. Want to be Retweeted? Large Scale Analytics on Factors Impacting Retweet in Twitter Network. In *Social Computing (socialcom), 2010 IEEE Second International Conference on*, 177–184. IEEE.

Taigman, Y.; Yang, M.; Ranzato, M.; and Wolf, L. 2014. Deepface: Closing the Gap to Human-level Performance in Face Verification. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, 1701–1708. IEEE.

Team, M. B. 2015. Fun with ML, Stream Analytics and PowerBI? Observing Virality in Real Time.

Thelwall, M. 2008. Social Networks, Gender, and Friending: An Analysis of Myspace Member Profiles. *Journal of the American Society for Information Science and Technology* 59(8):1321–1330.

Thomas, W., and Thomas, D. S. 1928. *The Child in America; Behavior Problems and Programs*. A. A. Knopf.

Thomas, K.; Grier, C.; Song, D.; and Paxson, V. 2011. Suspended Accounts in Retrospect: An Analysis of Twitter Spam. In *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, 243–258. ACM.

Tilly, C. 1999. *Durable Inequality*. Univ of California Press.

Tomaskovic-Devey, D.; Zimmer, C.; Stainback, K.; Robinson, C.; Taylor, T.; and McTague, T. 2006. Documenting Desegregation: Segregation in American Workplaces by Race, Ethnicity, and Sex, 1966–2003. *American Sociological Review* 71(4):565–588.

Tufekcioglu, Z. 2014. Big Questions for Social Media Big Data. In *The AAAI Press*.

Velichety, S., and Ram, S. 2013. Examining Lists on Twitter to Uncover Relationships between Following, Membership and Subscription. In *Proceedings of the 22nd international conference on World Wide Web companion*, 673–676. International World Wide Web Conferences Steering Committee.

Vergeer, M., and Hermans, L. 2013. Campaigning on Twitter: Microblogging and Online Social Networking as Campaign Tools in the 2010 General Elections in the Netherlands. *Journal of Computer-Mediated Communication* 18(4):399–419.

Wagner, C.; Garcia, D.; Jadidi, M.; and Strohmaier, M. 2015. It's a Man's Wikipedia? Assessing Gender Inequality in an Online Encyclopedia. *arXiv Preprint arXiv:1501.06307*.

Ward, K. B.; Gast, J.; and Grant, L. 1992. Visibility and Dissemination of Womens and Mens Sociological Scholarship. *Social Problems* 39(3):291–298.

Weng, J.; Lim, E.-P.; Jiang, J.; and He, Q. 2010. Twitterrank: Finding Topic-sensitive Influential Twitterers. In *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, 261–270. ACM.

Williams, J. C. 2005. The Glass Ceiling and the Maternal Wall in Academia. *New Directions for Higher Education* 2005(130):91–105.

Willis, J., and Todorov, A. 2006. First Impressions Making Up Your Mind After a 100-ms Exposure to a Face. *Psychological Science* 17(7):592–598.

Wu, S.; Hofman, J. M.; Mason, W. A.; and Watts, D. J. 2011. Who Says What to Whom on Twitter. In *Proceedings of the 20th International Conference on World Wide Web*, 705–714. ACM.

Yu, K.; Lu, Z.; and Stander, J. 2003. Quantile Regression: Applications and Current Research areas. *Journal of the Royal Statistical Society: Series D (The Statistician)* 52(3):331–350.

Zaman, T. R.; Herbrich, R.; Van Gael, J.; and Stern, D. 2010. Predicting Information Spreading in Twitter. In *Workshop on Computational Social Science and the Wisdom of Crowds, Nips*, volume 104, 17599–601. Citeseer.

Zhou, E.; Cao, Z.; and Yin, Q. 2015. Naive-Deep Face Recognition: Touching the Limit of LFW Benchmark or Not? *arXiv Preprint arXiv:1501.04690*.