Unlocking Unstructured Social Media Data in Marketing

William Rand

Assistant Professor of Bussiness Management

In Collaboration with

Kelly Hewett, Roland Rust, and Harald J. van Heerde

NC STATE

Poole College of Management

MENU

DEALBREAKER

NEWS

Bank of America Will Steal Your Parrot And Then Make *You* Look Like The Crazy One

By BESS LEVIN

21 Comments / Mar 10, 2010 at 6:15 PM

No friends of Ken Lewis In its storied history, Ba
foreclosed on a few hous
of deadbeats out there and really no good central
homes are supposed to padlocked and whose are
really something we can hold against them. Wha
however, is, for instance, foreclosing on someone
confiscating his/her beloved parrot. Except that
America really did that.

THE WALL STREET JOURNAL.

ess Tech Markets Opinion Arts Life

king Parrot

as Vacant, Padlocked It and

Privacy and cookies | Jobs | Dating | Offers | Shop | Puzzles | Investor | Log In | Register | Subscribe

The Telegraph

HOME » NEWS » NEWS TOPICS » HOW ABOUT THAT?

Bank of America apologises for repossessing parr

Bank of America apologised to a woman after one of its contractor
trashed her house and took her parrot while wrongly repossessing

166 0 0 166 Email

Photo: REUTERS

4:37AM GMT 10 Mar 2010

Forty-six-year-old Angela Iannelli sued the bank on Monday.

She claims her mortgage was up-to-date when one of the banking giant's contractors damaged furniture, took her pet parrot, Luke, and padlocked her door in October.

In a statement, the bank said it "sincerely apologises" and has tried for months to resolve the issue.

The bank said it has "zero tolerance for this kind of error" and said it will

Troy McComas troy48 @tlm1948 · 28 Nov 2012
Woman loses **parrot**, has house ransacked by **Bank of America** in mortgage **foreclosure** mix-up: lawsuit - Bank of... tmblr.co/Z9AlVxYGuJrh
Expand

Daily Legal News @DailyLegalNews · 12 Mar 2010
Bank of America faces 50K lawsuit after stealing woman's **parrot**, causing other damage, in mistaken **foreclosure** raid http://uslaw.com/-/r6Cto
Expand

Alpha Dog @Alpha_Dog2001 · 12 Mar 2010
www.nydailynews.com/news/national/2010/03/11/2010-03-11_womans_**parrot**_taken_house_ransacked_by_**bank_of_america**_in_**foreclosure**_mixup.html
Expand

New York Daily News @NYDailyNews · 11 Mar 2010
Woman's **parrot** taken, house ransacked by **Bank of America**: lawsuit - A bird in the hands of a **foreclosure** busines... http://su.pr/2wzR90
Expand

Eric Milgram @EricMilgram · 11 Mar 2010
"**Bank of America** Steals Pet **Parrot** from Wrongly Foreclosed Home - #Gawker" (http://bit.ly/9Ou6ph) #BoA #**Foreclosure**
Expand

Managers' perspectives

*"Media is an **echo chamber** . . . you have **advertising** that is perceived one way, **news** perceived another way, and **social media** perceived yet another way and they are different durations."*

Marketing executive in a Fortune 50 FSI

The “echoverse”

Firm Actions

Consumer Sentiment

Social Media Word of Mouth

News Media

Business Performance

Industry setting: Top 4 US financial services firms

June 2007 – December 2013

Data

- Firm Communications:
  - All press releases volume and valence
  - All corporate Tweets (volume and valence)
  - Ad spending from Ad$pender (Kantar)
- Traditional Media News Stories (volume and valence)
- Consumer Sentiment – YouGov's **BrandIndex**
- Social Media WOM – Tweets about firms (volume and valence)
- Firm Performance – consumer deposits

Textual data overview

- Size of Collections
  - Press releases (n = 5,376)
    - All of the press releases from the 4 banks over the time period
  - News Articles (n = 65,261)
    - Provided by Factiva
    - Drawn from 14 of the largest circulation newspapers in the US
  - Tweets (n = 18,577,733)
    - Gathered using Gnip queries regarding the 4 banks
- Coding of Data
  - Manual Coding is impossible
    - Volume and Variety of Data
  - Data coding must be done automatically using state of the art computational linguistics

Automated data extraction

- Conversion of all data from either Microsoft Word documents or PDFs to plain text, and preprocessing of Twitter data
- Data extracted from each document:
  - Date
  - Associated bank(s) – source of the press release, banks mentioned in articles and Tweets
  - Sentiment
- Aggregated data at the week level for each dataset

Twitter data collection

- Both from the bank and about the bank
- Used 72 different Power Track Searches from Gnip
  - 27 BofA, 22 Chase, 15 Wells Fargo, 18 Citibank
- Examples
  - Different Bank Twitter Accounts and Mentions of those Accounts
    - @BofA_News, @BofA_Help, @BofA_Careers, @BofA_Community, @BofA_Tips, @bankofamerica, @MerrillLynch
  - References to the Bank
    - BAC and (bank or banking), BOA and (bank or banking), “Bank of America”, BofA
  - References to Executives
    - Ken Lewis and (CEO or BAC or BOA), Moynihan (Brian or CEO or BAC or BOA)
  - References to Products
    - (BAC or BOA) and loan, (BAC or BOA) and ATM, (BAC or BOA) and mortgage
- All terms were chosen to minimize the probability of false positives but to gather as much as possible (maximizing recall and precision)

Press release and news article sentiment analysis

- Based on LIWC (Linguistic Inquiry and Word Count; Pennebaker et al. 2007)
- Contains a Dictionary of Positive and Negative Emotion Terms, including stems, e.g., ugly and ugliness map to ugl*
- Each target word in a document is examined and see if it matches any of the terms in the dictionaries
- Outputs a percentage of positive and negative words and a word count

Twitter sentiment analysis

- Long Form Sentiment Tools, such as LIWC, do not work for Tweets
  - Misspellings, e.g., “happniess”, “citybank”
  - Acronyms, e.g., LULZ, IMHO
  - Dropped characters, e.g., “hppy”, “
- We used a 10-fold, cross-validated, Naïve Bayes Classifier (Go, Bhayani, and Huang, 2009)
  - Trained on tweets that contain emoticons, e.g., :) and :(
  - Tweets are preprocessed
    - Removed whitespaces
    - Added a feature for all CAPS
    - Lowercase everything
    - Remove punctuation at beginning and end of words, but keep emoticons
    - Reduce repeated letters, e.g., sooooo becomes so
    - Remove common stop words
  - In the end wind up with ~8000 features, e.g., common words, timestamps, links
  - Output is -1 to 1, measuring how positive or negative the tweet is

Creation of weekly data

- Twitter Data
  - Valence: Sum of the sentiment scores associated with tweets by bank by week
  - Volume: Count of the number of tweets by bank by week
  - Separate out Tweets from the Bank vs. Tweets by others
- Press Releases and News Articles
  - Valence: Sum of percentage of positive words minus the sum of percentage of negative words by bank by week
  - Volume: Count of the number of documents by bank by week

Consumer Sentiment

- Data from BrandIndex database (YouGov)

- Weekly measure of YouGov's overall buzz metric:
  - Which of the following banks are in the news?
  - Are they in the news in a positive or a negative way?
  - Average across thousands of responses

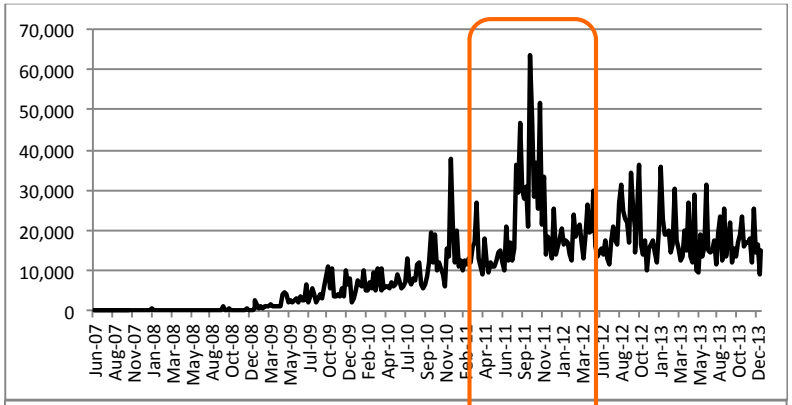
What we observed

The echoverse at work

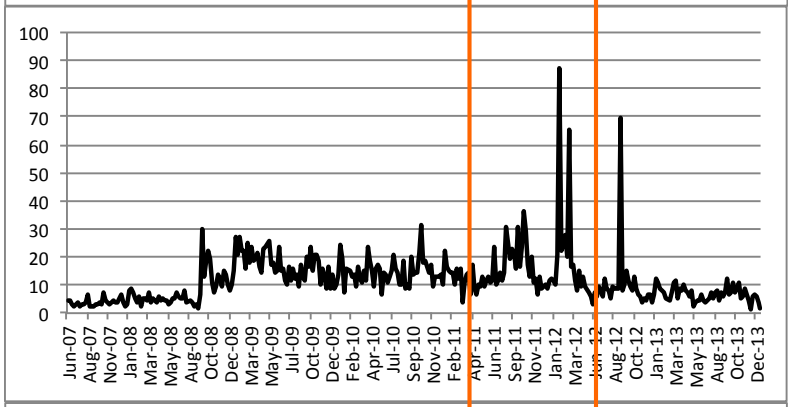
The case of

Bank of America

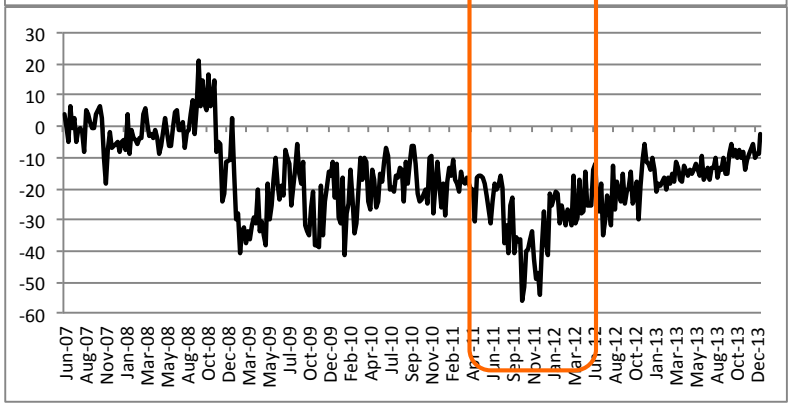
Twitter volume



News article volume



Consumer sentiment



June 2007 – December 2013

The brand echoverse

Online Word of Mouth: Twitter posts

Firm Communications:
- Press Releases
- Advertising
- Twitter posts

Business Outcome: Customer deposits

Traditional Media News Stories

Consumer Sentiment BrandIndex

Statistical Method

- Echoverse conceptualization says all variables are endogenous
- Time lags exist - 1 lag is optimal
- Appropriate methodology is VAR Model: Vector AutoRegressive Model

- Calculate Generalized Impulse Response Functions:
  - Shock one variable by 1 standard deviation
  - Simulate effect on all variables in system

The Estimation Equations

$$\begin{pmatrix} \text{News Article Volume} \\ \text{New Article Valence} \\ \text{Consumer Sentiment} \\ \text{Public TwitterVolume} \\ \text{Public Twitter Valence} \\ \text{Company TwitterVolume} \\ \text{Company Twitter Valence} \\ \text{Press Release Volume} \\ \text{Press Release Valence} \\ \text{Advertising Spend} \\ \Delta\text{Customer Deposits} \end{pmatrix}_{\text{bank } b, \text{week } t} = \sum_{l=1}^{L} B_{bl} \begin{pmatrix} \text{News Article Volume} \\ \text{New Article Valence} \\ \text{Consumer Sentiment} \\ \text{Public TwitterVolume} \\ \text{Public Twitter Valence} \\ \text{Company TwitterVolume} \\ \text{Company Twitter Valence} \\ \text{Press Release Volume} \\ \text{Press Release Valence} \\ \text{Advertising Spend} \\ \Delta\text{Customer Deposits} \end{pmatrix}_{\text{bank } b, \text{week } t-l} + Controls + e_{bt}$$

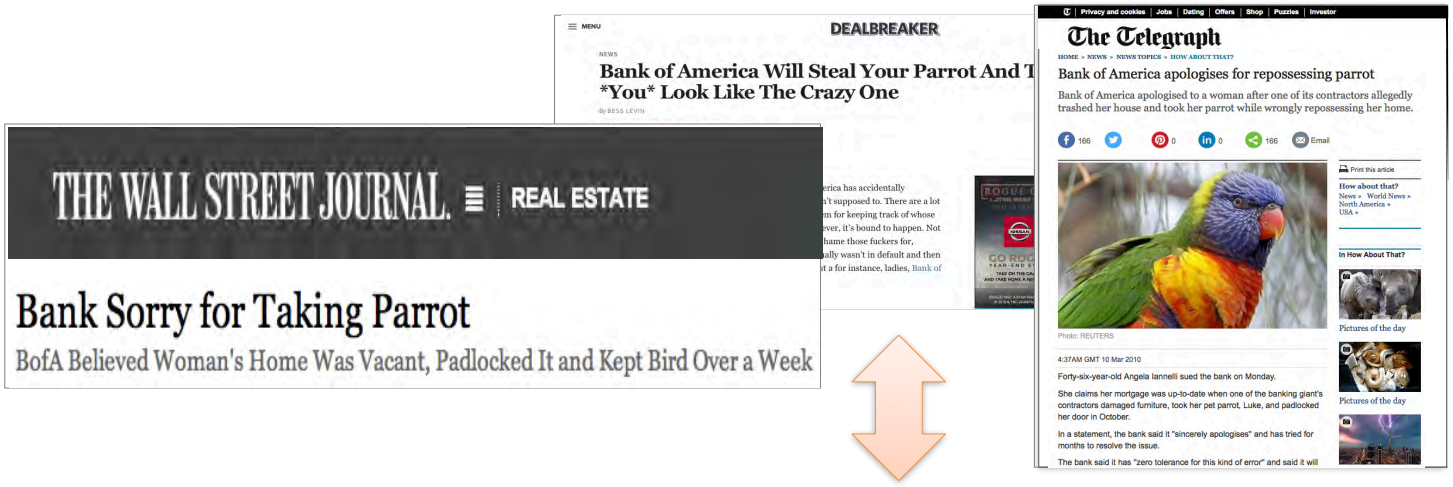
Controls: Fixed effects for Banks, Competitor Actions, Consumer Confidence, Unemployment Rate; Interest Rate, Time trend

First, the bad news . . . .

1) Bad news spreads fast and wide

2) Online word of mouth hurts firm performance (customer deposits)

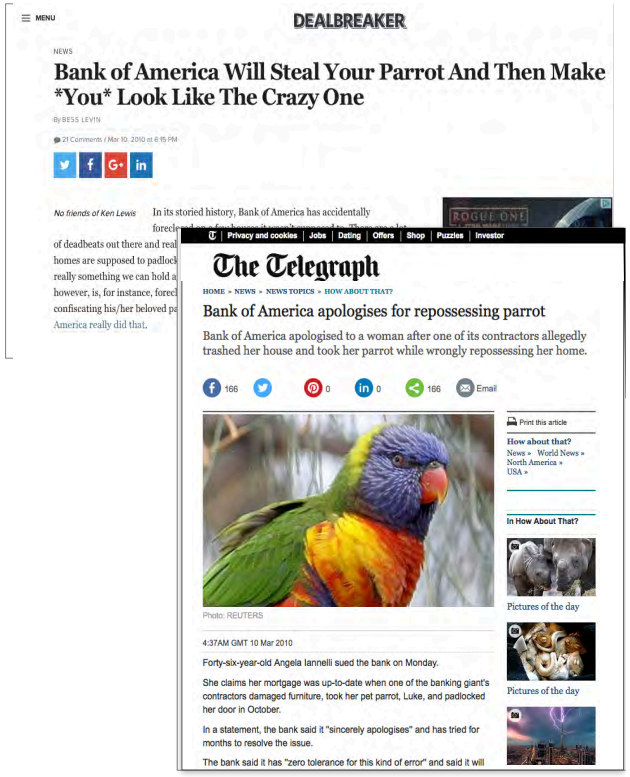
1) Bad news spreads fast and wide.



A. ***More negative news articles lead to more public tweets.***

B. ***More negative public tweets lead to more news articles.***

C. More negative consumer sentiment



&

D. More negative company tweets

More news articles

2) Online word of mouth hurts firm performance.

What can companies do?

The good news . . .

A) Press releases are surprisingly effective. . . .

More press releases

1

More/more positive news articles

More/more positive word of mouth

2

3

Higher customer deposits

BUT . . . press release <u>strategy</u> matters.

Bank of America

Let's lay low and watch.

BAD IDEA!

Let's keep our name out there

WELLS FARGO

GOOD IDEA!

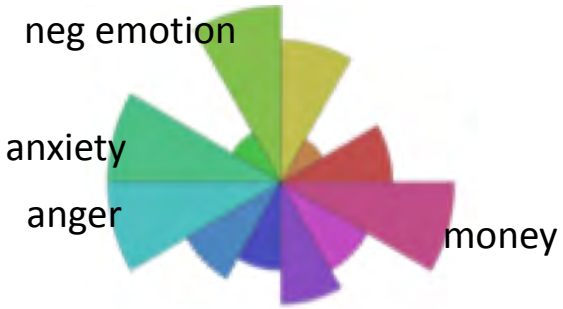
AND . . . press release <u>language</u> matters.

Categories of words used in press releases:

Differences in press release language emerged.

Bank of America

PRESS RELEASE

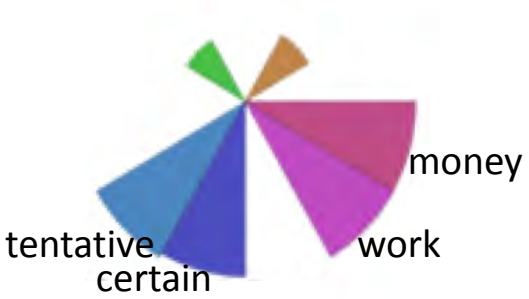


*Tendency to **react** to negativity in the media*

BAD IDEA!

WELLS FARGO

PRESS RELEASE



*Tendency to **ignore** negativity in the media*

GOOD IDEA!

B) Company tweets can calm things down.

More company tweets

Bank of America @BankofAmerica · Jan 5
This morning @MeredithVerdone will be talking about innovation in service of our customers @shingy @brynmooser @sherrillkaplan #ces2017

Bank of America Retweeted

Jeff Barker @JeffBarkerNY · 24 Dec 2016
Glad to see one of our teammates featured in @TheAdvocateMag bit.ly/2igH8OP #LGBT #equality. Congrats on a great article!

Bank of America @BankofAmerica · Jan 4
Looking forward to watching #Atlanta area nonprofits' continued success in 2017. Thanks @PhilWHudson for helping spread the word!

> **AtlBizChron** @AtlBizChron
> Bank of America awards more than $4M to #Atlanta #nonprofits in 2016 $BAC bizjournals.com/atlanta/news/2...

Bank of America @BankofAmerica · Jan 1
Here's to a 2017 filled with new experiences, new adventures and new memories you'll cherish for years.

Fewer news articles

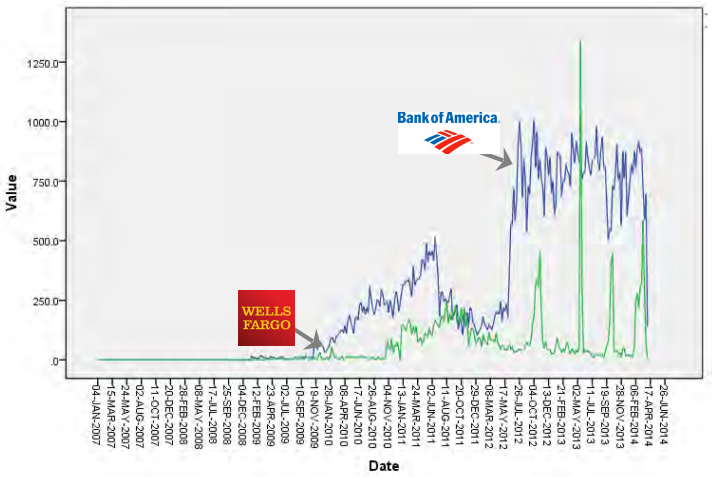
Politics101 @NBfromLB · 23 Dec 2016
Who else has had issues with @BankofAmerica would appreciate hearing from you #BankOfAmerica

Miriam Ogembo @miridee21 · 30 Dec 2016
Anyone planning on banking with #bankofamerica RUN don't do it. They will never help you when there's a fraud committed to your account

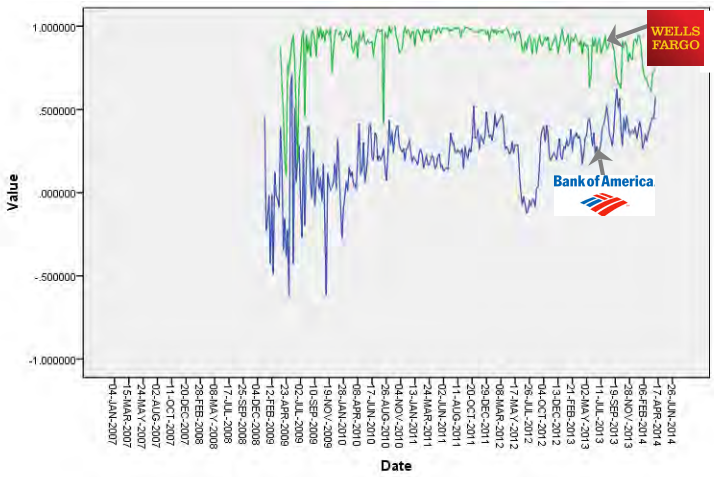
Valleygurl_T @passport2hauteT · 21 Dec 2016
Per @NBCNews #BankofAmerica #Chase & #WellsFargo all have increased #bankfees with an avg. transaction of $4.50 for using #ATM outside yours

BUT . . . Twitter <u>strategy</u> matters.

Twitter Volume



Twitter Valence



Bank of America

- Greater volume
- More personalized
- Customer service-focused; more <u>neutral</u>

GOOD IDEA!

- Low volume
- Mass audience
- Promotional; more <u>positive</u>

BAD IDEA!

WELLS FARGO

C) Advertising bypasses the echoverse.

1

No impact on news articles

2

No impact on word of mouth

3

No impact on consumer sentiment

C) Advertising bypasses the echoverse.

BUT Can still influence customer deposits

4

Thanks!

Questions?

wmrand@ncsu.edu
@billrand

NC STATE
Poole College of Management