Measuring Organizational Legitimacy in Social Media

Assessing Citizens' Judgments With Sentiment Analysis

Etter, Michael; Colleoni, Elanor; Illia, Laura; Meggiorin, Katia; D'Eugenio, Antonino

Document Version
Accepted author manuscript

Published in:
Business & Society

DOI:
10.1177/0007650316683926

Publication date:
2018

License
Unspecified

Citation for published version (APA):
Etter, M., Colleoni, E., Illia, L., Meggiorin, K., & D'Eugenio, A. (2018). Measuring Organizational Legitimacy in Social Media: Assessing Citizens' Judgments With Sentiment Analysis. *Business & Society*, *57*(1), 60–97. https://doi.org/10.1177/0007650316683926

Measuring Organizational Legitimacy in Social Media: Assessing Citizens' Judgments With Sentiment Analysis

Michael Etter, Elanor Colleoni, Laura Illia, Katia Meggiorin, and Antonino D'Eugenio

Journal article (Accepted manuscript*)

Please cite this article as:
Etter, M., Colleoni, E., Illia, L., Meggiorin, K., & D'Eugenio, A. (2018). Measuring Organizational Legitimacy in Social Media: Assessing Citizens' Judgments With Sentiment Analysis. *Business & Society*, *57*(1), 60-97. https://doi.org/10.1177/0007650316683926

DOI: https://doi.org/10.1177/0007650316683926

Measuring Organizational Legitimacy in Social Media
Assessing Citizens' Judgments with Sentiment Analysis

Michael Etter, King's Business School
Elanor Colleoni, Reputation Institute
Laura Illia, IE University
Katia Meggiorin, IE university
Antonino D'Eugenio, UniCredit SA

Article accepted at *Business and Society*: https://doi.org/10.1177/0007650316683926

Abstract

Conventional quantitative methods for the measurement of organizational legitimacy consider mainly three sources that make judgments about organizations visible: news media, accreditation bodies, and surveys. Over the last decade, however, social media have enabled ordinary citizens to bypass the gatekeeping function of these institutional evaluators and autonomously make individual judgments public. This inclusion of voices beyond functional and formally organized stakeholder groups potentially pluralizes the ongoing discussions about organizations. The individual judgments in blogs, tweets, and Facebook posts give indication about the broader fit between an organization's perceived behavior and heterogeneous social norms and therefore constitute an indicator of organizational legitimacy that can be accessed and measured. We propose the use of social media data and sentiment analysis to study the affect-based responses to organizational actions by citizens. We critically discuss and compare the method with existing quantitative methods for legitimacy measurement and apply them to a recent case in the banking industry. We discuss the value of the method for studying the process of legitimacy construction as the expression and negotiation of normative judgments about organizations by various evaluators.

Keywords: Organizational legitimacy, measurement, social media, sentiment analysis, Twitter

Organizational legitimacy is generally defined as the social acceptance of organizations and their actions (Suchman, 1995). Based on Suchman (1995), business and society scholars have conceptualized the process of legitimation, or legitimacy construction, as the expression and negotiation of normative judgments by various evaluators in a public discourse (Scherer & Palazzo, 2006; Schultz, Morsing, & Castelló, 2013). Conventional quantitative measurements have taken into account three sources that make these normative judgments visible. With quantitative content analysis, researchers have assessed the evaluation of organizations by *news media* (e.g., Pollock & Rindova, 2003), a powerful evaluator that influences the public opinion about organizations (Carroll & McCombs, 2003) and gives direct, selective access to individual judgments (Lee & Carroll, 2011); measurements based on *accreditation bodies* reveal the balanced expert-evaluation by influential evaluators that legitimize organizations through visible linkages (e.g., Baum & Oliver, 1991); *surveys* give access to the evaluation of organizations by the general public or specific stakeholder groups whose judgments enter the public domain through condensed public rankings and indices (Fombrun, 2007).

These measures are valuable to assess evaluations about organizations based on the norms, evaluation criteria, and selection processes of respective evaluators. However, they only partly account for the plurality of norms, values, expectations, and concerns of ordinary citizens, actors that have been regarded as increasingly important from a normative perspective that calls for a stricter democratic accountability of corporate behavior (Matten & Crane, 2005). News media content, for example, is shaped by various selection processes (Shoemaker & Reese, 2014) and its analysis therefore doesn't account for opinions, organizations, or events that are not selected (Vergne, 2011). The observation of linkages to accreditation bodies gives, at best, indirect indication about to the evaluation by ordinary citizens. Survey-based measures, while giving direct access to citizens' judgments, are typically only calculated and published once a year (e.g., Fombrun, 2007) and have been criticized to cover limited and only predefined organizational aspects, for which respondents might have also insufficient knowledge (Helm, 2007).

In sum, conventional measures of organizational legitimacy capture a mere fraction of the plurality of citizens' judgments, because the possibilities for citizens to express their opinions through institutional evaluators are limited.

With the rise of social media, however, ordinary citizens have begun to create autonomous public arenas, where organizational activities are continuously discussed and evaluated (Whelan et al., 2013). While opinion pages of newspapers provide selected citizens a limited forum to express opinions (Lee & Carroll, 2011), social media have given citizens the possibility to bypass this gatekeeping function of news media (Papacharissi, 2009). Although some scholars have questioned the democratic potential of Internet technologies (e.g., Hindman, 2009), others conceive the variety of debates in social media as potentially more democratic (Whelan et al., 2013; Castells, 2007). Recently, researchers have started to grasp organizational legitimacy in social media. For example, Castelló et al. (2016) and Colleoni (2013) have analyzed stakeholder tweets about organizations to measure the outcomes of different communication strategies. More generally, scholars have studied how stakeholders raise their voices and express their opinions about organizations online (Etter & Vestergaard, 2015; Hunter, Le Menestrel, & De Bettignies, 2008)[1]. We argue that the systematic and direct access to these voices can contribute to the understanding of the construction of legitimacy in a "normative context that becomes transnationalized, fragmented, pluralized, more complex and less understandable" (Palazzo & Scherer, 2006, p. 79). A legitimacy measure based on social media data, then, has the potential to complement extant measures and contribute to a more encompassing understanding of legitimacy based on the judgments by various evaluators.

In order to measure citizens' judgments in social media, we suggest the use of sentiment analysis, a method that has recently gained attention by business and society scholars (e.g., Castelló et al., 2016), yet still lacks a thorough and critical discussion. This method bases on the techniques of natural language processing, text analysis, and computational linguistics, and measures the affective orientation of sentences towards an object (Pang & Lee, 2008). It therewith gives indication about "affect-based responses" through which ordinary citizens bestow legitimacy to organizations (Haack, Pfarrer, & Scherer, 2014, p. 634). The method allows assessing judgments about organizations 1) in large amounts of social media data, 2) to develop time-sensitive measures, and 3) to access judgments about understudied organizational issues by important actors of civil society.

After a discussion of conventional measurements, we will introduce the method of sentiment analysis and show with an illustrative case study based on Twitter data how it can complement conventional measurements of organizational legitimacy.

Organizational Legitimacy

The Construction of Organizational Legitimacy

Suchman's (1995, p. 577) notion of moral legitimacy as the "explicit normative evaluation of the organization and its activities" has gained particular attention by business and society scholars, who argue that legitimacy is constructed through the expression and negotiation of normative judgments in a public discourse (Castelló et al., 2016; Palazzo & Scherer, 2006; Schultz et al., 2013). This conceptualization has met particular resonance in business and society research, because it accounts for the pluralization of norms and expectations towards organizations. For some scholars, then, legitimacy emerges from a deliberative discourse among various actors with the active participation of organizations (Scherer & Palazzo, 2006). For others, legitimacy is communicatively constructed, often involving dissent and beyond corporate knowledge and control (Castello et al., 2016; Schultz et al., 2013). Inherent to both views is that legitimacy assessments are expressed in a discourse of evaluators (Bitektine & Haack, 2015). The evaluators, or the sources of legitimacy, are "the internal and external audiences who observe organizations and make legitimacy assessments" (Ruef & Scott, 1998, p. 880).

There are many evaluators that asses an organization's conformity to specific standards and norms (Ruef & Scott, 1998). To reach these judgments, evaluators take into consideration various aspects, such as organizational accomplishments, the way organizations operate, the way organizations are structured, and the evaluation of their leaders (Suchman, 1995, pp. 579-582). For example, news media evaluate organizational outcomes and actions, including various dimensions of (ir-) responsible behavior (e.g., Lee & Carroll, 2011); accreditation bodies assess organizational compliance with professional norms, such as a bank's ability to protect depositor savings (e.g., Deephouse & Carter, 2005); and survey-based ranking asses the public's evaluation for various aspects, such as the protection of the environment (e.g., Fombrun, 2007). Because different evaluators tend to have varying interests and use diverse criteria and norms in their evaluation, attention to various evaluators is important (Meyer & Scott, 1983).

How Citizens' Judgments in Social Media construct Legitimacy

Social media have given ordinary citizens the possibility to express and negotiate judgments about organizations online (e.g., Castelló et al., 2013; Whelan et al., 2013). This evaluation of corporations is often expressed in conversations, which, for example, take place under a certain hashtag, a linguistic marker (e.g., #panamapapers) that is used to categorize conversations by topics (Albu & Etter, 2015). Studies have shown that approximately 20 percent of all social media content is about organizations (Jansen, Zhang, Sobel, & Chowdury, 2009). Furthermore, judgments in social media are impactful because ordinary citizens increasingly use them as information sources to assess and negotiate the appropriateness of organizational actions (Castelló et al., 2013; Whelan et al., 2013). Twitter, for example, is generally regarded as a reliable source of information (Westerman, Spence, & Van Der Heide, 2014), and the evaluation of corporation, their actions and services are deemed as credible (Sotiriadis & van Zyl, 2013). Because of their impact, credibility, and increasing volume and importance, judgments in social media then can be considered as contributing to the co-construction of organizational legitimacy.

In contrast to institutional evaluators, such as news media and accreditation bodies, that typically endorse or question organizations with more or less rational and balanced evaluations (e.g., Meyer & Scott, 1983; Deephouse & Carter, 2005), citizens bestow legitimacy to organizations through affect-based responses (Haack et al., 2014). Affect-based responses, such as joy (positive affect) or disappointment (negative affect), are the affective reactions to a perceptional input, such as organizational actions, and form the basis of individual judgment formation (Haidt, 2001). Affect-based responses are verbalized and expressed as evaluative judgments about organizations and their actions. Similar to endorsing articles in news media (Deephouse & Carter, 2005), positive judgments, then, can be considered as legitimizing organizations, while negative judgments can be considered as de-legitimizing organizations. Consequently, we argue that legitimacy is constructed through individuals' expression of positive and negative judgments in social media.

While institutional evaluators cover citizens' judgments to a limited degree, social media facilitate their expression in various ways (Papacharissi, 2014). Different than institutional evaluators that follow professional rules, norms, and conventions (e.g., Shoemaker & Reese, 2014;

Baum & Oliver, 1991), ordinary citizens are less restricted, which allows them to express their individual feelings and opinions as semi-private utterings, such as outrage and praise, in a personal style of writing (Papacharissi, 2014).

In sum, additional to judgments by institutional evaluators, social media make the judgments of ordinary citizens public that contribute to the construction of organizational legitimacy (for an overview of sources see table 1).

Insert Table 1 about here

Conventional Quantitative Measurements of Organizational Legitimacy

Applying quantitative methods, researchers have so far mainly considered news media, accreditation bodies, and surveys for the measurement of organizational legitimacy.

News Media Content Analysis

News media are the most extensively explored sources of organizational legitimacy (e.g., Bansal & Clelland, 2004; Deephouse, 1996; Vergne, 2011), also called "media legitimacy" (Bitektine, 2011, p. 154). News media (de-)legitimize organizations by making organizational activities visible and evaluating them (Suchman, 1995). Consequently, legitimacy measurements based on news media content measure how organizations are covered and framed. The method assigns meaning to a selection of words to assess the positive, neutral, or negative tone of news media coverage about an organization.[2] The aim of the analysis is to build a tonality or favorability index, such as the Janis-Fadner imbalance coefficient (Janis & Fadner, 1943; Deephouse & Carter, 2005).

Because news media are a major influencer of the public opinion about organizations (Carroll & McCombs, 2003), it can be argued that news media analysis is a valid method to assess organizational legitimacy. Furthermore, studies have shown that news media influence discourses about organizations on social media, for example by setting the frames during a crisis (Etter & Vestergaard, 2015; Van de Meer & Verhoeven, 2013), and therefore indirectly cover citizens' voices too. The method can also give selective direct access to citizens' opinions through the analysis of letters to the editor (Lee & Carroll, 2011). Nevertheless, news media have their own agendas and content production is influenced by various factors and selection processes (Shoemaker & Reese, 2014). Therefore, news media analysis provides foremost a valuable measure for the evaluation of a powerful institutional evaluator, but gives only limited indication about the judgments by citizens.

Analysis of Organizational Compliance with Accreditation Bodies

Other researchers have examined the role of accreditation bodies which establish and monitor the rules and norms that determine the way organizations in a given domain should perform their activities (Deephouse, 1996; Pfeffer & Salancik, 1978). It has been argued that organizations gain legitimacy, also called "linkage legitimacy" (Bitektine, 2011, p. 156), through visible institutional linkages to these prominent and legitimate actors (Baum & Oliver, 1991).

Linkage measures capture the organizational compliance with regulations and standards by quantifying compliance with rules, standards, and regulations. Frequently used measures include longitudinal data on the number of registered, licensed, rated, or certified organizations (Baum & Oliver, 1991; Pfeffer & Salancik, 1978; Deephouse, 1996; Deephouse & Carter, 2005). Scholars have developed longitudinal indexes related to accreditations or ranking that are repeated over time to analyze the evolution of legitimacy. Organizational linkages are typically coded through either binary variables (e.g., Deephouse & Carter, 2005) illustrating the absence or presence of regulator's actions, accreditation, collaborative relationship or registration of licenses, or ordinal variables that allow to build a ranking system (e.g., Ruef & Scott, 1998)

The method is a valuable measure for organizational legitimacy, because accreditation bodies are powerful institutions that judge organizations with a balanced evaluation based on many dimensions and expert knowledge (Deephouse, 1996; Baum & Oliver, 1991). They provide a more or less balanced evaluation, on which other actors, such as investors, journalists, or citizens build their judgment. However, the method gives, at best, indirect indication about the evaluation by ordinary citizens. Furthermore, because organizations are typically accredited once a year, the analysis does not account for the short- to medium-term dynamism of legitimacy formation.

Surveys

A third quantitative method consists of surveys that assess the evaluation of organizations by the general public or particular stakeholder groups (e.g., Ponzi, Fombrun, & Gardberg, 2011). Prominent examples of survey-based rankings are Fortune's most Admired Corporations (FMAC), the RepTrak (RT), and other polls for decision makers (e.g., Qiu & Welch, 2004).[3]

The merit of this method is that it gives access to the evaluation of organizations by ordinary citizens (e.g., RT) or particular stakeholder groups, such as executives, directors, and analysts (e.g., FMAC). Furthermore, representative samples guarantee that the public is represented in a valid way. Nevertheless, the evaluation is limited to the predefined aspects and items respondents are asked to rate. For example, the FMAC has been criticized for being overly focused on financial aspects (Wartick, 2002). Another widely held critique of items and scaling techniques is that surveys directed at members of the general public do not consider that participants might not have distinct knowledge of the criteria for assessing organizations (Rindova et al., 2005). Schultz et al. (2001) observed that respondents often use intuition when answering scales. Similarly, in referring to the halo effect (Thorndike, 1920), Helm (2007) argues that respondents judge unknown organizational aspects with aspects familiar to them. Finally, because surveys are costly, they are mostly conducted on yearly basis, which restricts timely monitoring and insights into the short-term development of organizational legitimacy.

In sum, we conclude that conventional legitimacy measures capture the compliance with relatively homogenous expectations that are based on established standards and predefined criteria, institutionalized selection and evaluation processes, and on particular interests and agendas of respective sources. While these evaluations might reflect or influence the judgments of ordinary citizens to certain degrees, they only give access to a fraction of the "growing complexity of globalized social networks [that] is accompanied by an internal pluralization of postindustrial societies" (Palazzo & Scherer, 2006, p. 77). In the next paragraphs we describe how social media make this pluralization visible.

Pluralization of Discourses in Social Media

Potential Democratization of Online Arenas

Because citizens can express their opinions in social media, scholars have heralded these technologies for pluralizing public discourses (Castells, 2007). While institutional actors, such as news organizations (Hermida, 2012), NGOs (Blumell, 2016), corporations (Etter, 2014), and organized social movements (Etter & Vestergaard, 2015), use social media to disseminate messages, it is particularly ordinary citizens that express themselves online and outnumber institutional actors. Indeed, only 0.05 percent of current 300 million Twitter accounts are official verified accounts (businesses, news organizations, celebrities, etc.).

By referring to Crane and colleagues' (2008, p. 9-12) notion of "citizenship arenas", Whelan and colleagues (2013) propose that social media enable the creation of democratic public online arenas. The multi-modal and transnational nature of social media foster the autonomy of these arenas and alter, if not diminish, control over communications and media (Castells, 2007). In contrast to news media that are typically profit-oriented organizations (McManus, 1994), tend to follow similar topics (Dearing & Rogers, 1996), and report events only if they are newsworthy (Galtung & Ruge, 1965), ordinary citizens have typically no institutionalized pressure to produce content that sells. The publication of content is rather dependent on the personal interest involvement in a topic (Papacharissi, 2009; Boyd, 2010). As a result, online arenas are populated by a wide set of social concerns and interests (Castelló et al., 2013). Similarly, while surveys and accreditation bodies evaluate organizations with predefined criteria for predefined issues and organizations, social media allow citizens the expression of judgments about any organization and issue from the top of their head.

Furthermore, ordinary citizens are less subject to direct organizational control and influence than traditional news media (Castells, 2007). Corporations influence news media exposure through public relations tactics, such as press releases or personal interactions with journalists (Westphal & Deephouse, 2011). In contrast to journalists, ordinary citizens are less or only indirectly targeted by these public relations tactics and often base evaluations on personal experiences (Boyd, 2010). As a consequence "legitimacy is not only formed in separate spheres of society, within hierarchical orders of stable institutions or powerful elites (…), but co-constructed by networked publics" (Schultz et al., 2013, p. 685).

Limitations of Pluralization

Despite the heralded potential for increasing participation among previously unengaged citizens, scholars have also argued that social media fall somewhat short of the expectations held by those most optimistic on behalf of their democratic and disruptive potential (e.g., Larsson & Moe, 2012). For example, studies have shown that certain user types contribute with larger volume of content to online discourses, and that these heavy users often belong to elite groups with prominent positions in society (e.g., Larsson & Moe, 2012). Furthermore, communities or movements mobilizing around particular topics give certain opinions more weight, might influence the agenda of news media, potentially resulting in an echo chamber between social media and traditional media. (Pfeffer, Zorbach, & Carley, 2013). Similarly, social connections tend to base on homophily and therefore act as a filter overemphasizing the importance of single topics or opinions (Pariser, 2011). The resulting information echoes create the impression of everybody talking about the same topics or having the same opinions (Sunstein, 2009). This phenomenon is increased by technological filters and algorithms that expose individuals to information based on previous interest and social connections (Pariser 2011). Furthermore, it has to be acknowledged that, while most citizens have access to social media, not all make use of them to the same degree (Wei & Hindman, 2011). Usage patterns vary depending on age, gender, education, race, and personality (Hargittai & Litt, 2011; Roblyer et al., 2010). Nevertheless, despite these digital divides, research also shows that those groups actually expressing political opinions online are diverse (Bakker & De Vreese, 2011)

Finally, utterances in social media are subject to several biases. For example, social media users aim at creating favorable images of themselves (Boyd, 2010) and therefore stretch the truth, sometimes to its outer limits (Marwick, 2005). Similarly, citizens often enact self-censorship to accommodate the expectations of an imagined audience (Marwick & Boyd, 2011) or when they feel that their opinion is incongruent with the perceived majority-opinion of online peers (Nekmat &

Gonzenbach, 2013). Furthermore, studies show that organizations increasingly try to influence what employees express in their capacity as citizens online, which is likely to influence what they dare to express in social media (Stohl et al., 2015).

In sum, we conclude that social media have the potential for pluralizing discourses that construct legitimacy. Nevertheless, it has to be acknowledged that not all judgments are necessarily publicly expressed (Bitektine & Haack, 2015). Those, which are expressed, however, contribute to the formation of organizational legitimacy.

Measuring Organizational Legitimacy with Sentiment Analysis

Social media are transforming individual judgments from something private (or at least concentrated to face-to-face interactions) into something public. The result is a tangible manifestation of judgments in tweets, blog entries, and other forms of texts, that can be collected and analyzed.

Collecting Social Media Data and Sampling Techniques

According to our review, social media data have been collected in three ways: Through application programming interfaces (API), data crawlers, or paid services (Lomborg & Bechmann, 2014). APIs are software tools that enable the download of information from social media platforms (De Souza & Redmiles, 2009). Additional to text-data that emerge from online interactions and discussions, APIs can give systematic access to author details, time of publication, and geo-location, which can be used to filter the data for certain regions, actors, or time periods. APIs are regulated by the platform owners, who decide what information is accessible.

While Twitter is widely used by academics not only for its relevance, but also for its generous policy regarding data access, other platforms are less open. For these platforms, researchers can use data crawlers, which collect systematically publicly available information by simulating user behavior. For instance, "Issue crawler" is a crawler developed by the Digital Method Initiative at the University of Amsterdam and collects information about specific issues. The limitation of this approach consists in the need to adapt the crawler for every platform.

Finally, social media data can be collected by paying the service of third-part companies specialized in the systematic data collection.

Regardless of the method, data collection from social media platforms present some challenges to empirical research in terms of sampling (Gillespie, 2010). For example, the interconnectedness of the content and its embeddedness in a network of interrelated discussions and actors challenges the definition of the boundaries of discourses, groups, and often platforms (Jenkins, 2006). This poses the question of *delimitation*, i.e. the selection of coherent and representative sample for subsequent analysis (Gerlitz & Rieder, 2013). The two most common sampling techniques are topic-based sampling and snowball sampling.

Topic-based sampling techniques extract information containing specific key words (Bruns & Stieglitz, 2012). In other research fields, topic based sampling has been used to study the emergence and evolution of topical concerns (Burgees & Bruns, 2012), to explore brand communication (Stieglitz & Krüger, 2011), to observe homophily (Adamic & Glance, 2005), or communication practices around CSR (Colleoni, 2013). The definition of key words aims at delineating a discussion around a specific topic or a specific entity. To this end, a list of the representative keywords is defined by the researcher (e.g., "BP oil spill"), usually with the support of domain experts on the basis of their prior knowledge of the topic (e.g., Colleoni, 2013). Snowball sampling builds on predefined lists of discussions defined by domain experts, and extend these through "snowballing" or triangulation, by investigating the networks of the initially identified actors (Rogers, 2009). This sampling strategy has been used to sample topic- or activity-based user groups (Paßmann, Boeschoten, & Schäfer, 2013), word of mouth networks (Pedersen, Razmerita, & Colleoni, 2014), or CSR communities (Fieseler, Fleck, & Meckel, 2010), and helps identifying relevant actors and respective content, that were not known before. For example, Fieseler and colleagues (2010) have applied this technique to investigate McDonald's engagement in the CSR blog's network.

All the described methods are non-probability selection techniques, particularly suited for case-based analysis, which does not guarantee the selection of a representative sample and therefore limits the generalization of the results (Gilbert, 2008). This does not invalidate the results of the analysis, but it does raise questions about the generality of derived claims in relation to the entire population (Boyd & Crawford, 2012).

After the sample has been extracted, the data are transformed into a form that can be automatically analyzed by the algorithm. This pre-processing phase consists of data cleaning, language detection, and text standardization. The data cleaning process is a particular delicate phase, as a consistent quota of the communication in social media is made by fake accounts. For example, Facebook estimated that around 10% of its accounts are fake (Thier, 2012). Based on prior research on spam and fake accounts (Stringhini, Kruegel, & Vigna, 2010), accounts that grow too fast in terms of friend requests and messages sent, are eliminated, as well as accounts that are created and closed within a day. The text is then cleaned to extract only relevant text, deleting punctuation, numbers, and stop-words, such as parentheses and conjunctions. Once this first data cleaning is completed, an algorithm for language detection is applied to the text in order to select only content in one language. Finally, each word is reduced to its word stem, so that words like computer and computing both become the same term, “comput”. At the end of this procedure, the data consist of a set of tokens that can be analyzed.

Sentiment Analysis of Social Media Data

Once the sample has been properly transformed, the data are analyzed through sentiment analysis. Sentiment analysis is part of a computing paradigm that assesses subjectivity (e.g., affect or opinions) in texts (Pang & Lee, 2008). The method usually detects the polarity of sentences, assessing whether individuals are expressing any form of positive or negative sentiment toward an object. A sentiment measure in the form of a single score (for example, ranging from -100 to +100) or a label (i.e. positive or negative) attached to a word or a set of words, represents the aggregated evaluative judgments in a large number of texts for a pre-defined entity. If a mixed tone is detected, as for instance in the case of two sentences with different sentiment in a single text, the sentiment algorithm computes the final score comparing the amount of negative and positive sentiment. Detection of mixed sentiment, however, is less relevant when dealing with social media data, because it tends to appear in the form of short sentences, such as tweets, which mostly do not include mixed sentiment.

One approach to classify sentiment in texts is lexicon-based. With pre-coded word lists, occurrences of subjective words within texts are computed to detect polarity of a sentence (Rao & Ravichandran, 2009; Thelwall, Buckley, & Paltoglou, 2011). The overall sentiment score represents the sum of all the evaluative expressions of individuals, usually represented as a nominal number normalized by the number of words in a given sentence. The major limitation of lexicon based approaches is that they rely on pre-coded lists and are unable to detect novel expressions.

A more complex approach is represented by the linguistic analysis which bases inferences on the semantic and syntactic structure of the text to extract its sentiment valence (e.g., Cambria, et al., 2013b). In particular, by defining the structure, morphology, and relationship between words, these approaches can identify superlatives, negations, context, and idioms as part of the prediction process (Thelwall et al., 2011) and therefore detect subtly expressed sentiments, overcoming several challenges related to the investigation of social media, such as the use of informal language and uncommon abbreviations. By extracting the structure of the discourse, these algorithms try to detect the holder of the sentiment, the target, and the context (Saif, He, & Alani, 2012). When moving from one context to another, opinions and affect can shift from positive to negative, or neutral. Semantics can capture this evolution and differentiate its results accordingly (Gangemi, Presutti, & Reforgiato, 2014). By relying on large semantic knowledge bases, such approaches advance techniques that blindly use pre-coded words and instead rely on the implicit meaning associated with natural language concepts (Cambria et al., 2013a). The major limitation of these approaches is their limited flexibility and therefore their reduced results in terms of generalization to other domains due to the strong dependency of the results to the original semantic structure used to build the analytical model (Pang & Lee, 2004).

The most widely used approach for sentiment analysis is machine learning (Pang et al., 2002; Zhang, Shang, & Jia, 2015; Agarwal & Bhattacharyya, 2005), which bases on the automatic discovering of useful information, or novel patterns in large data sets. By using this approach, researchers can create algorithms that, properly trained, can classify unforeseen data and therefore analyze real-time data that are constantly created in social media. Most importantly, by using this approach, the set of words that are evaluated as positive or negative is not confined to the pre-defined words list but emerges from the ongoing discourses.

In order to train the algorithm for sentiment analysis, supervised classification is applied (Pang et al., 2002). This requires two data sets: a training set that consists of a corpus of manually labeled text; and a test set in form of a corpus of unlabeled text, which is used by the researcher to validate the performance of the classification. A supervised learning algorithm, then, results from the training set as the mathematical model that best maps the data. In a first step, the training set is created by categorizing text based on the affective orientation of the sentence. This process is driven by the expertise of the researcher, who identifies those features to be included in the analysis. From this initial set of features, those that best describe and discriminate the different affective judgments are extrapolated. The algorithm provides a percentage of accuracy and the researcher choses the combination of features with highest accuracy. These features are used to train a classification model which then is able to automatically detect unforeseen data.

The second step consists in the researcher's assessment of the accuracy of the prediction of the classification model by testing its ability to correctly classifying unlabeled data. To do so, cross-validation is applied that involves partitioning the test set into complementary subsets, and performing the analysis with the classification model on one subset. The classification model is assessed on a set of quality measures, namely accuracy, precision and recall. Through the combination of these measures the overall ability of the classification model to predict the sentiment of a new corpus of text is assessed. The indexes are calculated by the algorithm, the researcher choses the model with best quality. On average, the performances of sentiment classification of social media data range from 70% to 95% correctly classified entities in terms of f-score (Zhang et al., 2015).[4] Once the text has been classified, the data is graphically represented as trend over time.

Measuring Citizens' Judgments as a Source of Legitimacy

Through the detection of affective orientations in large amounts of texts, the advanced techniques of sentiment analysis can be considered as a valid method for the measurement of citizens' judgment in social media data. The method may be recommended for those scholars interested in: (1) exploiting the potential of large scale social media data; (2) developing real time legitimacy indicators; (3) assessing judgments by new or understudied actors.

Exploiting the Potential of Large Social Media Data Sets

The access to big databases in social media allows analyzing the broadness of (de-)legitimizing issues and conversations about corporations. Online conversations and actors are highly fragmented (Takeshita, 2005) and therefore potentially draw a highly complex picture of the multiplicity of issues (de)legitimizing an organization. A large data set is therefore essential to reliably and directly analyze plural conversations about an organization. This might not have been an issue for those scholars analyzing legitimacy of corporations through news media content analysis, because media agendas are less fragmented than the social media ones (Castelló et al., 2013; Papacharissi, 2009). Here, small-medium databases of media reports cover the issues addressed in news media. Measuring legitimacy through sentiment analysis instead allows researchers to directly assess judgments of ordinary citizens. The big amount of data allows computing queries of databases which extrapolate an overall picture of how ordinary citizens evaluate organizations from the multiplicity of conversations.

Developing time-sensitive Legitimacy Indicators

Sentiment analysis may represent a valuable measurement for those researchers who are interested in analyzing (de-)legitimizing content that affects corporations' legitimacy. Previous studies using evaluator-based measures with means of news media or accreditation bodies assessed legitimacy on a monthly or yearly basis. Consequently, these measures are not suitable to assess legitimacy in a communication environment, where speed and dynamism of communications has increased significantly (Castells, 2007; Castelló et al., 2013). Indeeed, yearly measures of legitimacy are helpful in assessing how the dominant voices of institutional actors (de-)legitimize corporations over a longer time period. However, (de-)legitimizing conversations in social media might reach big audiences within a few hours or minutes (Chumley, 2014). Hence, scholars interested in analyzing the evolution of legitimacy in more dynamic situations, such as crises, might use sentiment analysis of social media data.

Assess Judgments of previously understudied Issues and Actors

Sentiment analysis of social media data gives access to the judgments of previously understudied issues and actors. Past studies have followed the logic of first identifying the actors (e.g., New York Times) and then analyzing the content (e.g., news media coverage endorsing vs. challenging corporations). Such an approach bases on the existence of few established and identifiable institutional evaluators that for a long time were the only actors that had the resources to publicly evaluate corporations on a regular basis. Today, however, anybody who is able to attract the attention of other users upon any topic can become an important actor in (de-)legitimation processes. Groups that emerge transnationally around an issue raised by a few users may mutate quickly, when new actors join these groups (Hunter et al., 2008; Chumley, 2014). Hence, the assessment of legitimacy in social media requires a measurement that recognizes how evaluations are expressed not only by clear-well defined actors, but also by a spontaneous aggregation of actors that express judgments around a variety of topics and issues. In this regard, sentiment analysis is a powerful method, because it allows to first analyzing the judgments and opinions of actors for a certain entity. In a second step, researchers can identify interest groups propagating such (de-

)legitimizing sentiment. This allows scholars to have a flexible measurement of legitimacy that allows grasping the fast evolving nature of social media conversations.

Table 2 displays an overview of discussed measurements of organizational legitimacy.

Insert table 2 about here

Illustration: Application of Sentiment Analysis

With a case study we illustrate how citizens' judgments in social media are measured with sentiment analysis. We then triangulate (Jick, 1979) these results with those from conventional quantitative measures that are applied to the same case and time period. The chosen entity under study is a major Italian bank, ItalyBank (pseudonym) that has recently evoked public criticism and praise for various issues. We analyze the judgments by 5.991citizens in 14.179 tweets. Additionally we analyze 722 newspaper articles about ItalyBank, monitor its evaluation in a national survey, and observe its institutional linkage to a credit rating agency.

Methodology

Sentiment analysis of tweets**:** To assess organizational legitimacy as the normative judgment by citizens in Twitter, we first collected tweets through Twitter's search API for a time period of one year (1st May 2013 to 30th April 2014). The keyword search resulted in totally 36.092 tweets in Italian language mentioning the word "ItalyBank". In order to analyze ongoing conversations that construct organizational legitimacy (Schultz et al., 2013), we selected those tweets that included a hashtag (e.g., #italybank or #statehelp). Furthermore, we excluded all tweets that were produced by ItalyBank's twelve Twitter accounts (totally 423 tweets). This selection process resulted in a total of 14.179 tweets by 5991 authors for 389 different hashtags, whereby each hashtag typically refers to a specific topic (see appendix A and B for examples).

In order to identify positive and negative judgments that construct organizational legitimacy (Haack et al., 2014), the sentiment of each tweet was detected through a machine learning approach based on a supervised classification model (Crammer et al., 2006). In a first step, to create the training set for the mathematical model, three individual coders coded the same 1459 tweets (10% of the dataset of re-tweets) for a unique sentiment value: negative sentiment (-1), neutral sentiment (0)[5], and positive sentiment (+1) towards ItalyBank (inter-coding reliability: Kalpha= .81, p= .026)[6]. In a next step we employed a Passive–Aggressive (PA) classifier (Crammer et al., 2006) implemented with a Pairwise Coupling with majority voting method (Hastie & Tibshirani, 1998). Pairwise coupling is an approach that combines different binary classifiers to obtain a multi-class classification (Platt, Cristianini & Shawe-Taylor, 2000). The quality of the feature extraction and classification model was confirmed by the experimental results obtained through 10-fold cross-validation on the training dataset: f-measure 0.75 and accuracy 0.8. The learned model and the feature extraction process were then employed on all tweets to compute the sentiment values. The sentiment analysis revealed that 34% percent of the 14.179 tweets were affectively charged, a rate that was found in previous Twitter studies (e.g., Hansen et al., 2011). Based on the idea that individuals bestow legitimacy to organizations through positive and negative judgments (Haack et al., 2014; Haidt, 2001) and following previous research for news media legitimacy (Deephouse, 1996; Deephouse & Carter, 2005), we calculated the Janis-Fadner coefficient of imbalance, ranging from -1 to +1, for monthly measures (Janis & Fadner, 1965). This coefficient gives indication about the overall evaluation of Italybank by citizens for the respective months.

News media content analysis: Because ItalyBank is a major national bank and therefore its actions are likely to be covered by national newspapers, we chose the two largest national newspapers of Italy based on the criteria of circulation and source authority (Deephouse, 1996): *Corriere della Sera* and *La Repubblica*. We searched Factiva database with the keyword "Italybank" for the case period, resulting in totally 1444 articles; 846 articles from *Corriere della Sera* and 598 articles from *La Repubblica*. For the study every second article was coded, resulting in a total of 722 articles. A recording unit was defined as the evaluation of ItalyBank in a single article. Consistent with past research (Deephouse, 1996; Deephouse & Carter, 2005) each recording unit was rated as either endorsing the bank or questioning its legitimacy and given equal weight in the subsequent analysis. A recording unit was rated as questioning when there was evidence that the bank's action, structure, mission, or performance was being questioned or challenged. Otherwise, the recording unit was rated as endorsing the bank. A subsample of 75 articles, or 10% of the sample, was coded by two coders resulting in a sufficient intercoder-reliability of 0.95. Monthly measures of legitimacy were calculated using the Janis-Fadner coefficient of imbalance (Janis & Fadner, 1965; Deephouse & Carter, 2005).

Accreditation: To observe institutional linkages (Baum & Oliver, 1991) we used secondary data provided by Standard & Poor's (S&P), one of the most influential credit rating agencies in the global (and Italian) financial sector (Langohr & Langohr, 2010). S&P provides expert-evaluation about the likelihood of (financial) institutions to meet their financial commitments, which are regularly reported through public reports. For this study we consulted the reports about the long-term credit ratings for ItalyBank that can range from D (lowest) to AAA (highest).

Survey: As a survey measurement we use secondary data provided by the RepTrak, one of the most used corporate reputation rankings in research and practice (Ponzi et al., 2011). The RepTrak assesses evaluations by the general public for totally seven aspects (“dimensions”). To measure organizational legitimacy, we selected the aspects “Citizenship” (items: “supports good causes”, “positive societal influence”, and “environmentally responsible”), “Governance” (“open and transparent”, “behaves ethically”, and “fair in way it does business”), and “Workplace” (“rewards employees fairly”, “employee well-being”, “equal opportunities”). The data were originally collected through computer-assisted web interviewing (CAWI) during the first quarter of the years 2013 and 2014, assessing the reputation of ItalyBank and additional four Italian banks. The sample includes 3000 respondents among Italian citizens and is representative for age and gender. Respondents rate companies on a 7-points likert-scale. The computed overall score ranges from 1 to 100 index points.

Results

The results of the sentiment analysis are displayed in figure 1 that shows the development of judgments in Twitter. The monthly measures indicate how the overall judgment, or sentiment, is persistently negative, with a monthly average of -0.29. Furthermore, a slight increase from May 2013 to April 2014 can be detected. In contrast, the results of the news media analysis, also displayed in the figure 1, reveals that the overall judgment calculated as monthly measures is mainly positive (i.e. endorsing), with a monthly average of 0.49 (see figure 1). Also here a slight increase can be identified with a significant drop in August 2013 (0.03). The survey-based results also reveal a slight increase of the overall evaluation from 56.5 index points in the first quarter 2013 to 57.5 index points in the first quarter 2014. Finally, the observation of institutional linkages reveals a BBB+ rating of ItalyBank in May 2013. On July 12th 2014 ItalyBank was downgraded to BBB with no further changes until end of April 2014.

Insert figure 1 about here

Overall, we find an increase of organizational legitimacy as reflected in the judgments of all evaluators during the studied time period, except for the expert-evaluation by the credit rating agency. News media more or less continuously endorse ItalyBank on a high level. However, we were also identified a significant drop for August 2013. In contrast to news media, a negative sentiment prevails in Twitter. This can be interpreted with the fact that, even though ItalyBank was only partially affected by the economic and financial crisis, citizens – in contrast to news media - have still and continuously blamed it in a myriad of conversations that took place on social media, for example under the hashtags #BankItalia and #badbank (see Appendix A and B). The evaluation by the general public represented in the survey can be interpreted, if not negative, as rather modest considering the possible maximum of 100 index points. Finally, the credit rating by S&P starts on the intermediate rating BBB+ and then decreases to BBB, indicating that ItalyBank is still found to have "adequate capacity to meet its financial commitments", yet to a lower degree than before. (Standard & Poor's, 2015).

The case study shows how the analysis of four different sources of organizational legitimacy reveals varying and different sorts of judgments. This is not surprising, because these sources base their evaluation on different criteria, norms, and standards and make them visible through different processes. These differences become apparent, when contrasting the mostly endorsing evaluation in news media with the negative sentiment in Twitter, the development of which we were able to monitor in a detailed manner. In contrast, the presented insights through the survey are based on judgments that are only published once a year, and hence, are more suitable for long-term studies.

Nevertheless, the survey based results somehow confirm the negative sentiment of the public in Twitter and give indication about the increase from one year to the next.

In sum, if we understand organizational legitimacy as constructed through the expression and negotiation of normative judgments by various evaluators, the triangulation of results from different data sources gives a multifaceted picture of organizational legitimacy and its development over time. Each data source on its own gives insights for the construction of legitimacy by respective evaluator, and each evaluator has its own particular influence on the perception of organizations. In concerted use the measures can complement each other, because they deliver complementary judgments from different angles, frequencies, based on different evaluation criteria, and with different time sensitivity.

Discussion

Because there is rarely a single measure or authority that can pronounce the judgment for the whole society (Bitektine & Haack, 2015), attention to various evaluators can enhance the scholarship around organizational legitimacy. Conventional methods, while powerful in measuring the evaluation by institutional evaluators, capture only a fraction of the possible voices, norms, and aspects in the evaluation of organizations. In this article we have discussed sentiment analysis of social media data as a method that gives access to the plurality of judgments by ordinary citizens. Our comparison with existing sources and the empirical analysis have shown how using this method can enrich our understanding of the legitimation of organizations "whose legitimacy is based on civil society discourses" (Palazzo & Scherer, 2006, p. 78).

The case study has illustrated how the number of over 14.000 tweets by almost 6000 citizens and categorized under 389 hashtags stands in contrast to a few, yet influential institutional evaluators and their judgments. The consideration of the many voices accounts for an understanding of organizational legitimacy that is based on a plurality of expressed opinions (Palazzo & Scherer, 2006; Schultz et al., 2013). Our study has shown how citizens in social media judge the actions of the studied organization more negatively than news media. Similarly, the decrease in citizens' judgment over time stands in contrast to the increase of positive judgment by accreditation bodies. These diverging evaluations illustrate how judgments of citizens are not necessarily captured by an analysis of institutional evaluators. As an additional method, then, sentiment analysis of social media data can contribute to a more encompassing understanding of legitimacy that is based on the judgments by various evaluators (Ruef & Scott, 1998).

Implications for Research

The use of sentiment analysis and large scale social media data is appealing and offers tremendous potential for various research avenues and theory development. The method allows directly accessing the voices of ordinary citizens, and therefore gives also better indication for individual perceptions (Bitektine, 2011) than measures that are solely based on institutional evaluators' judgments. Consequently, the method can be used to address a wide array of research questions in the business and society field.

While a sentiment-score can be criticized for being reductionist, it is at the same time appealing for researchers, because it allows testing of hypotheses. As a *dependent variable*, it might be of interest to investigate the antecedents of legitimacy in social media. For example: How do organizational strategies of legitimation (Scherer et al., 2013) or certain events impact organizational legitimacy? To answer these kinds of questions, the method can add additional value to existing measurements, because corporate actions and events, such as scandals, are increasingly discussed and evaluated in social media (e.g., Etter & Vestergaard, 2015). As our study has shown, organizations are discussed differently in social media than in news media, and these evaluations might not be detected by surveys that are conducted long after an event has occurred. Sentiment analysis, then, can give fruitful insights into the development of citizens' judgments about certain events.

Furthermore, over the last years corporations have invested heavily in social media channels in order to influence public opinion through various strategies (e.g., Castello et al., 2016; Etter, 2014; Illia et al., 2015). If and how citizens in social media are affected by these strategies or rather develop counter-discourses, where organizational actions are questioned and criticized, can be detected by sentiment analysis of social media data. Similarly, sentiment analysis allows analyzing citizen's perceptions of business, financial, and CSR initiatives, additional to conventional legitimacy measures. For example, while CSR initiatives might find not much attention by news media because of lacking newsworthiness, they might trigger enthusiasm (or critique) in social media. Measuring legitimacy in social media might therewith help to detect, if CSR initiatives meet heterogeneous expectations of citizens.

When using legitimacy as an *independent variable*, the relationship between citizens' judgments in social media and various organizational outcomes, such as financial performance, stock market value, sales, etc. might be of interest. Indeed, judgments in social media are used as information source by various stakeholders, give an indication about the affective orientation towards an organization, and hence might have predictive power for various forms of organizational performance. The big advantage of sentiment analysis, then, is that it can account for the dynamic development of citizens' judgments, which is more difficult to grasp with costly surveys. Nevertheless, a concerted use with other legitimacy measures, such as a news media based measurement, can detect if possible effects are stronger, when other evaluators are involved. For example, in the specific case of a triangulation of social media data with news media data, a researcher can assess whether judgments in social media only affect corporate performance, when these judgments are amplified by news media coverage.

Indeed, it has to be acknowledged that various evaluators interact and influence each other: judgments published in rankings or reports about accreditations and ratings might be covered by news media, and news media coverage can influence social media conversations, as well as vice versa (e.g., Pfeffer et al., 2013). These interrelations might be considered as mediating effects, and not as mere triangulation, when analyzing the relationships described above.

Research questions might be addressed through case studies or with representative samples, for example from a particular industry, which might give generalizable insights about the various relationships outlined above. Furthermore, quantitative sentiment analysis can be used as a starting point for qualitative analyses that investigate more in detail the affective dynamics of legitimation processes. Here, sentiment analysis of social media data allows accessing emerging legitimacy issues in a heterogeneous context, which might not be captured with other measures.

Finally, future research might also deepen the understanding of some of the theoretical underpinnings that have been presented in this article. For example, based on the notion of affect-based responses (Haack et al., 2014) that are expressed in social media, we have argued that the expression of affective judgments and opinions by ordinary citizens constitute a source of organizational legitimacy. The relationship between affect and legitimacy needs more conceptual groundwork, which, however, was beyond the scope of this article.

Implications for Practice

Gaining legitimacy is an increasingly complex process for organizations (Castelló et al., 2015). In today's world, organizations face a multiplicity of expectations and norms due to the expansion of corporate activities and the individualization of society (Scherer & Palazzo, 2006). It therefore requires a multi-faceted understanding of many concerns, voices, and conceptions of truth, and an ability to engage across independent and conflicting interpretations of issues (Schultz et al., 2013). Particularly in regards to CSR issues, legitimacy of organizations is challenged (e.g., Sharma & Henriques, 2005; Scherer & Palazzo, 2013). To address CSR issues stakeholder engagement has become an important organizational activity (e.g., Sharma & Henriques, 2005; Scherer & Palazzo, 2013) that is increasingly enacted through social media platforms (Etter, 2014; Illia et al., 2015).

Managing CSR and legitimacy through engaging in discourses is limited due to the decentralization of communication in indeterminate networks and the complex dynamics enfolded therein (Schultz et al., 2013). Our analysis has shown that news media clippings alone might not give indication about the concerns of ordinary citizens. In social media, ordinary citizens might discuss issues in a negative way long before – or after – they have been picked up by the news media or been detected by costly surveys. Monitoring social media through sentiment analysis therefore might be a valuable tool to detect (de-)legitimizing discourses around a variety of (CSR) topics. Additional to extant measures, sentiment analysis and the exploration of social media data can give organizations a multifaceted understanding of the many concerns and judgments that ordinary citizens express in social media. The presented study, for example, identified around 400 hashtags, or topics, to which citizens expressed their opinions. The identification of certain issues can be considered as a necessary first step for the co-creation of CSR agendas through participation in non-hierarchical, open social media platforms (Castelló et al., 2016). Furthermore, with its time-sensitive approach, organizations can use sentiment analysis for detecting (de-)legitimizing conversations and issues in a timely manner. Indeed, continuous monitoring is crucial, since conversations can emerge anytime and anywhere.

In sum, sentiment analysis can help to detect sentiment and important issues in social media. As an additional legitimacy measure to surveys, rankings, and media clippings, it gives immediate indication about the judgments of ordinary citizens represented in social media.

Limitations

The limitations of the presented method are the sampling and the delimitation of inherently connected data (discourses or actors). In our study we analyzed tweets that mentioned the bank's name, thereby possibly excluding tweets that make reference to the bank without explicitly mentioning it. Similarly, challenges arise from data cleaning processes, such as excluding texts from certain actors. In our study we found and excluded 423 tweets - less than 3% of originally collected tweets - stemming from the bank itself. There is, however, a possibility that further undetected tweets stem from other institutional evaluators. Furthermore, while we have no indication for our study, there exists the possibility of manipulation through public relations firms distributed positive messages about companies under fake profiles. Generally, manipulations can often be detected or, considering the large amount of sourced data, may not significantly impact results. Finally, the level of accuracy that algorithms can attain in the classification of sentiment in text is limited, because mining sentiment from natural language is challenging. However, techniques are constantly improving, and the accuracy levels with today's technologies, such as used in this study, range around 80%.

Conclusion

This article has shown how conventional quantitative methods for the measurement of organizational legitimacy are valuable for assessing judgments that are made visible by institutional evaluators. We have also shown how these sources give only limited indication about judgments by ordinary citizens that have been regarded as increasingly important from a normative perspective that calls for a stricter democratic accountability of corporate behavior (Matten & Crane, 2005). To assess these judgments we have proposed, critically discussed, and illustratively applied the use of social media data and sentiment analysis.

Our article does not favor one method over the other. We believe that each method has value on its own and provides a valid measure for organizational legitimacy. Sentiment analysis accommodates researchers and practitioners that aim at exploring the power of large scale data sets that include a multitude of conversations and judgments by ordinary citizens and therefore give indication about the broader fit between an organization's perceived behavior and heterogeneous social norms. Furthermore the method allows the investigation of developments in a very timely manner and therewith is suited for a communication environment that is characterized by speed and connectivity. Finally, by focusing on the content first and on the actors in a second step, the method is perfect for a context, where established and identifiable evaluators are not the only source of legitimacy, but where any actor with access to social media can evaluate corporations publicly and where conversations can emerge everywhere in a transnational network of actors that contribute to the (de-)legitimation of corporations. Alone or in combination with extant measures, the method is powerful to study organizational legitimacy as constructed through the expression and negotiation of normative judgment by ordinary citizens.

Acknowledgments

We thank Andrew Crane and the other editors of this special issue for their helpful guidance during the development of this article. Furthermore, we would like to thank Glen Whelan, Patrick Haack, and Alessandro Rozza for their fruitful comments and inputs for prior versions of this article, as well as the three anonymous reviewers for their constructive critique and comments.

Funding Acknowledgment

This work was supported by the Carlsberg Foundation; and the UniCredit & Universities Foundation.

[1] Hunter and colleagues (2008) call media, which stakeholder groups use to raise their voices, "stakeholder media".

[2] "Neutral" is either coded as the presence of positive and negative (e.g., Pollock & Rindova, 2003), or the absence of positive and negative (e.g., Lee & Carroll, 2011).

[3] These measures typically assess the evaluation of various organizational aspects and aggregate them to an overall score. Surveys are mostly used to measure organizational reputation (e.g., Fombrun, 2007), a concept that has a "substantial conceptual overlap" with organizational legitimacy (Deephouse & Carter, 2005, p. 330). It can be argued that the evaluation of aspects such as "support of good causes" or "environmental responsibility" also gives indication about organizational legitimacy.

[4] When dealing with social media, another valuable information about the affective orientation of a content is represented by emoticons (Pak & Paroubek, 2010). Emoticons and punctuations are key when detecting ambivalence, such as irony, sarcasm and negation, which remains one of the challenges for sentiment analysis.

[5] Neutral was coded as absence of negative and positive sentiment.

[6] The sentiment corresponds to "focal media favorability", which is directly linked to the firm, in contrast to "peripheral media favorability", which refers to the general tone independent from the evaluation of the organization (Carroll, 2009).

References

Adamic, L., & Glance, N. (2005). The Political Blogosphere and the 2004 U.S. Election: Divided They Blog. Proceedings of the 3rd international workshop on Link discovery, New York: ACM.

Agarwal, A., & Bhattacharyya, P. (2005). Sentiment analysis: A new approach for effective use of linguistic knowledge and exploiting similarities in a set of documents to be classified, In Proceedings of the International Conference on Natural Language Processing (ICON).

Albinger, H. S., & Freeman, S. J. (2000). Corporate social performance and attractiveness as an employer to different job seeking populations. *Journal of Business Ethics*, *28*(3), 243-253.

Bakker, T. P., & de Vreese, C. H. (2011). Good news for the future? Young people, Internet use, and political participation. *Communication Research, 38*(4), 451-470.

Bansal, P., & Clelland, I. (2004). Talking trash: Legitimacy, impression management, and unsystematic risk in the context of the natural environment. *Academy of Management Journal*, *47*(1), 93-103.

Baum, J. A., & Oliver, C. (1991). Institutional linkages and organizational mortality. *Administrative science quarterly*, 36(2), 187-218.

Blumell, L.E. (2016). *Birds of a feather…Gatekeeping human rights online news coverage via traditional and humanitarian news media*. PhD dissertation, Texas Tech University at Lubbock.

Bitektine, A. (2011). Toward a theory of social judgments of organizations: The case of legitimacy, reputation, and status. *Academy of Management Review, 36*(1), 151-179.

Bitektine, A., & Haack, P. (2015). The "macro" and the "micro" of legitimacy: Toward a multilevel theory of the legitimacy process. *Academy of Management Review*, *40*(1), 49-75.

Boyd, D. (2010). Social network sites as networked publics: Affordances, dynamics, and implications. In Z. Papacharissi (ed.), *Networked Self: Identity, Community, and Culture on Social Network Sites* (39-58). New York: Routledge.

Boyd, D., & Crawford K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, 15(5), 662–679.

Brown, N., & Deegan, C. (1998). The public disclosure of environmental performance information—a dual test of media agenda setting theory and legitimacy theory. *Accounting and business research*, *29*(1), 21-41.

Bruns, A., & Stieglitz, S. (2012). Quantitative Approaches to Comparing Communication Patterns on Twitter. *Journal of Technology in Human Services*, 30(3-4), 160–185.

Burgess, J., & Bruns, A. (2011). The Use of Twitter Hashtags in the Formation of Ad Hoc Publics. Paper presented at the European Consortium for Political Research conference, Reykjavik, 25-27 Aug. 2011.

Cambria, E., Mazzocco, T., & Hussain, A. (2013b). Application of Multi-Dimensional Scaling and Artificial Neural Networks for Biologically Inspired Opinion Mining. *Biologically Inspired Cognitive Architectures*, 4(1), 41–53.

Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013a). New Avenues in Opinion Mining and Sentiment Analysis. *Intelligent Systems*, 28(2), 15-21.

Carroll, C. E. (2009). The relationship between firms' media favorability and public esteem. *Public Relations Journal*, *3*(4), 1-32.

Carroll, C. E., & McCombs, M. (2003). Agenda-setting effects of business news on the public's images and opinions about major corporations. *Corporate reputation review, 6*(1), 36-46.

Castelló, I., Etter, M., & Årup Nielsen, F. (2016). Strategies of Legitimacy Through Social Media: The Networked Strategy. *Journal of Management Studies, 53(3), 402–432*.

Castelló, I., Morsing, M., & Schultz, F. (2013). Communicative dynamics and the polyphony of corporate social responsibility in the network society. *Journal of Business Ethics*, *118*(4), 683-694.

Castells, M. (2007). Communication, power and counter-power in the network society. *International journal of communication*, *1*(1), 29.

Chumley, C. (2014, November, 19). Barilla pasta maker — humbled by gay pressures — gives LGBTs rights. Retrieved Mai 21, 2015, from http://www.washingtontimes.com/news/2014/nov/19/barilla-pasta-maker-humbled-by-gay-pressures-gives/

Colleoni, E. (2013). CSR communication for organizational legitimacy in social media. *Corporate Communications: An International Journal*, 18(2), 228 - 248.

Crammer, K., Dekel, O., Keshet, J., Shalev-Shwartz, S., & Singer, S. (2006). Online passive-aggressive algorithms. *Journal of Machine Learning Research*, 7,551–585.

Crane, A., McWilliams, A., Matten, D., Moon, J., & Siegel, D. S. (Eds.). (2008). *The Oxford handbook of corporate social responsibility*. New York: Oxford University Press.

Davies, G., Chun, R., Da Silva, R. V., & Roper, S. (2003). *Corporate reputation and competitiveness*. London: Routledge.

de Souza, C.R.B., & Redmiles, D. (2009). On the Roles of APIs in the Coordination of Collaborative Software Development. *Journal of Computer Supported Cooperative Work*, 18(5/6), 445-475.

Dearing, J. W., & Rogers, E. (1996). *Agenda-setting* (Vol. 6). New York: Sage.

Deephouse, D. L. (1996). Does isomorphism legitimate?. *Academy of management journal*, *39*(4), 1024-1039.

Deephouse, D. L., & Carter, S. M. (2005). An Examination of Differences Between Organizational Legitimacy and Organizational Reputation. *Journal of Management Studies*, *42*(2), 329-360.

Etter, M. (2014). Broadcasting, reacting, engaging–three strategies for CSR communication in Twitter. *Journal of Communication Management*, *18*(4), 322-342.

Etter, M. A., & Vestergaard, A. (2015). Facebook and the public framing of a corporate crisis. *Corporate Communications: An International Journal*, *20*(2), 163-177.

Fieseler, C., Fleck, M., & Meckel, M. 20(10). Corporate social responsibility in the blogosphere. *Journal of Business Ethics*, 91(4), 599–614

Fombrun, C. J. (2007). List of lists: A compilation of international corporate reputation ratings. *Corporate Reputation Review*, *10*(2), 144-153.

Galtung, J., & Ruge, M. H. (1965). The structure of foreign news the presentation of the Congo, Cuba and Cyprus Crises in four Norwegian newspapers. *Journal of peace research*, *2*(1), 64-90.

Gangemi, A., Presutti, V., & Reforgiato, D. (2014). Framebased detection of opinion holders and topics: a model and a tool. *IEEE Computational Intelligence Magazine,* 9(1), 20– 30.

Gerlitz, C., & Rieder, B. (2013). Mining one percent of Twitter: collections, baselines, sampling, *M/C Journal*, 16(2), 1-12.

Gillespie, T. (2010). The Politics of Platforms. *New Media & Society*, 12(2), 347–364.

Haack, P., Pfarrer, M. D., & Scherer, A. G. (2014). Legitimacy-as-Feeling: How Affect Leads to Vertical Legitimacy Spillovers in Transnational Governance. *Journal of Management Studies*, *51*(4), 634-666.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological review*, *108*(4), 814.

Hansen, L. K., Arvidsson, A., Nielsen, F. Å., Colleoni, E., & Etter, M. (2011). Good friends, bad news-affect and virality in twitter. In *Future information technology* (pp. 34-43). Berlin, Germany: Springer.

Hargittai, E., & Litt, E. (2011). The tweet smell of celebrity success: Explaining variation in Twitter adoption among a diverse group of young adults. *New Media & Society*, *13*(5), 824-842.

Hart, R.P., & Carroll, C.E. (2015) *DICTION 7, The Text Analysis Program: Help Manual*, Austin:Digitext.

Hastie, T., & Tibshirani, R. (1998). Classification by pairwise coupling. *The Annals of Statistics*, 26(2), 451-471.

Helm, S. (2007). One reputation or many? Comparing stakeholders' perceptions of corporate reputation. *Corporate Communications: An International Journal*, *12*(3), 238-254.

Hermida, A. (2012). Social journalism: exploring how social media is shaping journalism. In E. Siapera & A. Veglis (eds.), *The handbook of global online journalism* (309-328). New York: John Wiley & Sons.

Highhouse, S., Broadfoot, A., Yugo, J. E., & Devendorf, S. A. (2009). Examining corporate reputation judgments with generalizability theory. *Journal of Applied Psychology*, *94*(3), 782.

Hindman, M. (2009). The Myth of Digital Democracy. Princeton, NJ: Princeton University Press.

Hoffman, A. J., & Ocasio, W. (2001). Not all events are attended equally: Toward a middle-range theory of industry attention to external events. *Organization science*, *12*(4), 414-434.

Hunter, M. L., Le Menestrel, M., & De Bettignies, H. C. (2008). Beyond control: Crisis strategies and stakeholder media in the Danone Boycott of 2001. *Corporate Reputation Review*, 11(4), 335–350.

Illia, L., Romenti, S., Rodrıguez-Cánovas, B., Murtarelli, G., & Carroll, C. E. (2015). Exploring Corporations' Dialogue About CSR in the Digital Era. *Journal of Business Ethics*, 1-20.

Janis, I. L., & Fadner, R. H. (1943). A coefficient of imbalance for content analysis. *Psychometrika*, *8*(2), 105-119.

Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American society for information science and technology*, *60*(11), 2169-2188.

Jenkins, H. (2006) Convergence Culture: Where Old and New Media Collide. New York: New York University Press.

Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative science quarterly*, *24*(4), 602-611.

Langohr, H., & Langohr, P. (2010). *The rating agencies and their credit ratings: what they are, how they work, and why they are relevant* (Vol. 510). New York: John Wiley & Sons.

Larsson, A.O. & Moe, H. (2012). Studying political microblogging: Twitter users in the 2010 Swedish election campaign. *New Media & Society*, 14(5), 729–747.

Lee, S. Y., & Carroll, C. E. (2011). The Emergence, Variation, and Evolution of Corporate Social Responsibility in the Public Sphere, 1980–2004: The Exposure of Firms to Public Debate. *Journal of Business Ethics*, 114(1), 115-131.

Lomborg, S., & Bechmann, A. (2014). Using APIs for data collection on social media. *The Information Society*, *30*(4), 256-265.

Martin, R.M., & Boynton L. A. (2005) From liftoff to landing: NASA's crisis communications and resulting media coverage following the Challenger and Columbia tragedies. *Public Relations Review* 31(2), 253–26
Marwick, A. E. (2005). 'I'm a Lot More Interesting than a Friendster Profile': Identity Presentation, Authenticity and Power in Social Networking Services. Conference Proceedings *Association of Internet Researchers 6*.0. Chicago, Il.
Marwick, A. E. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New media & society*, *13*(1), 114-133.
Matten, D., & Crane, A. (2005). Corporate citizenship: Toward an extended theoretical conceptualization. *Academy of Management review*, *30*(1), 166-179.
McManus, J. H. (1994). *Market-driven journalism: Let the citizen beware?* Thousand Oaks, CA: Sage Publications.
Meyer, J. W., & Scott, W. R. (1983). *Organizational environments. Ritual and rationality*. Beverly Hills, Ca: Sage.
Nekmat, E., & Gonzenbach, W. J. (2013). Multiple Opinion Climates in Online Forums Role of Website Source Reference and Within-Forum Opinion Congruency. *Journalism & Mass Communication Quarterly*, *90*(4), 736-756.
Palazzo, G., & Scherer, A. G. (2006). Corporate legitimacy as deliberation: A communicative framework. *Journal of business ethics*, *66*(1), 71-88.
Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and trends in information retrieval*, 2(1-2), 1-135.
Papacharissi, Z. (2009). The virtual sphere 2.0: The Internet, the public sphere, and beyond. In A. Chadwick, & P. Howard (Eds.), *Routledge handbook of internet politics* (230-245). Abingdon: Routledge.
Papacharissi, Z. (2014). *Affective publics: Sentiment, technology, and politics*. Oxford: Oxford University Press.
Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin UK.
Paßmann, J., Boeschoten. T., & Schäfer, M..(2013). The Gift of the Gab. Retweet Cartels and Gift Economies on Twitter. In K.Weller, A. Bruns, J. Burgees, M. Maahrt, & C. Puschmann (Eds.), *Twitter and Society.* New York: Peter Lang.
Pedersen, S., Razmerita, L. & Colleoni, E. (2014). Electronic Word of Mouth communication and consumer behaviour: an exploratory study of Danish social media communication influence, *Professional Communication Knowledge Management Cognition*, 5(1), 112-131.
Peterson, D. K. (2004). The relationship between perceptions of corporate citizenship and organizational commitment. *Business & Society*, *43*(3), 296-319.
Ponzi, L. J., Fombrun, C. J., & Gardberg, N. A. (2011). RepTrak™ pulse: Conceptualizing and validating a short-form measure of corporate reputation.*Corporate Reputation Review*, *14*(1), 15-35.
Pfeffer, J., & Salancik, G. R. (1978). *The External Control of Organizations: A Resource Dependence Perspective*. New York, NY: Harper and Row.
Phillips, N., Lawrence, T. B., & Hardy, C. (2004). Discourse and institutions. *Academy of management review*, *29*(4), 635-652.
Platt, C., Cristianini, N. & Shawe-Taylor, J. (2000) Large margin DAGs for multiclass classification. Proceedings of Neural Information Processing systems, pg. 547-553, MIT Press, Cambridge.

Pollock, T. G., & Rindova, V. P. (2003). Media legitimation effects in the market for initial public offerings. *Academy of Management Journal*, *46*(5), 631-642.
Qiu, L., & Welch, I. (2004). *Investor sentiment measures* (No. w10794). National Bureau of Economic Research.
Rao, D., & Ravichandran, D. (2009). Semi-supervised polarity lexicon induction. In Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics.
Rieder, B. (2012). The Refraction Chamber: Twitter as Sphere and Network. *First Monday*, 11(5).
Roblyer, M., McDaniel, M., Webb, M., Herman, J., & Witty, J. V. (2010). Findings on Facebook in higher education: A comparison of college faculty and student uses and perceptions of social networking sites. *The Internet and Higher Education*, *13*(3), 134-140.
Rogers, R. (2009). *The End of the Virtual – Digital Methods*. Amsterdam, Netherlands: Amsterdam University Press.
Ruef, M., & Scott, W. R. (1998). A multidimensional model of organizational legitimacy: Hospital survival in changing institutional environments. *Administrative science quarterly*, 42(3), 877-904.
Saif, H., He, Y., & Alani, H. (2012). Semantic sentiment analysis of twitter. In Proceedings of the 11th international conference on The Semantic Web. Boston, MA.
Sakaki, T., Okazaki, M., & Matsuo, Y. (2010) Earthquake shakes Twitter users: real-time event detection by social sensors. In Proceedings of the 19th international conference on World wide web , WWW '10, (pp. 851–860), New York, NY, USA.
Scherer, A. G., Palazzo, G., & Seidl, D. (2013). Managing legitimacy in complex and heterogeneous environments: Sustainable development in a globalized world. *Journal of Management Studies*, *50*(2), 259-284.
Schultz, F., Castelló, I., & Morsing, M. (2013). The construction of corporate social responsibility in network societies: A communication view. *Journal of business ethics*, *115*(4), 681-692.
Schultz, M., Mouritsen, J., & Gabrielsen, G. (2001). Sticky reputation: Analyzing a ranking system. *Corporate Reputation Review*, *4*(1), 24-41.
Sharma, S., & Henriques, I. (2005). Stakeholder influences on sustainability practices in the Canadian forest products industry. *Strategic Management Journal*, *26*(2), 159-180.
Sotiriadis, M. D., & van Zyl, C. (2013). Electronic word-of-mouth and online reviews in tourism services: the use of twitter by tourists. *Electronic Commerce Research*, *13*(1), 103-124.
Stieglitz, S. & Krüger, N..(2011). Analysis of Sentiments in Corporate Twitter Communication – A Case study on an Issue of Toyota. In ACIS 2011 Proceedings.
Stohl, C., Etter, M., Banghart, S., & Woo, D. (2016). Social Media Policies: Implications for Contemporary Notions of Corporate Social Responsibility. *Journal of Business Ethics, 20(4),* 431-447.
Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of management review*, *20*(3), 571-610.
Sunstein, C. R. (2009). *Republic. com 2.0*. Princeton University Press.
Swanson, D. L. (1999). Toward an integrative theory of business and society: A research strategy for corporate social performance. *Academy of Management Review*, *24*(3), 506-521.
Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology*, *62*(2), 406-418.
Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of applied psychology*, *4*(1), 25-29.

Van Deursen, A., & Van Dijk, J. (2011). Internet skills and the digital divide. *New media & Society*, *13*(6), 893-911.
Vergne, J. P. (2010). Toward a new measure of organizational legitimacy: Method, validation, and illustration. *Organizational Research Methods*, *14*(3), 484-502.
Wartick, S. L. (2002). Measuring corporate reputation definition and data. *Business & Society*, *41*(4), 371-392.
Wei, L., & Hindman, D.B. (2011). Does the digital divide matter more? Comparing the effects of new media and old media use on the education-based knowledge gap. *Mass Communication and Society, 14*(2), 216–235
Westerman, D., Spence, P. R., & Van Der Heide, B. (2014). Social media as information source: Recency of updates and credibility of information. *Journal of Computer‐Mediated Communication*, *19*(2), 171-183.
Westphal, J. D., & Deephouse, D. L. (2011). Avoiding bad press: Interpersonal influence in relations between CEOs and journalists and the consequences for press reporting about firms and their leadership. *Organization Science*, *22*(4), 1061-1086.
Whelan, G., Moon, J., & Grant, B. (2013). Corporations and citizenship arenas in the age of social media. *Journal of Business Ethics*, *118*(4), 777-790.
Zhang, Y., Shang, L. and Jia, X. (2015). Sentiment Analysis on Microblogging by Integrating Text and Image Features. *Advances in Knowledge Discovery and Data Mining Lecture Notes in Computer Science*, *9078*(1), 52-63.

Table 1: Overview of Sources of Organizational Legitimacy

Sources of Organizational Legitimacy				
Source	News media	Accreditation bodies	Surveys	Citizens in social media
Evaluator type and impact	Institutional evaluator with impact on public opinion about organizations (Carroll & McCombs, 2003) and influence on social media discourses (e.g., Etter & Vestergaard, 2015)	Institutional evaluator that gains news media attention; judgments are used by financial analysts, investors, decision makers, and other stakeholders (e.g., Pollock & Rindova, 2003)	Institutional evaluator that gains news media attention; judgments are used by decision makers, managers, and other stakeholders (Fombrun, 2007)	Individual evaluator that can gain news media attention (e.g., Pfeffer et al., 2013); used mainly by online peers due to high credibility (e.g., Banning & Sweetser, 2007)
Selection of organizations and evaluation criteria/norms of evaluator	Selection of organizations and their actions based on news value (e.g., Galtung & Ruge, 1965); evaluation based on journalistic standards, such as fact-checking; norms, such as objectivity; and processes, such as editorial meetings (e.g., Shoemaker & Reese, 2014)	Standardized and balanced expert evaluation of a selected range of organizations based on predefined criteria (Haack et al.., 2014), such as such as a bank's ability to protect depositor savings (e.g., Deephouse, 1996) or staff qualifications of healthcare institutions (e.g., Baum & Oliver, 1991)	Standardized (non-) expert evaluation of a selected range of organizations based on predefined criteria, such as positive social impact, governance, and environmental responsibility (e.g., Fombrun, 2007)	Non-standardized selection of organizations based on personal interest and involvement (e.g., Boyd, 2010); non-standardized evaluation, based on individual norms and feelings (e.g., Papacharissi, 2014; Haack et al., 2014)
Frequency of publication of judgments	Continuously through various (online) news media outlets	Typically yearly accreditation, licensing, renewal of membership; published through reports and additionally news media	Typically published in yearly rankings/indices, often through news media and magazines (e.g., *Fortune's Most Admired Companies*)	Continuously through various social media platforms
Indication for legitimacy development	Short- to long-term	Long-term	Mid- to long-term	Short- to long-term
Indication for citizens' judgments	Limited by news media agenda (McCombs, 2013); selected, direct expression of citizens' judgments through letters to the editor (Lee & Carroll, 2011)	Indirectly, at best; ordinary citizens hardly impact judgments of accreditation bodies, but they are influenced by their judgments (Baum & Oliver, 1992)	Limited by predefined evaluation criteria and organizations, for which respondents might have insufficient knowledge (Schultz et al., 2001)	Judgment expression in any tone or style; biased by various factors, such as self-censorship (Nekmat & Gonzenbach, 2013), exageration (e.g., Marwick, 2005), digital divides (e.g., Wei & Hindman, 2011)

Table 2: Quantitative Measurements of Organizational Legitimacy

Quantitative Measurement of Organizational Legitimacy				
Source of legitimacy	News media	Accreditation bodies	Surveys	Citizens in social media
Method to assess judgments	Content analysis	Archivals, longitudinal analysis of secondary data	Questionnaire	Sentiment analysis
What it does	Selects text and assigns meaning to a selection of words to assess favorability, no favorability, or neutrality of media coverage	Assesses presence/absence or different levels of compliance with regulator rules (e.g., membership/accreditation)	Assesses the evaluation of various organizational aspects by respondents	Assesses affective responses towards organizations that are expressed in texts
Sources of data	News media; articles are usually sourced through databases, such as Lexis Nexis or Factiva	Governments; charities; non-governmental organizations; associations	Representative survey panels; members of the general public or specific stakeholder groups	Social media data, such as tweets, blogs, Facebook posts, etc.; sourced through APIs and data crawlers
(Automated) elaboration of data	Coding software for automated coding of text with assistance of dictionaries/wordlists, such as DICTION* (e.g., Hart & Carroll, 2015); Coding software that supports human coding, such as Nvivo or Atlas (e.g., Etter & Vestergaard, 2015); Statistical software for data management and quantitative analysis, such as SPSS, STATA, SAS, or R (e.g., Martin & Boynton, 2005)	Statistical software for data management and statistical analysis, such as RATE (e.g., Sing, Tucker, & House, 1986)	Statistical software for data management and quantitative analysis, such as SPSS, SPSS, STATA, SAS, or R (e.g., MacMillan, Money, Downing, & Hillenbrand, 2005)	Software based on programming languages, such as Python, that are able to process huge amounts of data (e.g. Castello et al., 2015)
Unit of analysis	News media article (e.g., Deephouse & Carter, 2005)	Reports, rankings, ratings, institutional linkages (e.g., Baum & Oliver, 1992)	Survey items covering an organizational aspect (e.g., Helm, 2007)	Comments, such as tweets, facebook posts, blog posts, etc. (e.g., Castello et al., 2015)

Continuation Table 2: Quantitative Measurements of Organizational Legitimacy

Variable types	Binary variable, such as endorsing or questioning (e.g., Deephouse & Carter, 2005); Ordinal variables, such as positive/negative/neutral (Lee & Carroll, 2011); ordinal varables, such as qualification of adjectives, "good environmental actions", "bad environmental actions" (e.g., Brown & Deegan, 1998)	Binary variable about absence or presence of regulator's actions, accreditation, collaborative relations or registration of licenses (e.g., Ruef & Scott, 1998); ordinal variable (e.g., Deephouse & Carter, 2005)	Ordinal variable, typically likkert scale for various items representing organizational aspects/dimensions (e.g., Fombrun, 2007); from 3 (Highouse et al. 2009) to 51 items (Davies et al. 2003)	Scale variable with the sentiment score that variate from negative to positive (Colleoni, 2013)
Coefficient/index	Favorability index; Tonality index; media endorsement index (Janis-Fadner imbalance coefficient); raw legitimacy vector (Vergne, 2011); typically calculated on annual basis (e.g., Deephouse & Carter, 2005)	Longitudinal index; rankings over times; longitudinal accreditation index; registration/acquisition index (e.g., Deephouse & Carter; Oliver & Baum, 1992; Pfeffer & Salancik, 1978)	Typically yearly ranking; overall score/index calculated based on evaluation of various items/organizational aspects (e.g., Fombrun, 2007; Helm, 2007)	Sentiment score/index calculated based on the detected sentiment for units of analysis. (e.g., Castello et al., 2015)
Examples of data sets	News media articles of regional newspapers (e.g., Deephouse & Carter, 2005), national newspapers (e.g., Brown & Deegan, 1998), or mix of national and regional newspapers (e.g., Lee & Carroll, 2011)	Reports by regulators about sanctions against organizations (e.g., Deephouse, 1996); public information about licenses/registers/certifications (e.g., Baum & Oliver, 1991; Pfeffer & Salancik, 1978)	Representative samples of the general public or particular stakeholder groups (e.g., Fombrun, 2007)	Facebook posts from the "wall" of a corporate Facebook site; blog posts and comments about organizations; Tweets about organizations (Castello et al., 2015; Collenoni, 2013)

Continuation Table 2: Quantitative Measurements of Organizational Legitimacy

Examples of data volumes (samples)	110 coded news media articles (Brown & Deegan, 1998); 460 coded news media articles (Lee & Carroll, 2011); 1277 coded news media articles (Deephouse & Carter, 2005)	Accreditation of 143 organizations by 7 accreditation bodies (Ruef & Scott, 1998); Accreditation of 1028 organizations by one accreditation body (Baum & Oliver, 1991)	Evaluation of global companies by aprox 4000 executives (Fortune World's Most Admired Companies); Evaluation of 2000 organizations across 15 countries by aprox 100 respondents per country (RepTrak)	43.000 tweets (Castello et al., 2015); 326.000 tweets (Colleoni, 2013)
Examples of time periods	5 years (Deephouse & Carter, 2005); 10 years (Brown & Deegan, 1998); 25 years (Lee & Carroll, 2011)	5 years (Deephouse, 1996); 17 years (Baum & Oliver, 1991); 46 years (Ruef & Scott, 1998)	5 years (Brammer & Millington, 2005); 15 years (Roberts & Dowling, 2002)	6 months (Colleoni, 2013); 41 months (Castello et al., 2015)
Examples of research objectives	Effect of news media legitimacy on IPOs (Pollock & Rindova, 2003) or on strategic communication (Brown & Deegan, 1998); effect of downsizing on organizational legitimacy (Lammertz & Baum, 1998)	Effect of organizational legitimacy on survival of organizations (Baum & Oliver, 1992); Effect of isomorphism on organizational legitimacy (Deephouse & Carter, 2005)	Effect of executives' judgments (FMAC) on financial performance (Roberts & Dowling, 2002); effect of CSR on public evaluation (Brammer & Millington, 2005)	Outcomes of communication strategies on organizational legitimacy (Colleoni, 2013; Castello et al., 2015)

*DICTION uses thirty-one wordlists to compute pre-defined formulas that analyse the text based on five master variables use– i.e.Activity, Optimism, Certainty, Realism and Commonality (Hart & Carroll, 2015). For instance, optimism is computed by "standardiz[ing] six variables and then add[ing] or subtract[ing] them (e.g., [praise + satisfaction + inspiration) - (blame + hardship + denial])" (Hart & Carroll, 2015, p. 4).

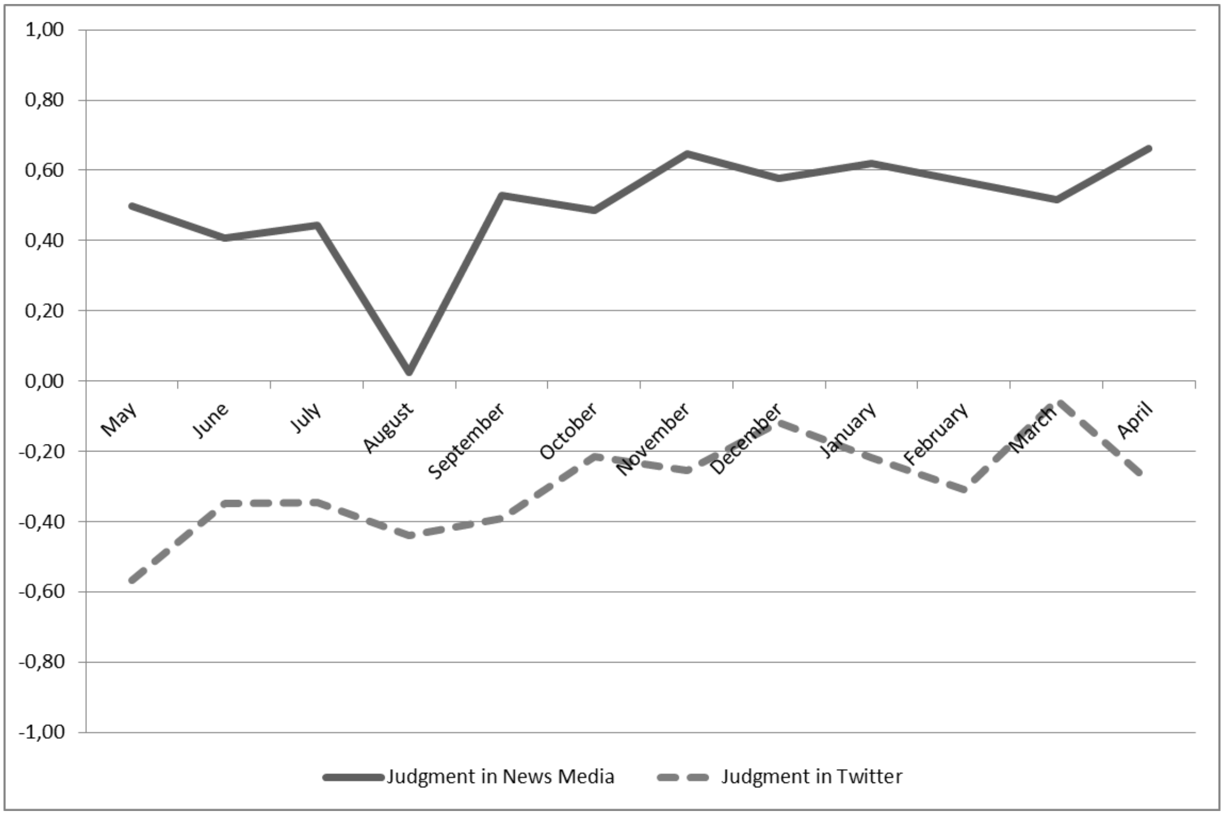
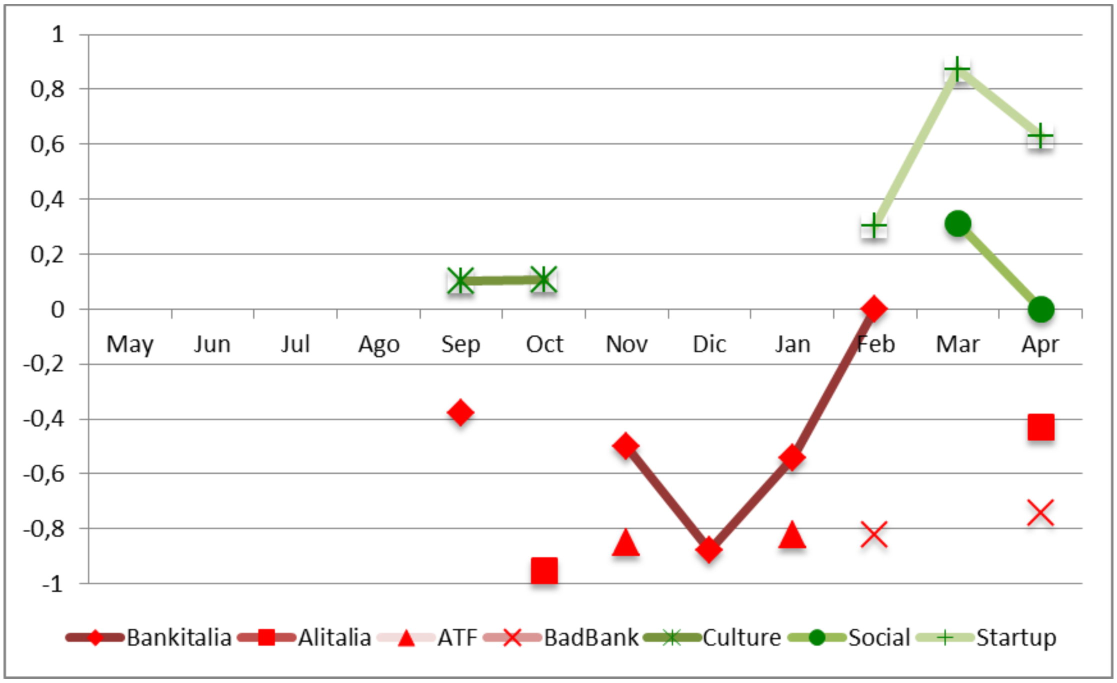


Figure 1: Monthly Measures (Janis-Fadner coefficient) of normative Judgments about Italy Bank in News Media and Twitter 1st May 2013 - 30th April 2014

Appendix A: Selected Twitter Conversations under particular Hashtags with varying Sentiment



Appendix B: Selected Hashtags and Explanation

Hashtag*	**Issue / Conversation**
#bankitalia	ItalyBank was involved in the decree of Bank Italia, the National Bank of Italy, which has re-financed its capitals through national reserves for the amount of about 4 milliards Euro. As ItalyBank is one of the two major owners of BankItalia, people assumed that part of that money went to ItalyBank too .
#alitalia	The Alitalia issue points to involvement of the bank in the recent sauvetage of Alitalia, the Italian airplane national company, an event which attracted ample attention and generated heated discussions and conflicting opinions around it.
#atf	The acquisition of an Ukrainian bank by ItalyBank implied a big speculation. The financial loss through the deal evoked negative sentiment.
#badbank	The intention to create an ad-hoc institution with other major Italian banks to externalize risks evoked mostly negative sentiment.
#culture**	The bank is repeatedly mentioned for its cultural initiatives, such as the art exposition in Bologna organized with artworks owned by ItalyBank.
#social**	Under the hashtag #social, social initiatives like the collection of money in favor of flooded populations are discussed.
#startup	Highly appreciated and shared information about bank economically supporting startups and innovative technological ideas

* hashtags are originally in Italian language and translated into English for the purpose of this study
** for these hashtags ItalyBank launched social media campaigns that promoted the positive discussions and sentiment