International Public Management Journal

ISSN: 1096-7494 (Print) 1559-3169 (Online) Journal homepage: https://www.tandfonline.com/loi/upmj20

Barriers for Crowd's Impact in Crowdsourced Policymaking: Civic Data Overload and Filter Hierarchy

Kaiping Chen & Tanja Aitamurto

To cite this article: Kaiping Chen & Tanja Aitamurto (2019) Barriers for Crowd's Impact in Crowdsourced Policymaking: Civic Data Overload and Filter Hierarchy, International Public Management Journal, 22:1, 99-126, DOI: 10.1080/10967494.2018.1488780

To link to this article: https://doi.org/10.1080/10967494.2018.1488780

Accepted author version posted online: 29 Jun 2018.
Published online: 16 Oct 2018.

International
Public
Management
Journal

Check for updates

BARRIERS FOR CROWD'S IMPACT IN CROWDSOURCED POLICYMAKING: CIVIC DATA OVERLOAD AND FILTER HIERARCHY

KAIPING CHEN AND TANJA AITAMURTO
STANFORD UNIVERSITY

ABSTRACT: *While crowdsourcing is an increasingly common method of open-government practices to strengthen participatory democracy, its impact on governance is unclear. Using data from a crowdsourced city-plan update by the City of Palo Alto, California, this article examines the impact of a crowd's input on policy changes. We used an enacted policy change to quantify government's response to crowd suggestions, whether crowd suggestions are adopted in the policy changes or not. While the city responded to less than half of the crowd's suggestions, the likelihood of its doing so increased by 51.42 percentage points when the crowd's ideas were amplified by a citizen advisory committee (CAC), a panel of residents working with the city in the policy update. We also found that the government is more likely to respond to crowd suggestions that are perceived as actionable. These two factors—CAC and the perceived data quality—constitute a filter which the crowd's suggestions have to pass to make into the policy. This filter created a hierarchy in the participatory practice. Although crowdsourcing intends to create equality and inclusiveness in policymaking, our findings reveal that the civic data overload and the filter hierarchy complicate the adoption of crowdsourcing as a democratic innovation in governance.*

INTRODUCTION

Crowdsourcing in policymaking has been heralded as a democratic innovation that holds potential for contributing to a more participatory democracy (Pateman 1970, 2012; for democratic innovations, see Smith 2009). In crowdsourcing, governments ask citizens to put forth their ideas, knowledge, and opinions regarding a policy (Aitamurto and Landemore 2015, 2016; Noveck 2015). Crowdsourcing, when conducted on digital platforms, has a dual advantage over many other civic engagement practices: it reaches citizens (1) in large numbers and (2) at low cost.

International Public Management Journal, Vol 22(1), pages 99–126
DOI: 10.1080/10967494.2018.1488780

ISSN: 1096-7494 print/1559-3169 online

This has led to more efficient knowledge-search and -discovery practices in several realms, such as crowdsourcing patent information (Noveck 2009) and solutions for companies' R & D problems (Jeppesen and Lakhani 2010). More recently crowdsourcing has been introduced to democratic practices, including policymaking (Aitamurto 2012; Aitamurto and Saldivar 2017). In crowdsourced policymaking, the goal is to use the crowd's input as an additional data point in the policymaking process, not to adopt all of the crowd's suggestions.

In addition to its epistemic import, crowdsourcing can contribute to a more responsive, collaborative, and transparent relationship between government and citizens in the hope of decreasing the "democratic deficit" facing representative governance (Fung 2006). Participatory democracy and responsive government are mutually reinforcing (Fung and Wright 2001; Gaventa 2004). Participatory practices can lead to more effective governance by advancing the public good and achieving a fairer and more equitable distribution of public services. On the other hand, some scholars (Fung 2015; Konisky and Beierle 2001; Michels and De Graaf 2010) argue that crowd participants exercise little real influence over the outcomes and, thus, that their participation is trivial.

To harness the potential of crowdsourcing, the government has to adopt crowdsourcing into the established policymaking cycle and overcome challenges such as incompatibility with the current organization and practices (Rogers 1962). The government must also develop strategies and models for how to use new methods and technologies (Chun et al. 2010). These challenges have become more common in public management in the era of technological innovations, including digital democracy practices and processing of "big data" (Chun et al. 2010; Thomas and Streib 2003).

Despite the proliferation of crowdsourcing in policymaking, its impact as a democratic practice in government policymaking remains unclear. This article focuses on addressing the following research questions: To what extent does crowdsourced citizens' input matter in policymaking? What factors affect government responsiveness in crowdsourced policymaking? Addressing these questions will enable us to understand the role of crowdsourcing as a participatory practice in policymaking and the factors complicating its use and adoption in government.

We use the concept of "government responsiveness" to examine to what extent crowdsourced citizens' input matters in policymaking. We investigate government responsiveness in crowdsourced policymaking drawing on data from a crowdsourced policymaking process in the City of Palo Alto, California. In 2015 and 2016, the city of Palo Alto crowdsourced ideas and opinions online from its residents for a comprehensive city plan (CCP) update. CCP is a fundamental urban-planning blueprint covering transportation, housing, and growth management. The Palo Alto City residents were invited to share their ideas on these topics online and to comment on policy drafts.

In the first section of this article, we review key concepts and literature about crowdsourcing and civic participation. We then introduce the case profile of the crowdsourced urban-planning-strategy process in Palo Alto. In the second section, we present our data and methods. The third section delineates findings, limitations, conclusions, and an agenda for future research.

THEORETICAL FRAMEWORK AND KEY CONCEPTS

Digital Crowdsourcing in Public Policymaking

By applying digital crowdsourcing to policymaking, local and national governments seek to gather facts, to improve policy, and to engage citizens. Crowdsourcing is an open call for anyone to participate in an online task by sharing information, knowledge, or talent (Brabham 2013; Estellés-Arolas and González-Ladrón-de-Guevara 2012; Howe 2008). Crowdsourcing reaches out to large, diverse crowds at low cost, which can make it an efficient knowledge search method (Aitamurto and Landemore 2016).

In theory, crowdsourcing creates *democratic*, *epistemic*, and *economic* value in policymaking by making it more inclusive, transparent, and collaborative, while enabling the government to access citizens' needs efficiently (Aitamurto and Chen 2017; Janssen et al. 2012). Crowdsourcing's transparency, accountability, inclusiveness, deliberativeness, and civic empowerment create *democratic* value. The knowledge sharing and development create *epistemic* value. Finally, crowdsourcing can create *economic* value by providing the government fast and cost-efficient access to citizens' needs, helping it to design better-fitting policies.

Iceland and Finland have experimented with crowdsourced lawmaking (Aitamurto and Landemore 2015; Landemore 2015). Iceland crowdsourced a part of its constitution reform process in 2010 to gather people's ideas and comments for the constitution drafting process. The constitution committee published drafts of the reformed constitution online, and people were invited to comment on the drafts. The partially crowdsourced constitution was deemed as more progressive than Iceland's then-current constitution, and the Icelandic people voted for its acceptance. However, due to political controversies, the constitution has not been accepted to date (Landemore 2017). In Finland, the Ministry of Justice crowdsourced information about problems in housing companies and the law governing them in 2014 and 2015. The law in question was the Limited Liability Housing Company Law. The Ministry invited citizens to participate in multiple stages of the policymaking process. The crowdsourced input was implemented in policy measures to address the issues (Aitamurto and Saldivar 2017). In the United States, the national government has applied crowdsourcing as a part of strategy reforms in federal agencies, such as EPA and FCC, and local governments have used crowdsourcing to detect topics for policy agendas and urban planning. In California, Lieutenant Governor Gavin Newsom crowdsourced people's input for a policy agenda in 2013. Based on the crowd's input, public awareness campaigns about the detected issues, such as disaster preparedness, were conducted (Aitamurto and Chen 2017; Nelimarkka et al. 2014; Goel et al. 2015).

In crowdsourced policymaking, governments typically ask citizens to contribute their ideas, knowledge, and opinions. The crowd's input is then synthesized and channeled into policy (Aitamurto 2012). The crowd can be invited to participate in one or several steps in the policymaking cycle, which has multiple stages: problem identification and definition, data gathering, developing options and proposals, consultation, designing and drafting the policy, decisions, evaluation, and implementation (Edwards 2001; Howlett, Ramesh, and Perl 1995; Peters 1999). Crowdsourcing is not typically utilized during the decision-making and implementation stages. The elected representatives still decide policy, even when the crowd participated in earlier parts of the process. In local government, this is the city council. At the national level, the elected representatives—in the House of Representatives and the Senate in the United States and the parliament in most European countries—make the final decisions. Crowdsourcing is not, therefore, a method of *direct* democracy (Frey 1994), in which citizens or residents directly determine policy. (Direct democracy does appear in democratic innovations such as participatory budgeting, in which participants have final say over a portion of the city budget.) Crowdsourcing, rather, is a normally method of *participatory* democracy (Pateman 1970, 2012) in that citizens participate in a more limited way, potentially influencing policy by providing additional information for policymakers, for example.

Crowdsourcing is an asynchronous, distributed, and depersonalized participation mechanism (Aitamurto and Landemore 2016) based on self-selection, enabling the dual benefit of reaching out to a large number of participants fast and with low cost. It provides citizens the flexibility to choose whether, how, and where to participate, and for how long. In these features, crowdsourcing differs from several other democratic innovations, such as deliberative polling and citizen juries (Fishkin 2011), which require an on-site and rather laborious participation procedure. These mini-public models of participation differ from crowdsourcing also in that they apply random sampling to ensure representativeness of the participant group. Their goal is to gauge the public opinion of the recruited mini-publics, assuming that reflects the opinion of the broader public. Deliberative polling also offers a predefined set of options to respondents that involve gauging "public opinion" or voting preferences, whereas crowdsourcing is applied in earlier stages in the policy cycle to search for solutions when the options are being developed. Crowdsourcing is thus similar to town hall meetings, public hearings, and consultations in which everybody is invited to participate, though only some do (Aitamurto and Landemore 2016).

Since crowdsourcing is based on self-selection, the method is not representative of the broader constituency and might, at first blush, be considered biased. However, when evaluated based on the usefulness of proposed solutions, it is not. The method leads to a set of proposed options, complemented by those of policymakers themselves. Ideally, of course, the process should be as inclusive as possible in order to maximize the breadth and efficiency of the knowledge search and to enhance democratic aspects in participation (Aitamurto and Landemore 2016).

Government's Responsiveness to Civic Participation

In order to use the crowdsourced input to benefit from its epistemic value, the government must analyze and synthesize the crowd's submissions. The larger the crowd, the more onerous the task, since the crowd's input is often atomic, divergent, and heterogeneous in content and format (Aitamurto 2016). As a result, it may ultimately be ignored by the government. This results in a "trivial" role for the crowd's input, thus following the unfortunate path of other participatory methods in which citizen voices are passed over or inadequately valued (Macnamara 2016; Fung 2015; Couldry 2010; Michels and De Graaf 2010; Konisky and Beierle 2001).

"Triviality" means that participants exercise "little influence over outcomes, the agenda of issues that they consider can be highly constrained, or the resources and authorities invested in a participatory process can be tiny" (Fung 2015:521). Therefore, even though citizen participation in democratic practices has increased, it has not been utilized in a meaningful way. This triviality has led to "a disappointment" among participants and a failure to meet the goals of increased legitimacy, effective governance, and social justice.

Empirical evidence has proved citizens with limited influence on policy outcomes. For instance, Michels and De Graaf (2010) studied two participatory democracy reforms in the Netherlands, discovering that individual citizens played only a minor role; citizen participation in policymaking did not lead to a fundamental new division of roles between citizens and politicians. Scholars such as Konisky and Beierle (2001:823) directly questioned whether citizen participation is worth the effort since "these processes [have] limited efficacy in changing policy, as most have only addressed issues outside the context of an actual policy decision." Similarly, Aitamurto (2016) found out that crowdsourced input in a Finnish off-road traffic-law-reform process remained unused by the government.

A key reason why citizen participation is undermined is that citizens' voices are not properly listened to (Couldry 2010). Rosanvallon and Goldhammer (2008:13) criticized the democratic practices as lacking the means by which citizens' voices can be valued within processes of policy development. Therefore, although governments increasingly invite citizen input, "little attention has been given to what listening involves: what would it mean both for governments to listen better to citizens and for citizens to listen better to each other?" (Couldry 2010:146). It is important for government to acknowledge, consider, and respond to citizen voices, rather than simply listening passively (Macnamara 2016:41–42). Responsive governance and participatory democracy could be mutually reinforcing, since a more responsive government can increase citizens' engagement in policymaking, and more engaged citizens can in turn benefit governance (Gaventa 2004). Responsive governance means not just listening but taking action by responding to citizens' needs, either through communication or by changing policies. Otherwise, citizens' voices will not affect policymaking.

Despite the prevailing pessimism surrounding citizen's impact on policymaking, some scholars have found evidence that the government sometimes considers public preferences in participatory policymaking. Kuklinski and Stanga (1979) found that the California Superior Courts responded to expressions of public preferences in the 1970s. After studying five experiments in the United States in the recent decade, Fung and Wright (2001) argued that participatory governance can help effective problem solving. Beyond the United States, Aitamurto and Saldivar (2016, 2017) examined a crowdsourced lawmaking process—namely, a reform of Limited Liability Housing Company Law in Finland. There, the civil servants evaluated the crowd's input and deemed it helpful in understanding the details and the scope of the problems in housing companies. The crowd's feedback was then incorporated into policy measures that were implemented shortly after the crowdsourcing process (Aitamurto and Saldivar 2017).

Reflecting upon the current literature studying citizens' impact on public policymaking, there are two gaps. First, the current literature largely regards government responsiveness as a black-and-white variable in which citizen input is or is not incorporated into final policy. However, responsiveness is rather a non-dichotomous variable—a matter of degree. The reason is that, when examining the impact of crowdsourcing, simply asking whether or not the government responded can be misleading due to the self-selected nature of crowdsourcing participation. In crowdsourced policymaking, the crowd participates voluntarily. Its participation is based on self-selection, and thus it does not form a representative sample. Furthermore, the crowd is not deciding about the policy. The government should take these aspects into account when analyzing the crowd's input, weighing the pros and cons of the suggestions and choosing to what extent to adopt the suggestions rather than simply responding or not responding. Therefore, it is more appropriate to ask *to what extent* government responds to citizens' voices. Second, a crucially related but understudied question is *what factors* determine government's degree of responsiveness. To review the factors that could help answer this question, we draw upon literature from innovation management and participatory/open government.

Challenges of Digital Democratic Innovation in Governance

Scholars have raised several key factors that determine whether an innovation will be adopted successfully (i.e., the extent of government responsiveness to an innovation).

First, as Rogers (1962) outlined in his work about diffusion of innovations, the adoption cycle follows a typical process before innovations are ultimately adopted—or rejected—by individuals or organizations. The rate of an innovation's adoption is influenced by at least five factors: (1) the perceived relative advantage of the innovation; (2) compatibility; (3) complexity or simplicity; (4) trialability; and (5) observability (Rogers 1962). In order to successfully implement a democratic innovation such as digital crowdsourcing, it must be compatible with the current governance processes so the government can add it to the established policymaking cycle; the government must perceive crowdsourcing's advantage; and the innovation itself must be not too complex or hard to implement. The adoption process can present challenges to government's practice and culture, since the government's mechanisms usually are not designed to adapt to changes and innovation (Besley and Ghatak 2003; Janssen and van der Voort 2016).

The second key factor is the data challenge that participatory open-government practices create. Crowdsourcing invites large-scale participation that is enabled by new technologies (Surowiecki 2005), similar to other digital two-way communication methods with citizens (Macintosh 2004; Bekkers and Homburg 2005; Chun et al. 2010). The rise of digital communication has resulted in large amounts of civic data, and these advancements push the government to be more transparent during policymaking processes. However, the scale and complexity of the civic data pose challenges for governments experimenting with digital democratic innovations. Governments need to develop systematic methods for listening and responding to participants (Macintosh 2004), overcome the lag between data collection and analysis (Mergel et al. 2016:934), understand the large amount of unstructured civic data (Chun et al. 2010; Janssen et al. 2012; Pirog 2014), and take into consideration the debatable quality of the citizens' input (Janssen et al. 2012). Without addressing these challenges, the government cannot develop a sufficient understanding of the crowd's input, harness its benefits, or meaningfully respond to it—and the innovation will fail.

The third key factor is the context—what participatory mechanisms are used in these innovations. Font et al. (2016, 2017) analyzed the adoption of 611 proposals that emerged from 39 participatory processes in Spain and found out that the participatory process—whether the proposals are generated from participatory budgeting or citizen councils or strategic planning or other temporary processes—has a significant influence on whether a proposal is implemented or not. This indicates that governments tend to choose inputs from certain participatory mechanisms more than others to inform their policy changes.

Moreover, for democratic innovation such as crowdsourcing, adoption also depends on how government perceives the quality of citizen suggestions (Font et al. 2016; Yang and Callahan 2005). Government responsiveness might not be determined by how many times a suggestion for a policy is raised but by what the government thinks of the suggestion: whether a suggestion is easy to implement or not and whether a suggestion is too broad or narrow. If a suggestion is too hard to implement—even if many citizens propose it—the government is less likely to respond (change the policy) because it does not have the capacity (i.e., budget, human resources) to implement it. If a suggestion is too broad (for instance, "transportation in Palo Alto should be improved"), the government is also less likely to respond since it is hard for the government to know what the citizens are really requesting. If a suggestion is too narrow (for instance, focusing only on the improvement of a specific street), the government might not change the policy because improving a certain street would not benefit a larger group.

Finally, media attention to the issue also influences the extent to which government adopts and responds to democratic innovation (Besley and Burgess 2001; Yanovitzky 2002; Qin et al. 2017). Issues that receive high media coverage can serve as the engine to drive government to take action to respond.

These factors provide us with the theoretical foundation to examine our research question on the factors that determine government's degree of responsiveness to citizen inputs.

Research Purpose

In this article, our overarching goal is to contribute new knowledge to the understudied issue of citizens' impact on crowdsourced policymaking (Aitamurto 2016; Aitamurto, Landemore, and Saldivar Galli 2017; Brabham 2013; Lee et al. 2014). To achieve this goal, we examined two related research questions: What is the impact of crowdsourcing on policymaking? What are the factors determining the government's responsiveness?[1]

The majority of existing studies employ interviews as the main method to analyze whether or not the government has adopted citizen suggestions. There is a lack of systematic analyses measuring the impact of civic participation (Font et al. 2016, 2017). We advance this situation through using mixed methods. We first use natural language processing to understand the main topics in crowd suggestions, citizen representatives' suggestions and the revised policies, and to help construct our independent variable. Although this text analysis method is useful in helping us explore the main concerns across three datasets, it is not adequate to answer directly the two research questions we posed. Therefore, we chose logistic regression in order to statistically test the extent to which a government responds to citizen suggestions and what factors matter for government responsiveness. We further confirmed and elaborated our findings from the logistic regression through interviews and participant observation. The interviews and participant observation are used to provide the reasons to explain the findings from the logistic regression. This mixed-methods approach enabled us to reveal a hierarchal filter system that arises in crowdsourced policymaking as a result of the bottleneck of civic data overload.

CASE PROFILE AND DATA

Case Profile

The empirical context of this article is the crowdsourced comprehensive city plan (CCP) reform in the City of Palo Alto, California.[2] Since August 2015, Palo Alto has been crowdsourcing residents' suggestions online for solving issues in the city. The issues included transportation, housing, land use, natural resources, business, and economics. The CCP covers these issues in sections that are called "elements" in the policy document. The elements contain programs and policies for the city to undertake. The CCP is a strategy that the city plans to use for the next 15 years, and it was last updated in 2002. The city has used a digital platform—Digital Commenter[3]—for crowdsourcing people's suggestions (i.e., feedback) for the existing (i.e., original) policy document. The city also established a "citizen advisory committee" (CAC), a voluntary panel to represent the residents of Palo Alto in the CCP. The CAC discusses policy reform and the related civic input in monthly meetings as a means of helping the city develop policy. After one year of online crowdsourcing and CAC meetings, the government finished revising the transportation element for the CCP in June 2016.

Data

The data used in the analyses are from the transportation element of the CCP update. We chose the transportation element for the analysis because it is the first element that has gone through a full cycle of feedback and revision (August 2015 to June 2016). Our data consist of three parts (see Figure 1). The first is the crowd's suggestions on transportation policy. They were exported from the Digital Commenter[4] in the xsl. file format. During the one-year crowdsourcing process, the government asked citizens to make suggestions on the original transportation policy twice. The first crowdsourcing sequence lasted six months, and the crowd provided 210 comments. The second crowd-sourcing lasted one month, and the crowd provided 33 comments. An example of a citizen comment left online was:

> Use trees as a traffic calming measure. Extensive research shows that that trees and landscape should be an integral part of the safety management of urban roads as they contribute to a safer street, with mid-block crashes, fewer injuries and fatalities.

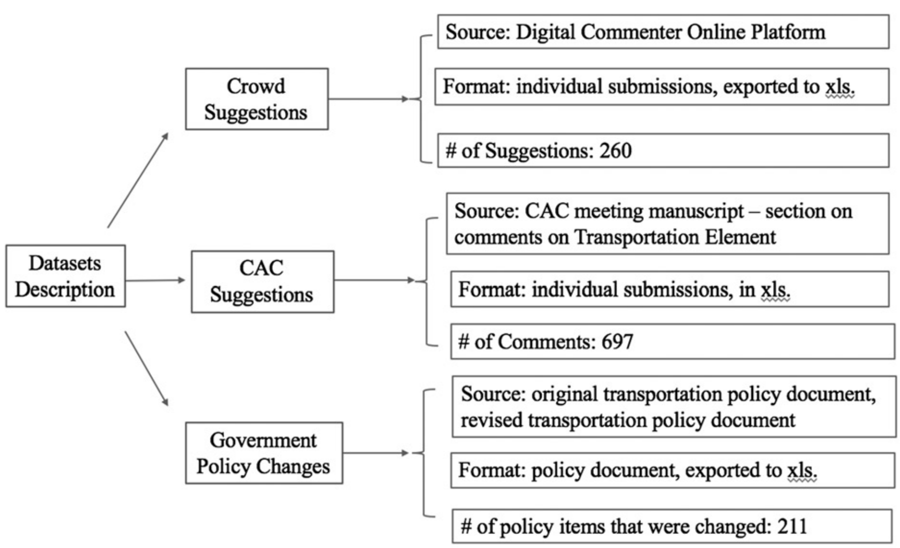


Figure 1. Description of the three datasets.

Sometimes, one citizen comment contained several suggestions, like the example which follows. Since we wanted to capture each suggestion the citizens raised, we divided this type of comment into separate suggestions and put them into different rows in a spreadsheet to ensure that we captured all of the suggestions from the citizens:

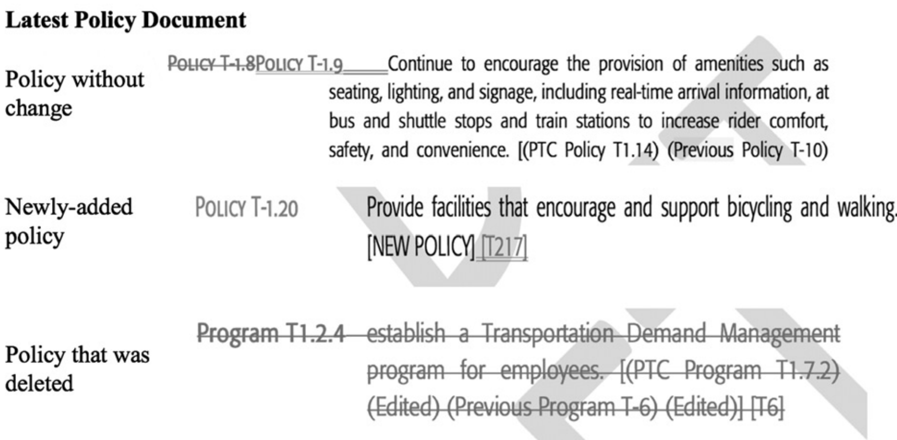
> (1) The parking lots near Cowper and University used to be empty during the day, presumably because they were reserved for monthly permits—I suggest increasing day-use, shared space (permit for different times during the day), etc. (2) Caltrain is wonderful if you work near a station; we have enough large employers who could provide a shuttle to outlying areas for their employees, and allow (paid or free) use for others as well.

Some comments were not suggestions on the transportation policy but focused on how people felt about using the online crowdsourcing platform. We excluded those comments. As a result, a total of 260 suggestions constituted the crowd data. Each row in the crowd spreadsheet represented a suggestion.

The second part of our dataset was the CAC's comments on the transportation policy. We downloaded their comments from the CAC website.[5] The CAC members collected suggestions from their neighborhoods, read online citizen suggestions, and then presented those comments to policymakers. We considered such a presentation a "response" by the CAC. There were total of 697 such responses, and we exported them into an Excel spreadsheet.

The third part of our dataset was government policies. We had the original transportation policy document, the one the government used to crowdsource citizen ideas online. We also had the finalized transportation policy document, which was released 10 months after the first crowdsourcing. We copied and pasted each policy item in the original transportation document into separate rows in a spreadsheet; there were total of such 123 policy items in the original policy document. We also copied and pasted each policy item in the finalized transportation document into separate row in a spreadsheet. There were total of 202 such policy items in the finalized policy document. For each policy item, the government tracked whether it was the same as that in the original policy document or was newly added or the original policy item was deleted in the revised policy document (Figure 2).

This detailed tracking enabled us to know and examine the differences between the original and the final policy document. Comparing the two, we found that 211[6] policy items were changed (i.e., deleted/edited severely/added). We can assume that this difference was caused by citizen participation and CAC internal discussion because (1) Palo Alto news coverage on the transportation issue in the crowdsourcing months did not significantly increase or decrease on the transportation issue compared to other years;[7] (2) there were no personnel or budget changes or other political reforms going on at the same time as the crowdsourced reform, so we can exclude the possibility of bureaucratic change on government policy; (3) the


Latest Policy Document

Policy without change	~~Policy T-1.8~~Policy T-1.9	Continue to encourage the provision of amenities such as seating, lighting, and signage, including real-time arrival information, at bus and shuttle stops and train stations to increase rider comfort, safety, and convenience. [(PTC Policy T1.14) (Previous Policy T-10)
Newly-added policy	Policy T-1.20	Provide facilities that encourage and support bicycling and walking. [NEW POLICY] [T217]
Policy that was deleted	~~Program T1.2.4~~	~~establish a Transportation Demand Management program for employees. [(PTC Program T1.7.2) (Edited) (Previous Program T-6) (Edited)] [T6]~~

Figure 2. Revised transportation policy document—tracking details.

crowdsourcing months in 2016 were between January and June, when the national primary election was going on, but local transportation issues were not a debated issue during that election. Also, considering that Palo Alto is in a stable blue state, its policy agenda on transportation was unlikely to be influenced by the general election in any event.

These three factors convinced us that changes in transportation policy could only have been caused by civic participation and/or CAC internal discussion with the government. There could possibly be other underlying variables, but we see a strong potential for the causality here.

METHOD AND ANALYSES

Exploring Three Datasets Using Tf-idf

We compared the key topics in the three spreadsheets—crowd suggestions, CAC comments, and changed government policies. We first turned each spreadsheet into a txt. file in which each row represented a citizen suggestion (or a CAC comment or a policy item). We then did data cleaning on the three txt. files by removing the stop words, stemming, unifying the capitalization, and tokenizing the document. To obtain the key phrases of each dataset, we performed Tf-idf on each dataset separately. Tf-idf is a method to inform researchers how important a word or a phrase is to a document by giving researchers the weighted term frequency for each document.[8] Different from a simple calculation of how often a term appears in a document, the importance of a word "increases proportionally to the number of times a word appears in the document but is offset by the frequency of the word in the corpus."[9] The purpose of this algorithm is to understand the key words from documents.[10] Taking the crowd suggestions dataset, for example, each row in the dataset is regarded as a document. Thus, there are 270 documents in this dataset. Citizens suggestions are the corpus. Tf-idf enables us to obtain the key word(s) across these documents for a specific dataset while giving more weight to the unique word(s) appearing in each document of a dataset. We used the Scikit-Learn TfidfVectorizer package to obtain the unigram, bigram, 3-grams, and 4-grams for each dataset independently.

Figure 3 indicates that there is similarity and difference comparing the crowd's input, the CAC's comments, and the government's responses. Three-grams and 4-grams provide more meaningful information than unigrams and bigrams, showing more comprehensive sentences that form full proposals rather than singular words as in unigrams and bigrams. The Tf-idf results show that issues related to bicycling and parking are shared concerns among the crowd, CAC, and the government's responses. However, the language to express these concerns differ in the crowd's, CAC's and government's revised policies. CAC and the government use more bureaucratic language, such as "management," "evaluate," "improving circulation safety," "support transit subsidies," "work merchants designate," while the crowd expresses ideas with lay terminology such as "underground parking space" and "two-way bike lanes."

Examining Responsiveness Using Logistic Regression

Dependent Variable

To test systematically the extent to which the government responded to citizens—in particular, the relationship between participation and government responsiveness—we conducted logistic regression. Our unit of analysis was each citizen suggestion in the spreadsheet of crowd ideas ($N = 260$). Our dependent variable—"responsiveness"—was defined as a changed policy that adopted the method suggested by a citizen on the transportation issue. Two researchers coded the government responsiveness, and we used Krippendorff's alpha to check inter-coder reliability (k. alpha = 0.787). Therefore, the dependent variable was a binary outcome. Overall, 46% of the crowd suggestions were responded in the changed policies. Compared to other governments' responsiveness rate (Distelhorst and Hou 2017; Table 2), this is in the middle range.

Independent Variables

We were primarily interested in two independent variables: (1) how many times each citizen suggestion was raised during crowdsourcing; and (2) whether or not the CAC responded to a citizen suggestion by passing it along to the policymakers.

Since many crowd suggestions mentioned similar issues, we categorized each citizen suggestion in the crowd spreadsheet. The government defined seven categories for the crowd's suggestions, and all were visible in the online platform for annotating the policy document. The digital crowdsourcing platform showed us which suggestions belonged to which policy category in the transportation policy document; the city defined a total of seven. When a participant submitted their

	Crowd Suggestions	CAC Comments	Revised Policies
Top 10 unigrams	traffic, bike, transit, shuttle, city, people, need, transportation, cars, public	transportation, need, transit, city, traffic, use, new, support, add, rail	bicycle, transit, transportation, use, encourage, including, traffic, pedestrian, safety, improvement

	Crowd Suggestions	CAC Comments	Revised Policies
Top 10 Bi-grams	traffic calming, bike lanes, public transit, shuttle service, el camino, parking spaces, stop signs, speed bumps, improving transit, parking lots	suggested wording, level service, new development, research park, public transportation, rail corridor, bicycle parking, safe routes, public transit, comp plan	bicycle pedestrian, santa clara, multimodal transit, transit stations, employee parking, alma street, bicycle parking, bus service, california avenue, employment districts

	Crowd Suggestions	CAC Comments	Revised Policies
Top 10 3-grams	live underground parks, highlight new ideas, traffic calming measures, yes traffic calming, yes transportation regional, transportation regional issue, transportation obviously regional, thank inviting residents, parking garage commercial, inviting residents input	restore originally written, need improve shuttle, stanford research park, vehicle bicycle parking, motor vehicle bicycle, transportation demand management, support local transit, ptc (planning and transportation commission) motor vehicle, bicycle parking council, suggested wording change	downtown palo alto, centers employment districts, el camino real, work merchants designate, merchants designate dedicated, employee parking areas, designate dedicated employee, dedicated employee parking, multimodal transit stations, stanford research park

	Crowd Suggestions	CAC Comments	Revised Policies
Top 10 4-grams	thank inviting residents input, parking garage commercial district, agree transportation obviously regional,	regional public transportation including, vehicle bicycle parking council, ptc motor vehicle bicycle, traffic counts good measure,	provide sufficient excessive parking, merchants designate dedicated employee, evaluate design baytofoothills path,
	adding dedicated bike infrastructure, underground parking space driveway, test idea extended version, promote underground parking space, permit parking closeby residential, parking closeby residential district, number commute vehicle trips, lanes twoway separated roadways	supportive satellite parking east, support subsidized ride sharing, support policies people cars, support parking management pricing, support parking meters downtown, safe routes bikes pedestrians	designate dedicated employee parking, dedicated employee parking areas, university avenuedowntown california avenue, commercial centers employment districts, provide local transit palo, maintenance bicycle pedestrian infrastructure, local transit palo alto

Figure 3. A comparison of crowd suggestions, CAC comments, and changed policies using Tf-idf.

suggestion, the online crowdsourcing system automatically categorized the suggestion according to the city's categorization.

However, we found that those seven categories were too general and could not accurately summarize each citizen suggestion. Thus, we used NLP capabilities to guide us to construct a more detailed and accurate category list. We drew upon the procedures suggested by Short et al. (2010; Table 2). First, we followed a "deductively derived word list." To do this, we created a working definition of the construct of interest—"citizen suggestions on transportation issues"—based on prior theory. The prior theory in our case are the pre-defined seven categories from the government policy document on what constitute the transportation issue, such as public transit, private transit, parking, etc. Then, we used a Tf-idf algorithm to generate the word list (including unigram, bigram, 3-grams, and 4-grams) from the citizen suggestion spreadsheet. After the "deductive content validity" step, we conducted the "inductive content analysis." Three researchers, as well as two city staff members preparing transportation policies, collaborated to evaluate the word list derived in the first step to categorize the words into the pre-defined seven categories and identify missing words, which are important to capture the transportation issue. After refining the word list, 44 categories were developed for all of the pre-defined seven transportation categories to capture each citizen suggestion more accurately. After the 44 categories were developed, to ensure reliability, multiple researchers coded random samples of citizen suggestions and further refined the category list until the inter-coder agreement was high (k. alpha = 0.89). This inter-rater reliability check can also be done using multiple computer-assisted text analysis methods (Kenny 2015; Short et al. 2010).

There are two limitations in the categorization. First, we cannot statistically assess the external validity of the categorization since we only have citizen suggestions from one crowdsourcing process. However, we have checked the news articles on the categories we developed in other cities through searching the category name and other cities near Palo Alto and in California. We found out that those categories (i.e., issues) are also major concerns in other cities. In the future, when there are more crowdsourced transportation policies, scholars will be able to access external validity by comparing categories developed from one dataset with another dataset. Second, we are also unable to evaluate predictive validity of our developed categories (i.e., the predictive power of our categories on other variables) since there are very little data on the key dependent variables, which are correlated with citizen participation, such as the quality of services, transparency of local government, and government responsiveness. However, in the future, with more practices of open governance and, in particular, more scholarly work examining the impact of citizen participation on open governance, there will be more data available for researchers to test the predictive validity. In fact, some scholars, such as Font et al. (2016, 2017), have already begun collecting large-scale datasets to trace policy changes across municipalities in Spain. We hope that this article can provide a valuable dataset on the dependent variable "government responsiveness" for future scholars to use and thus to assess the predictive validity of their categorization.

The 44 subcategories were used to calculate our first independent variable in the logistic regression analyses. All of the results held when we used the seven categories defined by the government. Coding each citizen suggestion into a subcategory enabled us to identify the frequency of the subcategory appearances in the dataset of crowd suggestions. If a citizen suggestion belonged to subcategory A, the number of times this suggestion appeared equaled the number of times subcategory A appeared in the crowd spreadsheet. Therefore, we developed our first independent variable: the number of times each citizen suggestion was mentioned.

The second independent variable is whether or not the CAC responded to a citizen suggestion. It is a categorical variable. The purpose of this independent variable was to examine how CAC's response to crowd suggestions influences government responsiveness—whether the government was more likely to respond to a citizen suggestion if CAC had first responded to it. Two research assistants coded CAC responsiveness, and we used Krippendorff's alpha to check inter-coder reliability (k. alpha = 0.785). To measure the effect of this variable on the outcome, we first conducted logistic regression to measure its effect size. Moreover, we also performed a moderator test (Kenny 2015) to measure how the CAC variable alters the strength of the relationship between the crowd variable and government responsiveness. We regard the variable of CAC as a moderator variable rather than a mediator variable because this variable (whether CAC responded to citizen suggestions or not) can only affect the strength of the relationship between the crowd and the government, rather than explaining the relationship between citizen suggestions and government responsiveness.

Control Variables

As discussed in the theory section, government responsiveness can also be influenced by how the government perceives the quality of citizen suggestions and media attention to the issues discussed during crowdsourcing. After interviewing key policymakers on what they perceived as high-quality citizen suggestions, two dimensions stood out: (1) how easily a suggestion could be implemented; and (2) how broad or narrow a suggestion was. City planners and two city staff members in charge of the crowdsourcing reform rated each citizen suggestion along these two dimensions on a scale of one to seven. To measure media attention to the transportation issue, we collected all of the news articles from the Palo Alto Online News for the crowdsourcing period ($N = 538$). To identify whether an article was about transportation, we used the Tf-idf method mentioned earlier to develop all of the key words related to the transportation issue. If an article title or an article abstract contained a key word, it was automatically coded as a transportation article. This resulted in 114 articles during the crowdsourcing period. We coded each transportation article along the 44 subcategories we had developed. Since we knew which subcategory each citizen suggestion belonged to and how many times that subcategory was reported by the media, we calculated an issue-attention value for each citizen suggestion.

Logistic Regression Results

Table 1 shows the coefficient estimates and standard errors of logistic regression with four specifications. In column (1), we estimated the effect on government responsiveness of the number of times each citizen suggestion was raised. There was no significant relationship. The government therefore did not respond based on the frequency of the crowd's proposal.

In column (2), we added in the squared term of the number of times each suggestion was raised to examine any nonlinear relationship between participation (i.e., frequency of suggestions) and government responsiveness. The coefficient of this term was negative, indicating an inverse U-shape between participation and government responsiveness, although the coefficient is not significant at the 0.05 level.

In column (3), we added the other independent variable: whether or not CAC had "responded" to a citizen suggestion by pushing certain proposals along to the government. The results showed a very interesting relationship. There was a positive and significant relationship between the responsiveness of CAC and that of the government. In other words, if CAC responded to a crowd suggestion, the government was likewise much more likely to respond. To further prove whether CAC is a powerful moderator, we used Cohen's d—the mean difference divided by pooled standard deviation (Kenny 2015). We found that the effect of the number of times each citizen suggestion was raised on government responsiveness when CAC responded to a citizen suggestion is 1.07 larger than when CAC did not respond to a citizen's suggestion, indicating a large effect (Cohen 1988).[11] Thus, CAC did serve as a powerful moderator role between the crowd and the government. CAC amplified the online crowd's voice, though selectively.

In column (4), we added the other covariates. The results of our key independent variables remain robust. Overall, we did not find any significant relationship between other covariates and government responsiveness at the 0.05 level. We did find that the coefficient of ease of implementation is close to significant at the 0.05 level ($p = 0.056$), indicating that if the government perceived a suggestion as more difficult to implement, then it was less likely to respond to it in policy changes.

To calculate the effect size of CAC responsiveness, we compared the probability of government responsiveness to a suggestion that CAC had responded to with one that it had not (0.5142), holding all other covariates at the mean. Thus, the government was 51 percentage points more likely to respond to a citizen suggestion to which CAC had first responded.

These results indicate that government responsiveness is not so much determined by the frequency of a crowd suggestion but by whether CAC had responded to it. This suggests that whether or not crowdsourcing influences government policy is highly dependent on whether the CAC valued the suggestion. Therefore, CAC plays the role of the amplifier and moderator.

The fact that the frequency of crowd suggestions did not determine government responsiveness is not surprising. Crowdsourcing is based on self-selection. Any

TABLE 1
Government Responsiveness and Determining Factors—Logistic Regression

	Government Responsiveness			
	(1)	(2)	(3)	(4)
Predictors of Government Responsiveness				
Citizen Suggestion Count	0.02 (0.010)	0.078 (0.010)	0.069 (0.060)	0.082 (0.061)
Citizen Suggestion Count Squared		−0.002 (0.001)	−0.001 (0.002)	−0.002 (0.002)
CAC Responsiveness			2.435*** (0.330)	2.442*** (0.337)
Issue Attention				0.005 (0.019)
Ease of Implementation of Suggestion				−0.203* (0.107)
Broadness of a Suggestion				0.017 (0.090)
Intercept	−0.197 (0.193)	−0.581 (0.452)	−2.264*** (0.452)	−1.400* (0.635)
McFadden's R-Squared	9.83E-05	6.50E-03	0.202	0.214
Hosmer-Lemeshow *p*-value	0.608	0.309	0.630	0.657
Observations	260		260	260

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Notes: The column numbers (1), (2), (3), and (4) refer to different model specifications. To evaluate model fitness for logistic regression, we used several methods. The McFadden's R-squared is similar to the adjusted R-squared in linear regression. The Hosmer-Lemeshow test uses *p*-value to evaluate the model fitness and $p < 0.05$ indicates a poor fit.

government response determined by the frequency of suggestions might have stemmed from input from a small set of self-selected people that is not representative of the general public. However, that does not justify the potential ignorance of the crowd's suggestions. Applying crowdsourcing correctly requires the government to evaluate citizens' suggestions qualitatively and to implement them based on thoughtful consideration rather than on frequency. The result of a non-significant relationship between the frequency of the civic input and the city's responsiveness suggests the possibility that the government was correcting the self-selection bias of the crowd. Nevertheless, there could be other reasons. To uncover why the government did not respond based on the frequency of the crowd's suggestions, we conducted in-depth interviews with the city staff members and CAC members.

Explaining Responsiveness with Interviews and Participant Observation

To examine the underlying mechanism that could explain responsiveness, we attended the CAC's monthly meetings and conducted in-depth interviews with 12 CAC members and leading city staff in the policy reform. The average monthly meeting lasted two hours, and the length of interviews was 45 minutes. During the meetings, we paid attention to how the government responded to CAC deliberation. During the interviews, we focused on two questions: (1) What is the CAC members' and city staff members' perspective on the role of the crowd's participation in the policy reform? and (2) How did CAC and the city process the crowdsourced input?

A key finding from the participant observation was that either the city staff took notes on what CAC discussed during the meeting or a government consultant took notes during the meeting. If a CAC member thought the notes were inaccurate, they would correct them. After every meeting, the city staff produced a draft of the latest CCP using the notes and transcripts. Thus, the CAC ideas were reviewed at the meeting and afterwards, and they were considered when creating the policy. This shows that the government paid great attention to what the CAC discussed. As a CAC member told us during a meeting, "The government is really doing a very careful job in listening to us and is doing a lot of preparation." This can explain why, in the logistic regression, the government was more likely to respond to a suggestion to which the CAC had responded.

The interviews also revealed an important paradox, which creates a challenge facing participatory democracy in the digital age. On the one hand, both the CAC members and the city staff perceived the online participation as a benefit to policymaking:

> That part [civic engagement] is great. Having some ability for people to come in and talk about the things they care about, and also see what other people think, I think it's all good. (CAC member, male)

> It reflects the community's values. We want to keep them continually engaged. (City staff member, female)

On the other hand, both CAC members and city staff noted their frustration in dealing with the large number of citizen suggestions they had to read and consider—what we termed "civic data overload." There were simply too many suggestions to process. In addition to the many other policy questions involved in crowdsourcing, the transportation issues alone generated hundreds of suggestions. City staff and CAC members described the problem as follows:

> To be honest, it is challenging. Just the volume of it, there is just this incredible volume. (City staff member, female)

> In addition to reading your stack of 150 pages of CAC draft, and these other 50 pages of notes other people have written, also go and check the online commentary and try to crystallize it yourself. I have a very hard time doing all of that. Seriously with the job and three kids, I don't have time to go through and do all of those things. (CAC member, male)

The interview results showed how the government's limitations in efficiently and meaningfully processing the crowdsourced data constrained its ability to respond to the crowd's proposals in policy reform. Therefore, although the city happily engaged citizens in policymaking and tried to read as many suggestions as possible, the lack of efficient analysis methods effectively rendered, to a degree, the participants' voices unheard and underprocessed. The civic data overload created an unfortunate paradox: the more actively people participated online, the more successful the crowdsourcing effort was in terms of knowledge discovery and civic engagement; simultaneously, however, processing the data became more challenging, and the crowd's total as well as average input was less likely to be heard and utilized.

DISCUSSION

The Data Bottleneck and the Hierarchy of Participatory Democratic Innovation

In this article, we provided empirical evidence to answer the unveiled issue of how to gauge the extent of government responsiveness to crowd's suggestions and how to identify the factors that determine such responsiveness. Despite policymakers' strong intent to encourage civic engagement and utilize crowd input to help draft and revise policy, we found that participatory digital innovation generated too many citizen suggestions for the government to synthesize, analyze, and evaluate. This civic data overload created a bottleneck that inhibited the government from fully processing and utilizing citizen suggestions. The data overload simply exceeded the government's existing skills and resources for dealing with it effectively. Such incompatibility in the organization is one of the factors hindering innovation adoption, as argued by Rogers (1962). For the government successfully to implement crowdsourcing as a democratic innovation, crowdsourcing should fit with the current governance processes. The civic data overload made crowdsourcing incompatible with the established policymaking practices and the traditional policymaking cycle: there was too much diverse input for the government to process efficiently. The features of crowdsourcing may not change in that the method will inevitably produce a large number of diverse citizen inputs, which will most likely cause challenges, even when efficient and reliable automated analysis methods are used, because human involvement is still needed to a certain degree. Moreover, the government's capacity to process crowdsourced input is not likely to improve drastically without extensive training.[12] If the government wants to continue using crowdsourcing in policymaking, it has to adjust its policymaking practices and to actively seek data processing mechanisms to make crowdsourcing a compatible method. Otherwise, the friction between crowdsourcing and the established policymaking practices will remain, hindering the adoption of crowdsourcing in the government.

We also found that, when facing the civic data overload, the government chose a hierarchal filter system to help it process the civic data. This filtering system had two components: (1) a self-selected citizen representative body, the CAC; and (2) high-quality citizen suggestions that the government perceived as relatively easily actionable. This hierarchal filter system echoes with Font et al.'s (2016, 2017) finding that governments cherry-pick projects from certain participatory mechanisms they prefer. The government tended to listen more carefully to what CAC (i.e., the self-selected citizen representative body) discussed than what the crowd suggested (i.e., the crowdsourcing body). The process of the filtering of crowd information is thorough: at its monthly meetings with the key government staff, the CAC discussed what it had heard from citizens and what it regarded as important changes to current policies; the government listened carefully to the CAC and tried to read some crowd suggestions; then, the government incorporated those actionable suggestions into policy changes.

Therefore, we see (Figure 4) that crowdsourced policymaking introduced a hierarchy of participation: the crowd was on the bottom; CAC filtered the crowd suggestions; and the government responded; i.e., decided which to implement. In other

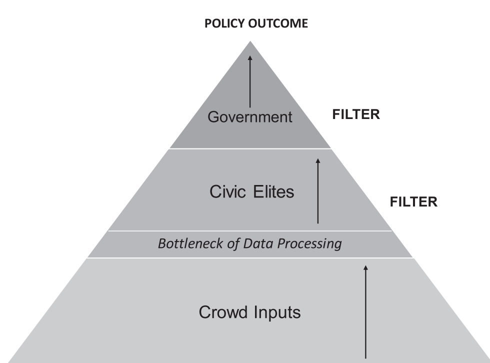


Figure 4. Hierarchy participation in crowdsourced policymaking.

words, without CAC's approval, a suggestion was less likely to end up being included in the policy.

Thus, the civic data overload poses a new challenge to a government wishing to implement crowdsourcing—how to synthesize civic data more efficiently and thereby value citizen voices more appropriately. As a response to this challenge, the government currently chose to use the filter system to process data, going against the original intention of crowdsourcing to create equality and inclusiveness in policymaking.

Problems of the Hierarchal Filter System

There are two main concerns with the hierarchical filter system in crowdsourced policymaking. These concerns figure into participatory democracy practices and digital democratic innovations in general.

The first is the civic data overload. Members of the CAC, as well as the city staff, expressed difficulty in analyzing and understanding the large number of crowd suggestions. Since CAC' s advisory job was voluntary and part-time, the members lacked the time and capacity to read each citizen suggestion carefully. As a result, they only read some of them, and they could only pass along to the government the ones they *had* read. The city staff also read only a few citizen suggestions and made policy changes based on their perception of these suggestions. Without understanding the crowd suggestions completely, the filter can distort what the crowd really wants.

The second concern is the role of the CAC as a filter of the crowd's input. To establish the CAC, the city launched an open call for residents to volunteer. The city then selected CAC members from among respondents. This process produced a CAC that was heavily populated with a "civic elite"—urban planners, architects, NGO activists, interest-group members, and former government agency officials—not residents from more diverse backgrounds and experiences. As previously noted (Table 1), using Tf-idf, we found out that the topics the CAC raised were more similar to the government's concerns than to the crowd's. The CAC discussed measurements and pilot studies on transportation rather than aspects of transportation closely related to daily life. Although the CAC's monthly meetings and materials were open to the public, they were not as easy for citizens to access as the crowdsourcing platforms. Therefore, the composition and process of the CAC were not as inclusive and transparent as they could and, perhaps, should have been. It is unclear whether the CAC accurately communicated the broader public's voices to the government, rather than representing its members' own interests or the interests of groups they represented. It is worrisome if they did not, because the government responded more often to citizen suggestions when CAC had first responded to them. Therefore, using citizen elites as a filter creates inequality in crowdsourced policymaking. The goal of enhanced inclusiveness, transparency, and collaboration thus falls short in part. On the other hand, simply by including the crowd and the CAC in policy development, the process undoubtedly is more inclusive, transparent, and collaborative than one without any civic participation at all.

Governments across the United States and around the world increasingly turn to crowdsourcing as a participatory method in policymaking, using online crowdsourcing platforms and CACs to engage large-scale citizen participation.[13] Open-government practices are supposed to bring equality and inclusiveness. The civic data overload challenge, however, creates an unequal structure for processing civic input in digital democratic innovations. In these crowdsourcing experiments, as well as in other digital democratic innovations, engaging citizens and effectively processing the data are common problems (Charalabidis et al. 2012; Janssen et al. 2012).

CONCLUSION

Crowdsourced policymaking is an increasingly common participatory method in the age of digital governance. Like many other open government practices and digital democratic innovations, it is supposed to generate useful civic data for the policymakers, allow more citizen voices to be heard, and develop a more inclusive, participatory, and collaborative government. However, this article unveils a hierarchical filter structure for processing crowdsourced input that results in the government being 1.5 times more likely to respond to crowd suggestions that were first responded to by the CAC, a citizen committee established by the government.

The crowdsourced transportation policy reform in Palo Alto offers several crucial lessons for effective participatory democratic innovation in the digital age. The first and foremost is how to design a better system to listen to, process, and respond to citizen suggestions. As Mergel et al. (2016) noted, “Big data accumulates quickly and seemingly exponentially; it can quickly overwhelm an analyst.” The policymakers themselves do not have the time and capacity to read each citizen suggestion and synthesize the civic data. Nor should full reliance be placed upon a CAC filter, which is often not a representative public voice and can lack transparency in its procedures.

This article introduced one example of a computational text analysis method—the Tf-idf algorithm of natural language processing—as a tool to efficiently understand the crowd input. Other methods in computational text analysis—such as automatic categorization and sentiment analysis—could also help both public managers and citizen committees similar to the CAC understand the main concerns of the crowd input. Hiring data scientists or outsourcing civic data analyses to data companies would more comprehensively reveal crowd suggestions than sole reliance on a CAC. Automated data analytic tools are another, more scientific and efficient, filter for synthesizing crowd inputs. In short, public managers need the capacity to “use a mix of staff, contractors, and personal resources to manage, analyze and interpret large-scale datasets” (Mergel et al. 2016:934).

Moreover, the government needs to educate citizens in how to develop constructive suggestions in a crowdsourced policymaking process. For instance, it should provide feedback to participating citizens, acknowledging their contribution, informing them how the government processes their input, and clearly communicating what kind of suggestions are helpful. This feedback loop would also be essential to motivate continuing civic participation. The online forum should serve not only to gather information or promote deliberation between citizens. It should also be an enhanced, two-way communication platform between the crowdsourcer—the government—and the crowd.

With the increasing practices of digital innovation in democratic participation, our article, by examining the extent to which policymakers respond to crowd participation and the factors determining that responsiveness, reveals a previously unknown side to these innovations: the civic data overload and the consequent filter bottleneck. Without efficiently synthesizing crowd suggestions in digital democratic innovations, citizen voices will remain unheard and undervalued.

NOTES

1. For the purpose of this article, we do not analyze the CAC's responsiveness to the crowd's ideas. We focus on the relationship between the government and the crowd, because that is the relationship between the crowdsourcer and the crowd.
2. http://www.paloaltocompplan.org/
3. https://www.peakdemocracy.com/portals/5/Issue_2871
4. https://www.peakdemocracy.com/portals/5/Issue_3378
5. http://www.paloaltocompplan.org/wp-content/uploads/2016/01/CAC-Comment-Matrix-Transportation_20160125.pdf
6. Please note that some policies from the original policy document were deleted, which means that they did not appear in the final document. This is why the number of policies that were changed can be larger than the number of policies in the final document ($211 > 202$).
7. We web-scraped all of the Palo Alto Online news articles (https://www.paloaltoonline.com/square/) from the time it was founded (year 2006) to the time this data was accessed in May, 2017 ($N = 6{,}366$). To identify whether or not a news article was about transportation, we used Tf-idf to develop all of the words and phrases that are related to transportation issues. Since the crowdsourcing period is from August 2015 to June 2016, we compared the percentage of transportation news coverage from August–December for the years 2010–2016 and we did not find that August–December 2015 had a significantly higher percentage of news coverage on transportation. Similarly, we also compared the percentage of transportation articles from January–June for the years 2010–2017. We also did not find that January–June 2016 had a significantly higher percentage of news coverage on transportation. The total number of news article that are about transportation during the crowdsourcing months is 103.

8. For how scholars in the field of public management use natural language processing techniques, see Pandey, Pandey and Miller (2017). For a more general use of machine learning techniques on text collection and analyses, see Manning, Raghavan, and Schütze (2008).
9. For more details on the meaning and computing of Tf-idf algorithm, see: http://www.tfidf.com.
10. Some readers might wonder how the approach of Tf-idf is similar or different from LDA, which is a popular topic modeling method in text mining. The purposes of LDA and Tf-idf are different. The goal of LDA is to help researchers generate the main topics and their associated words so that researchers can label the topic to classify a document. The purpose of Tf-idf is to give researchers weighted term frequency for each document and to inform researchers how important a word or a phrase is to a document. When we compare the three datasets—citizen suggestions, representative suggestions, and the policy changes—we are more interested to learn about the key words (phrases) for the three datasets and therefore we chose Tf-idf. When we tried to develop the category to classify citizen suggestions, we tried both LDA and Tf-idf. We found that the topics generated by LDA did not help us label the document in a meaningful way. Using the words generated by Tf-idf was better able to help the researchers to develop the category. Therefore, although both methods are helpful text analysis tools, for our data, using the word list Tf-idf generates is more useful than the topics (and associated words of each topic) from LDA.
11. As David Kenny (2015) explained, "when X and M are dichotomies f^2 equals the $d^2/4$ where d is the d difference measure described above. Cohen (1988) has suggested that f^2 effect sizes of 0.02, 0.15, and 0.35 are termed *small*, *medium*, and *large*, respectively." For details, see: http://davidakenny.net/cm/moderation.htm. In our case, $d = 1.07$, thus $f^2 = 0.29$, which is closer to 0.35. Thus, it is a large effect.
12. http://www.governing.com/commentary/col-training-policymakers-data-science-analysis.html
13. In Midland Michigan, http://www.cityofmidlandmi.gov/125/E-CityHall?pd_url=https%3A%2F%2Fwww.peakdemocracy.com%2Fportals%2F242%2FIssue_3878#peak_democracy, the government has adopted a similar online crowdsourcing platform as the city of Palo Alto to crowdsource citizen ideas. In Los Altos (http://www.losaltosca.gov/community/page/open-city-hall), Klamath Falls, Oregon (https://www.klamathfalls.city/i-want-to/find/city-hall/administration/open-city-hall,) Decatur, Georgia (http://www.decaturga.com/whats-new/open-city-hall), etc., the government has used Open City Hall to call for crowd suggestions on policy issues and government budget. Beyond the United States, crowdsourcing has also been used in Finland (www.thefinnishexperiment.com) and Iceland (https://democracyoneday.com/tag/iceland/). As scholars (Aitamurto 2015;

Aitamurto and Chen 2017) pointed out, governments across the world are facing an increasing civic data analysis challenge.

REFERENCES

Aitamurto, T. 2012. *Crowdsourcing for Democracy: New Era in Policy-Making*. Helsinki, Finland: Publications of the Committee for the Future, Parliament of Finland.

Aitamurto, T. 2016. "Collective Intelligence in Law Reforms: When the Logic of the Crowds and the Logic of Policymaking Collide." Pp. 1780–2789. Proceedings of the 49th Hawaii International Conference on System Sciences, Kauai, January 5–8.

Aitamurto, T., and K. Chen. 2017. "The Value of Crowdsourcing in Public Policymaking: Epistemic, Democratic and Economic Value." *The Theory and Practice of Legislation* 5(1):55–72.

Aitamurto, T., and J. Saldivar. 2017. "Motivating Participation in Crowdsourced Policymaking: The Interplay of Epistemic and Interactive Aspects. Motivating Participation in Crowdsourced Policymaking: The Interplay of Epistemic and Interactive Aspects." *Proceedings of the Acm on Human-Computer Interaction* 1(2):1–22. Article 18.

Aitamurto, T., and H. Landemore. 2015. "Five Design Principles for Crowdsourced Policymaking: Assessing the Case of Crowdsourced off-Road Traffic Law in Finland." *Journal of Social Media for Organizations* 2(1):1–19.

Aitamurto, T., and H. Landemore. 2016. "Crowdsourced Deliberation: The Case of the Law on off-Road Traffic in Finland." *Policy & Internet* 8(2):174–96.

Aitamurto, T., H. Landemore, and J. Saldivar Galli. 2017. "Unmasking the Crowd: Participants' Motivation Factors, Expectations, and Profile in a Crowdsourced Law Reform." *Information, Communication & Society* 20(8):1239. http://dx.doi.org/10.1080/1369118X.2016.1228993

Bekkers, V. J., and V. Homburg, eds. 2005. *The Information Ecology of e-Government: e-Government as Institutional and Technological Innovation in Public Administration*. Vol. 9. Amsterdam, The Netherlands: IOS Press.

Besley, T., and M. Ghatak. 2003. "Incentives, Choice, and Accountability in the Provision of Public Services." *Oxford Review of Economic Policy* 19(2):235–49.

Besley, T., and R. Burgess. 2001. "Political Agency, Government Responsiveness and the Role of the Media." *European Economic Review* 45 (4–6):629–40.

Brabham, D. 2013. *Crowdsourcing*. Cambridge, MA: MIT Press.

Charalabidis, Y., A. Triantafillou, V. Karkaletsis, and E. Louki. 2012. "Public Policy Formulation Through Non-Moderated Crowdsourcing in Social Media." Pp. 156–69. Paper presented at International Conference on Electronic Participation.

Chun, S. A., S. Shulman, R. Sandoval, and E. Hovy. 2010. "Government 2.0: Making Connections between Citizens, Data and Government." *Information Polity* 15(1):1.

Couldry, N. 2010. *Why Voice Matters*. London, UK: Sage Publications Inc.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Erlbaum.

Distelhorst, G., and Y. Hou. 2017. "Constituency Service under Nondemocratic Rule: evidence from China." *The Journal of Politics* 79(3):1024–1040.

Edwards, M. 2001. *Social Policy, Public Policy: From Problem to Practice*. Sydney, Australia: Allen and Unwin.

Estellés-Arolas, E., and F. González-Ladrón-de-Guevara. 2012. "Towards an Integrated Crowdsourcing Definition." *Journal of Information Science* 38(2):189–200.
Fishkin, J. S. 2011. *When the People Speak: Deliberative Democracy and Public Consultation*. Oxford, UK: Oxford University Press.
Font, J., S. P. del Amo, and G. Smith. 2016. "Tracing the Impact of Proposals from Participatory Processes: methodological Challenges and Substantive Lessons." *Journal of Public Deliberation* 12(1):3.
Font, J., G. Smith, C. Galais, and P. Alarcón. 2017. "Cherry-Picking Participation: Explaining the Fate of Proposals from Participatory Processes." *European Journal of Political Research* 57:615–636.
Frey, B. S. 1994. "Direct Democracy: Politico-Economic Lessons from Swiss Experience." *American Economics Review* 84(2):338–342.
Fung, A. 2006. "Varieties of Participation in Complex Governance." *Public Administration Review* 66(s1):66–75.
Fung, A. 2015. "Putting the Public Back into Governance: The Challenges of Citizen Participation and Its Future." *Public Administration Review* 75(4):513–522.
Fung, A., and E. O. Wright. 2001. "Deepening Democracy: Innovations in Empowered Participatory Governance." *Politics and Society* 29(1):5–42.
Gaventa, J. 2004. "Towards Participatory Governance: Assessing the Transformative Possibilities." Pp. 25–41 in *Participation: From Tyranny to Transformation*, edited by S. Hickey and G. Mohan, New York: Palgrave Macmillan.
Goel, A., A. Krishnaswamy, S. Sakshuwong, and T. Aitamurto. 2015. "Knapsack Voting." *Collective Intelligence 2015*, Santa Clara, May 2015.
Howe, J. 2008. *Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business*. New York: Crown Business.
Howlett, M., M. Ramesh, and A. Perl. 1995. *Studying Public Policy: Policy Cycles & Policy Subsystems*. Toronto: Oxford University Press.
Janssen, M., Y. Charalabidis, and A. Zuiderwijk. 2012. "Benefits, Adoption Barriers and Myths of Open Data and Open Government." *Information Systems Management* 29(4):258–268.
Janssen, M., and H. Van Der Voort. 2016. "Adaptive Governance: Towards a Stable, Accountable and Responsive Government." *Government Information Quarterly* 33(1):1.
Jeppesen, L. B., and K. R. Lakhani. 2010. "Marginality and Problem-Solving Effectiveness in Broadcast Search." *Organization Science* 21(5):1016–1033.
Kenny, D. 2015. "Moderator Variables." http://davidakenny.net/cm/moderation.htm
Konisky, D. M., and T. C. Beierle. 2001. "Innovations in Public Participation and Environmental Decision Making: Examples from the Great Lakes Region." *Society and Natural Resources* 14(9):815–26.
Kuklinski, J. H., and J. E. Stanga. 1979. "Political Participation and Government Responsiveness: The Behavior of California Superior Courts." *American Political Science Review* 73(4):1090–1099.
Landemore, H. 2015. "Inclusive Constitution-Making: The Icelandic Experiment." *Journal of Political Philosophy* 23(2):166–191.
Landemore, H. 2017. "Inclusive Constitution Making and Religious Rights: Lessons from the Icelandic Experiment." *The Journal of Politics* 79(3):762–779.
Lee, D., A. Goel, T. Aitamurto, and H. Landemore. 2014. "Crowdsourcing for Participatory Democracies: Efficient Elicitation of Social Choice Functions." Paper

presented at The Association for the Advancement of Artificial Intelligence conference on Human Computation and Crowdsourcing (HCOMP 2014), Pittsburgh, Pennsylvania.

Macintosh, A. 2004. "Using Information and Communication Technologies to Enhance Citizen Engagement in the Policy Process." *Promise and Problems of E-Democracy* 19:142.

Macnamara, J. 2016. *Organizational Listening: The Missing Essential in Public Communication*. New York: Peter Lang Publishing, Inc.

Mergel, I., R. K. Rethemeyer, and K. Isett. 2016. "Big Data in Public Affairs." *Public Administration Review* 76(6):928–937.

Michels, A., and L. De Graaf. 2010. "Examining Citizen Participation: Local Participatory Policy Making and Democracy." *Local Government Studies* 36(4):477–491.

Nelimarkka, M., B. Nonnecke, S. Krishnan, T. Aitamurto, D. Catterson, C. Crittenden, … J. Patel. 2014. "Comparing Three Online Civic Engagement Platforms Using the "Spectrum of Public Participation framework."" Pp. 25–6. Proceedings of the Oxford Internet, Policy, and Politics Conference (IPP).

Noveck, B. S. 2009. *Wiki Government: how Technology Can Make Government Better, Democracy Stronger, and Citizens More Powerful*. Washington, DC: Brookings Institution Press.

Noveck, B. S. 2015. *Smart Citizens, Smarter State: The Technologies of Expertise and the Future of Governing*. Harvard, UK: Harvard University Press.

Pandey, S., S. K. Pandey, and L. Miller. 2017. "Measuring Innovativeness of Public Organizations: Using Natural Language Processing Techniques in Computer-Aided Textual Analysis." *International Public Management Journal* 20(1):78–30.

Pateman, C. 1970. *Participation and Democratic Theory*. Cambridge, UK: Cambridge University Press.

Pateman, C. 2012. "APSA Presidential Address: Participatory Democracy Revisited." *Perspectives on Politics* 10(1):7–19.

Peters, G. 1999. *American Public Policy: Promise and Performance*. Chappaqua, NY: Chatham House/Seven Rivers.

Pirog, M. A. 2014. "Data Will Drive Innovation in Public Policy and Management Research in the Next Decade." *Journal of Policy Analysis and Management* 33(2):537–543.

Qin, B., D. Strömberg, and Y. Wu. 2017. "Why Does China Allow Freer Social Media? Protests versus Surveillance and Propaganda." *The Journal of Economic Perspectives* 31(1):117–140.

Rogers, E. M. 1962. *Diffusion of Innovations*. New York: Free Press of Glencoe.

Rosanvallon, P., and A. Goldhammer. 2008. *Counter-Democracy: Politics in an Age of Distrust*. Vol. 7. Cambridge, UK: Cambridge University Press.

Short, J. C., J. C. Broberg, C. C. Cogliser, and K. C. Brigham. 2010. "Construct Validation Using Computer-Aided Text Analysis (CATA): an Illustration Using Entrepreneurial Orientation." *Organizational Research Methods* 13(2):320–347.

Smith, G. 2009. *Democratic Innovations: Designing Institutions for Citizen Participation*. Cambridge, UK: Cambridge University Press.

Surowiecki, J. 2005. *The Wisdom of Crowds*. New York: Anchor.

Thomas, J. C., and G. Streib. 2003. "The New Face of Government: citizen-Initiated Contacts in the Era of E-Government." *Journal of Public Administration Research and Theory* 13(1):83–102.

Yang, K., and K. Callahan. 2005. "Assessing Citizen Involvement Efforts by Local Governments." *Public Performance & Management Review* 29(2):191–216.

Yanovitzky, I. 2002. “Effects of News Coverage on Policy Attention and Actions: A Closer Look into the Media-Policy Connection.” *Communication Research* 29(4):422–451.

ABOUT THE AUTHORS

Kaiping Chen (kpchen23@stanford.edu) is a Ph.D. candidate at the Department of Communication, Stanford University. Her research uses big data, machine learning, and text analysis to study civic engagement and government responsiveness across regimes. Her research has appeared or is forthcoming in journals such as *American Political Science Review*, *International Public Management Journal*, *Theory and Practice of Legislation*, and the *Russell Sage Foundation Journal of the Social Sciences*.

Tanja Aitamurto, Ph.D. (tanjaa@stanford.edu), is a postdoctoral fellow at the School of Engineering at Stanford. She examines the impact of civic technologies on human behavior and society. Tanja’s work has received a number of awards and has been published in highly ranked academic journals, including *New Media & Society*, *Information, Communication & Society*, and *International Journal of Communication*, and in venues such as the Conference on Human Factors in Computing Systems (CHI) and Computer-Supported Collaborative Work and Social Computing (CSCW). Tanja has attended meetings and given talks about her research at the White House, the United Nations General Assembly, the Wikimedia Foundation, OECD, the Council of Europe, and in several parliaments and governments.