ASSOCIATION FOR CONSUMER RESEARCH

When Words Sweat: Identifying Signals For Loan Default in the Text of Loan Applications

Oded Netzer, Columbia University, USA

Alain Lemaire, Columbia University, USA

Michal Herzenstein, University of Delaware, USA

We automatically process the raw text in thousands of loan requests from an online crowdfunding platform, and find that borrowers, consciously or not, leave traces of their intentions, circumstances, and personality traits in that text. Moreover, the text is predictive of default up to three years after it was written.

[to cite]:

[url]:

[copyright notice]:

Method

We analyze thousands of songs, testing how lyrical differentiation (i.e., how different a song's lyrics are from others within its genre) shapes popularity.

First, we collected song data. We captured consumer demand using Billboard's digital download rankings. This measures sales data from major paid song services (e.g., Apple iTunes). We obtained all songs ranked each quarter for three years (2014-16) for each of seven major genres (e.g., Pop, Rap, Rock). This resulted in a dataset of 4,200 song rankings and 1,879 unique songs.

Next, we used latent Dirichlet allocation (LDA, Blei 2012) to determine key lyrical features. LDA takes a number of texts (songs in our case), and by measuring word co-occurrence, determines the key latent topics or themes that make up those texts, and the words that make up each topic. In this case, some of the themes identified were "car cruising" (e.g., words like "road," "ride," and "drive"), "uncertain love" (e.g., "heart," "need," "love"), and "body movement" (e.g., "shake," "bounce," "clap").

Then, we used these latent topics to calculate how different each song is from songs in its genre. The LDA model outputs the proportion of each song that belongs to each latent topic. One song, for example might be 25% "car cruising," 45% "uncertain love," 10% "body movement," and so on. Averaging across all songs in a genre provides the genre's mean use of language. Country songs, for example, talk a lot about "girls and cars" (39%) and "uncertain love" (36%) and less about "body movement" (2%). We calculate the differentiation of each song by taking the difference between its score and the genre mean for a given topic, aggregating across topics.

Finally, we use an OLS regression to examine the relationship between lyrical differentiation and song performance.

Results

Results demonstrate a significant relationship between lyrical differentiation and popularity. The more different a song's lyrics are from its genre, the more successful it is. 16% lyrical differentiation from the genre mean corresponds to a one position improvement in chart ranking. These results hold controlling for numerous covariates (e.g., artist, radio airplay, prior chart appearance, genre, and presence in multiple genres) and alternative specifications, including those accounting for non-linearity or selection. The results are also robust to whether we compute topics across or within genres. Across all specifications, differentiation was linked to success.

More or Less Typical?

Differentiation could occur in one of two ways. While Rap songs with more differentiated lyrical topics are more successful, for example, this could be because they include even *more* rap-associated content than other Rap songs or because they include *less* Rap-associated content and more content associated with other genres.

Results suggest the latter. Successful songs tend to use less of their genre's own typical topics (e.g., "bad man" in Rap) and more of topics associated with other genres (e.g., country's "uncertain love" in Rap).

Moderation by Genre

Finally, we examined whether the relationship between lyrical differentiation and success varied by genre. Differentiation was good overall, but lyrics may matter less in Dance music, for instance, where attributes that drive movement (e.g., the beat) may be more important than lyrics (Fabbri 1981). Results are consistent with this prediction. Lyrical differentiation matters less in Dance, as well as Pop music, which, almost by definition, is more about mainstreaming than differentiation (Frith 1986).

Conclusion

In conclusion, analysis of thousands of song suggests that differentiation plays an important role in cultural success. More generally, this work demonstrates how natural language processing can be used to study cultural dynamics, and sheds light on psychological drivers of why things catch on.

How do Authenticity Meanings Evolve?
A Longitudinal Analysis of Music Album Reviews

EXTENDED ABSTRACT

Research has shown that consumers value authenticity (Grayson and Martinec 2004; Rose and Wood 2005; Valsesia et al 2015). As past research has emphasized, authenticity can be generally defined in one of two ways: (a) being true to one's self, which we call sincerity, but has also been known as expressive authenticity or creative authenticity (Dutton 2004; Ilicic and Webster 2014; Valsesia et al 2015) or (b) being true to some established type or genre, which we call heritage, but is also known as iconic authenticity (Grayson and Martinec 2004) or type authenticity (Davies 2001). In this research, we explore how changing field-level factors can influence the definition of authenticity preferred by music critics. By examining the context of music album reviews, we longitudinally analyze how authenticity meanings evolve over the development of a music genre. The music industry is one domain in which authenticity has particular value to consumers (e.g., Grazian 2003; Peterson 1997). Because audiences value authenticity in music, artists often seek to create authentic music and to perform it authentically, and critics often comment on the authenticity of a musician and his or her music (Peterson 1997; Valsesia et al 2015).

Although previous research has examined the individual and small-group effects involved with evaluating authenticity (Grayson and Martinec 2004; Grayson and Shulman 2000; Valsesia et al 2015) and coping with its paradoxes (Rose and Wood 2005), it has not accounted for field-level dynamics—changes in the age, number, and variety of alternatives—on criteria for evaluating authenticity. Using theoretical tools from cultural sociology and institutional theory, we examine the change of criteria for judging authenticity over time and in conjunction with changes in field age and size.

In our work, we empirically study the evolution of authenticity meanings over time by focusing on the music albums reviewed by the magazine *Rolling Stone* from 1967 to 2014. Our sample consists of 4,924 reviews, which were downloaded from the official website of the magazine. We first conducted a qualitative analysis to identify the most common ways in which the concept of authenticity was discussed. This first analysis enabled us to identify three different types of authenticity most common in our dataset, and which are similar to those identified by Napoli et al. (2014): quality commitment, sincerity, and heritage. Then, we conducted a quantitative, automated text analysis of *Rolling Stone*'s music album reviews (Humphreys 2014) to measure the change in these types of authenticity over time.

We find that, when a genre is new, music reviews tend to mention authenticity in association with quality commitment. As the genre matures, reviewers focus more on authenticity as sincerity, evaluating the motives of the artist. Then, as a genre matures further and faces competition from sub-genres and other genres, reviewers focus more on authenticity as heritage, evaluating it in relation to an ideal type for that genre. We argue that this progression is common across most genres because, over time, the resources available to critics change (as well as the resources available to artists and audiences). By "resources," we mean common understandings among stakeholders about the genre.

Criteria for evaluating authenticity in our dataset therefore has three distinct periods. Specifically, in Phase I (1960s-1970s), quality commitment resulted to be the dominant dimension. During Phase II (1980s-1990s), sincerity was the main dimension. In Phase III (2000s-2010s), heritage was the salient dimension. By drawing both historical and cross-sectional comparisons, we can assess differences according to field-level dynamics of both size and age of the genre. For example, avant-garde, which has remained a small genre, went through quality commitment and sincerity stages, but has not yet shifted to evaluation of authenticity according to heritage or type.

Earlier in the development of a music genre, critics, artists, and audiences do not have the resources to judge authenticity according to type. Instead, they rely on a definition of authenticity that does not require field-specific knowledge that people have developed across domains—namely, the common understandings they have regarding how to judge whether someone is being true to self or quality according to some abstract standard.

Using work from cultural sociology, we argue that cultural resources are needed to develop criteria for judging authenticity based on type. As Peterson (1997) points out, an artist cannot be true to an established model until a model has been established. For critics to feel comfortable referring to a musician's adherence to the standards of a genre, industry stakeholders (including both critics and audiences) must have a sufficient level of field-specific knowledge regarding a genre's history and tropes. Commenting about the difference between country music in the 1990s versus the 1920s (when the genre was in its infancy), Peterson argues that: "the most obvious change … is that there was then no clear tradition of country music with its own past, its own iconic progenitors, its own institutional delivery system, and its own self-conscious fan community" (Peterson 1997, p. 223).

This research makes several contributions to our understanding of authenticity. First, we identify the differential role of field-dependent and field-independent knowledge in criteria for assessing authenticity. That is, field-dependent knowledge leads to assessments of quality commitment and sincerity, but assessments of heritage require more sophisticated knowledge of fields and attributes. Second, the findings explore a kind of cultural capital that is different from that examined previously. Bourdieu defines cultural capital as a resource that an individual person builds up in order to gain advantage in a field. The cultural capital needed to define and promote heritage-based authenticity is not really "owned" by a particular person (although some can be more/less knowledgeable about it) but is instead a resource that everyone in the field can access and refer to when making aesthetic judgments. Although previous research has enumerated different types of authenticity, it has not specified how these are related or may shift in relation to institutional, field-level dynamics.

When Words Sweat: Identifying Signals for Loan Default in the Text of Loan Applications

EXTENDED ABSTRACT

Unprecedented loan default levels were at the heart of the financial crisis of 2008-9 (Lewis 2015). While the majority of these loans were mortgages, a significant amount was in consumer loans. As a result, it became difficult for consumers to secure a bank loan—as much as 30% of the people who wished to get a loan were not able to obtain one through conventional channels (Dogra and Gorbachev 2016). Consequently, Americans turned to other sources. In 2011, 14% of American households (over 16 million households) reported using "non-bank credit," which includes payday loans and pawn shops (Mills and Monson 2013). Arguably, a better alternative to these "shark" loans emerged in the form of online crowdfunding platforms, which serve as a marketplace for lenders and borrowers. These platforms have become a substantial financial market, with over 1,250 sites and $34.4 billion in raised funds in 2015 (Crowdfunding industry report, Massolution 2015).

While the idea of crowdfunding may be appealing to both borrowers, who may not be able to get a loan elsewhere, and lenders, who may be able to get better returns on their investments, default rates in crowdfunding sites are high, which puts the entire concept at risk. In comparison to more secured channels such as bank loans, default rates are especially important in crowdfunded loans because the crowdfunding process is riskier and more uncertain–lending money to strangers without any collateral. Furthermore, the online nature of crowdfunding platforms eliminates the human interactions around the financial transaction of granting a loan, which has been shown to reduce default rates. Indeed, Agarwal and Hauswald (2010) found that supplementing the loan application process with the human touch of loan officers decreases default rates significantly due to better screening on the bank's part and higher interpersonal commitment on the consumer's part.

We propose that in a similar vein to loan officers, who are able to assess default likelihood in the offline banking environment, the text that borrowers write when requesting an online crowdfunded loan may leave some traces—similar to body language detected by loan officers—that can improve predictions of future repayment behavior. Specifically, we investigate whether borrowers leave traces of their intentions, circumstances, emotional states, and personality traits in the text they write that are predictive of whether they will default on their loan up to three years after the text was written. To answer these questions we apply text-mining and machine-learning tools to a large dataset of loans from the crowdfunding platform Prosper.

We began by creating an ensemble of predictive models consisting of decision trees and regularized logit models. We find that the predictive ability of the textual information alone is of similar magnitude to that of the financial and demographic information. Moreover, supplementing the financial and demographical information with the textual information improves predictions of default by as much as 4.03%.

Next we used a multi-method approach to uncover the words and writing styles that are most predictive of default. Using a naive Bayes and an L1 regularization binary logistic model we find that loan requests written by defaulting borrowers are more likely to include words related to the borrower's family, financial and general hardship, mentions of God and the near future, as well as pleading lenders for help. We use a latent Dirichlet analysis (LDA) to identify the loan purpose, life circumstances, and writing styles that are most associated with loan default. We find that loans whose purpose is to help with a business or medical circumstances are riskier than other loan types in terms of their default likelihood. Consistent with the naive Bayes results, we find that lenders pleading for help and providing explanations are also associated with higher risk of default.

We further explore the writing styles and personality traces embedded in the loan request text using the Linguistic Inquiry and Word Count dictionary (LIWC; Tausczik and Pennebaker 2010). We find that defaulting loan requests are written in a manner consistent with the writing style of extroverts. This is not surprising as extroverts aspire to have exciting lives, more so than others, and hence spend their money on experiences and worry less about the financial consequences (Brown and Taylor 2014). We further find that defaulting loan requests are written in a manner consistent with the writing style of liars. While we are unable to claim that defaulting borrowers were intentionally deceptive when they wrote the loan request, we believe

their writing style may have reflected their doubts in their ability to repay the loan.

Our research shows that in an environment characterized by high uncertainty, verifiable and unverifiable data have similar predictive ability. While borrowers can truly write whatever they wish in the textbox of the loan application—supposedly "cheap talk" (Farrell and Rabin 1996)—their word usage reveals something about them that we could not have inferred from their financial dossier alone, namely their personality and intentions.

Make America Tweet Again:
A Dynamic Analysis of Micro-blogging
During the 2016 U.S. Republican Primary Debates

EXTENDED ABSTRACT

The 2016 presidential election illustrated the growing role that micro-blogging sites such as Twitter play in electoral politics. Whereas politicians once relied on stump speeches to reach potential voters, candidates can now directly speak to their audiences in real time from the convenience of their devices, whether it be to react to a comment made by another, or express a view about an emerging news event. Indeed, the pervasive influence of micro-blogging has been such that *New Yorker* writer Nathan Heller (2016) called the 2016 U.S. presidential campaign the country's first "Twitter Election"—one where the battle for the "Twittersphere" loomed almost as large as the election itself.

Whether this new channel of communication has helped better inform the electorate in elections, however, is unclear. On one hand, the increased volume of messages from candidates allows them to reach a broader cross-section of potential voters, something that could enhance knowledge about their positions on key issues. On the other hand, the fact that these messages come in the form of 140-character sound bites implies that this broader reach could come with a loss of depth.

In this paper we attempt to explore this issue by reporting an analysis of a unique dataset that characterizes how the substantive and affective content of Tweets evolved before, during, and after the course of three critical Republican Primary debates leading to the presidential election. We focus on the primary debates because past work has shown that they have the largest influence in shaping voter opinion about the candidates—more so than the final presidential debates, when voter opinions are typically solidified (McInney and Warner 2013). We examine how Tweeting behavior—who is Tweeting, what is being Tweeted, and the sentiment that is expressed in the Tweets—changed as the debates evolved over time.

We emerge with three major sets of findings that collectively offer a disquieting view of the ability of the "Twittersphere" to serve as a substantive source of information about the debates. First, we found that as the debates proceeded the Twittersphere provided an increasingly backward-looking account of the debates, as original content describing the event were increasingly dominated in number by re-Tweets of the more popular Tweets posted earlier. In other words, as time progressed the Twittersphere increasingly resembled a replaying of the "greatest hits" from earlier in the debate.

More critically, we found that that this shift to retweeting was also accompanied by a change in the tone and content of what was being posted. While Tweets posted during the debates often mentioned substantive policy issues, after the debate Twitter looked more like a news tabloid, dominated by re-Tweets of content focusing on the more sensationalistic events that occurred during the debates. For example, early in the August 2015 debate there was an initial flurry of Tweeting about the acidic exchange between Donald Trump

and Megyn Kelly about his views toward women. This quickly died out, however, as Tweets shifted their attention to the more substantive policy topics being discussed by the candidates. But by the end of the debates there was a sudden resurgence of interest in the Trump-Kelly exchange, such that one hour after the debates that single topic dominated the Twittersphere.

The third major finding was more theoretical in nature: statistical analyses of the drivers of re-Tweeting rates suggested that the mechanism that drove the popularity of a given Tweet was more complex than that reported in prior research on the drivers of content sharing (e.g., Berger and Milkman 2012). Surprisingly, simple negative and positive emotion tended to be *negatively* associated with re-Tweeting counts. In our case Tweet popularity was positively associated with more nuanced factors such as visual appeal, criticisms expressed through humor, and words related to reward and achievement (e.g., "won," "lost").

We suggest that this finding may be explained by the unique motivations for sharing content during live broadcast events on Twitter relative to other media. Unlike message sharing on more restricted networks (e.g., texting, email, or even Facebook), Tweets have the potential to reach particularly broad audiences, such that Tweeters may be motivated to create and/or share posts that allow them to gain as large a share of the Twittersphere as possible during the short span of the events (e.g., Barasch and Berger 2014). If Tweeters believed that the most popular content would be that eschewing substance and embracing sensationalism, as time progressed Tweets containing the latter should come to dominate. In essence, what emerges might be a norm of sensationalism (Heath 1996), where the Twittersphere is dominated by the most extreme instances of this norm; the funniest videos, the most sensationalist moments from the debates.

While our analyses here focused on the dynamics of Tweeting during political debates, we suggest that the findings may generalize to other televised live events such as sporting matches, election coverage, and awards shows — events where there is typically a high flurry of Tweeting activity during a relatively short period of time. There, like here, we might expect to see re-Tweeting come to dominate the generation of original content as events wear on, with the most sensationalist content having the longest staying power. We conclude the paper with a discussion of both the theoretical implications of the work for research on message sharing as well as substantive implications of the findings for political processes.

REFERENCES

Paper #1

Berger, Jonah, Eric T. Bradlow, Alex Braunstein, and Yao Zhang (2012), "From Karen to Katie: Using Baby Names to Understand Cultural Evolution," *Psychological Science*, 23(10), 1067-1073.

Bielby, William T. and Denise D. Bielby (1994), "All hits are flukes: Institutionalized decision making and the rhetoric of network prime- time program development," *American Journal of Sociology*, 99, 1287–1313.

Blei, David M. (2012), "Probabilistic Topic Models," *Communications of the ACM*, 55 (4), 77-84.

Fabbri, Franco (1981), "Musical Genres," in *Popular Music Perspectives*, ed. D. Horn and P. Tagg; Göteborg and Exeter: International Association for the Study of Popular Music, p. 52-81.

Flavell, John H., Patricia H. Miller, and Scott A. Miller (2001). *Cognitive development*. Upper Saddle River, NJ: Prentice Hall.

Frith, Simon (1986), "Why do songs have words?," *The Sociological Review*, 34(1), 77-106.

Gladwell, Malcolm (2006), "The formula," *The New Yorker*, Retrieved from http://www.newyorker.com/ archive/2006/10/16/ 061016fa_fact6.

Hahn, Matthew W., and R. Alexander Bentley (2003), "Drift as a mechanism for cultural change: An example from baby names," *Proceedings of the Royal Society B: Biological Sciences*, 270, S120–S123.

Heath, Chip, Chris Bell and Emily Sternberg (2001), "Emotional selection in memes: The case of urban legends," *Journal of Personality and Social Psychology*, 81, 1028–1041.

Hirsch, Paul M. (1972), "Processing fads and fashions: An organization set analysis of cultural industry systems," *American Journal of Sociology*, 77, 639–659.

Lieberson, Stanley (2000). *A matter of taste: How names, fashions, and culture change*. New Haven, CT: Yale University Press.

Monahan, Jennifer L., Shiela T. Murphy and Robert B. Zajonc (2000), "Subliminal mere exposure: Specific, general, and diffuse effects," *Psychological Science*, 11, 462–466.

Salganik, Matthew J., Peter S. Dodds and Duncan J. Watts (2006), "Experimental study of inequality and unpredictability in an artificial cultural market," *Science*, 311, 854–856.

Simonton, Dean K. (1980), "Thematic Fame, Melodic Originality, and Musical Zeitgeist: A Biographical and Transhistorical Content Analysis," *Journal of Personality and Social Psychology*, 38 (6), 972-983.

Tausczik, Yla R., and James W. Pennebaker (2010), "The psychological meaning of words: LIWC and computerized text analysis methods," *Journal of Language and Social Psychology*, 29(1), 24-54.

Zuckerman, Ezra (2016), "Optimal distinctiveness revisited," in *Oxford Handbook of Organizational Identity*, ed. M. G. Pratt, M. Schultz, B. E. Ashforth, D. Ravasi, (Oxford: Oxford University Press), 183-199.

Zuckerman, Marvin (1979), *Sensation seeking: Beyond the optimum level of arousal.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Paper #2

Davies, S. (2001). *Musical Works and Performances: A Philosophical Exploration*, Oxford, UK: Clarendon Press.

Dutton, Denis (2004), "Authenticity in Art," in *The Oxford Handbook of Aesthetics, ed. Jerrold Levinson*, Oxford, UK: Oxford University Press, 258–74.

Grayson Kent and Radan Martinec (2004), "Consumer Perceptions of Iconicity and Indexicality and Their Influence on Assessments of Authentic Market Offerings," *Journal of Consumer Research*, 31 (2), 296–312.

Grayson, Kent and David Shulman (2000), "Indexicality and the Verification Function of Irreplaceable Possessions: A Semiotic Analysis," *Journal of Consumer Research*, 27 (June), 17–30.

Grazian, D. (2003) *Blue Chicago: The Search for Authenticity in Urban Blues Clubs*, Chicago, IL: University of Chicago Press.

Humphreys, Ashlee, (2014) "The Use of Text-Analysis in Consumer Research," *Working Paper*.

Ilicic, Jasmina and Cynthia M. Webster (2014), "Investigating Consumer-Brand Relational Authenticity," *Journal of Brand Management*, 21 (4), 342–63.

Napoli, Julie, Sonia J. Dickinson, Micheal B. Beverland, and Francis Farrelly (2014), "Measuring Consumer-Based Brand Authenticity," *Journal of Business Research*, 67 (6), 1090–8.

Peterson, R.A. (1997) *Creating Country Music: Fabricating Authenticity*, Chicago, IL: University of Chicago Press.

Rose, Randall L. and Stacy L. Wood (2005), "Paradox and the Consumption of Authenticity Through Reality Television," *Journal of Consumer Research*, 32 (1), 284–96.

Valsesia Francesca, Joseph C. Nunes and Andrea Ordanini (2016), "What Wins Awards Is Not Always What I Buy: How Creative Control Affects Authenticity and Thus Recognition (But Not Liking)," *Journal of Consumer Research*, 42 (6), 897–914.

Paper #3

Agarwal, Sumit, and Robert Hauswald (2010), "Distance and Private Information in Lending," *Review of Financial Studies*, 23 (7), 2757-2788.

Brown, Sarah, and Karl Taylor (2014), "Household Finances and the 'Big Five' Personality Traits," *Journal of Economic Psychology,* 45, 197-212.

Dogra, Keshav and Olga Gorbachev (2016), "Consumption Volatility, Liquidity Constraints and Household Welfare," *The Economic Journal*, forthcoming.

Farrell, Joseph, and Matthew Rabin (1996), "Cheap Talk," The Journal of Economic Perspectives, 10 (3), 103-118.

Lewis, Michael (2015), *The Big Short: Inside the Doomsday Machine*. WW Norton & Company.

Mills, Gregory B. and William Monson (2013), "The Rising Use of Nonbank Credit among U.S. Households: 2009-2011," *Urban Institute*.

Tausczik, Yla R., and James W. Pennebaker (2010), "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods," *Journal of Language and Social Psychology,* 29 (1), 24-54.

Paper #4

Barasch, Alix, and Jonah Berger (2014), "Broadcasting and Narrowcasting: How Audience Size Affects what People Share," *Journal of Marketing Research,* 51(3), 286-299.

Berger, Jonah, and Katherine Milkman (2012), "What Makes Online Content go Viral," *Journal of Marketing Research,* 49 (2), 192-205.

Heath, Chip (1996), "Do People Prefer to Pass Along Good or Bad News? Valence and Relevance of News as Predictors of Transmission Propensity", *Organizational Behavior and Human Decision Processes*, 68(2), 79-94.

Heller, Nathan (2016), "The First Debate of the Twitter Election", *New Yorker*, September 27.

McInney, M. S., and Benjamin R. Warner (2013), "Do Presidential Debates Matter? Examining a Decade of Campaign Debate Effect," *Argumentation and Advocacy*, 49(Spring), 238-58.