The Medium Is The Message:
The Impact of Technological Innovation in
Municipal Coproduction Systems

Presented at the 2015 APPAM Conference
Miami, Florida
November 14, 2015

Matthew Young
Ph.D. Candidate
Sol Price School of Public Policy
University of Southern California

ABSTRACT
This study contributes to both coproduction and e-government literatures by examining how the integration of new technologies into an existing 311 complaint system improves its performance. Coproduction has recently received increased attention in the public management and administration field. While prior research on coproduction in general and 311-based coproduction in particular has explored questions of distributional biases, few studies empirically assess how adoption of emergent technologies affects coproduction outputs. This paper utilizes survival analysis to assess how the introduction of a smartphone app affects the length of time necessary to address reported issues. Preliminary results are sensitive to assumptions regarding hazard proportionality and baseline functional form, indicating the need for additional diagnostics. Differences in demographics and types of reports generated across technologies are also explored, with findings suggesting that users of the smartphone and Twitter systems may be more attentive to issues related to pedestrian traffic throughout the city.

DRAFT: Please do not distribute or cite without the author’s permission
INTRODUCTION

Over the past several decades, many local governments around the world have been faced with the unenviable task of maintaining or improving levels of service with fewer and fewer resources. How governments can maximize efficiency while still addressing concerns over equity is the subject of considerable scholarship. Some deal with institutional arrangements, while others are concerned with different technological tools and resources that can both augment existing arrangement and lead to new ones altogether. This paper focuses on the interaction between these two bodies of research. Specifically, it asks how the implementation of emergent information and communication technologies (ICTs) affect the output of coproduction systems, defined by Ferris (1984) as a system of voluntary citizen participation in the production of publically provided goods and services.

Coproduction has recently received increased attention in the public management and administration field. While earlier work focused on normative considerations and developing theories of coproduction, the second wave of scholarship is more empirical, with studies examining the factors that drive the decision to use coproduction and the levels of equity in its usage (Bovaird, 2007; Clark, Brudney, & Jang, 2013; Jakobsen & Andersen, 2013; Minkoff, 2013). This last area of interest includes work on how emergent technologies may facilitate or hinder coproduction, and dovetails with the literature on e-government. Indeed, e-government and coproduction share many normative justifications: both are theorized as ways for governments (especially municipal governments) to “do more with less” and maintain service levels in spite of dwindling budgets. Both coproduction and e-government have also been credited with offering ways to radically alter the relationship between government and citizens, although these claims are not without their skeptics (Bovaird, 2007; Brudney, 1985; Coursey & Norris, 2008; Dunleavy, Margetts, Bastow, & Tinkler, 2006; Linders, 2012; Norris & Reddick, 2013).
Understanding how emergent technologies affect the coproduction of public goods in particular is important for several reasons. First, given the normative argument for efficiency that motivates both coproduction and e-government adoption it is crucial to understand if the pairing of these ideas yields the sort of efficiency gains predicted by theory. Second, some prior research suggests that the implementation of emergent technologies in coproduction systems may reduce distributional biases in usage, challenging common understanding of the regressive effects of information technology. This may be sufficient cause to motivate public managers to adopt and implement these technologies, although in fiscally constrained governments, cost effectiveness may be necessary component for any argument for adopting new technologies. Yet for all of the way in which coproduction and e-government are complementary, little is known about whether adoption of e-government-style technologies translates to material changes in the output of coproduction arrangements. This study contributes to both coproduction and e-government literatures by examining how the integration of new technologies into an existing municipal coproduction system affects the length of time necessary to fix reported issues via survival analysis.

The rest of the paper is organized as follows. A discussion of the preexisting literature is presented first, culminating in proposed hypotheses. This is followed by a discussion of the data and methods used to test the hypotheses. Findings from the analysis are then presented, as well as a discussion of their implications. The paper then concludes with a summary of the results as well as an evaluation of its limitations and directions for further research.
BACKGROUND

This section provides an overview of both the coproduction and e-government literatures as they relate to public management and administration. It begins with coproduction in both general terms and in specific relation to the study of municipal 311 systems. It then transitions to a discussion of the e-government literature before generating the hypotheses that will be tested in the empirical analysis to follow.

Coproduction and 311 Systems

Research on coproduction has enjoyed a revival over the past several years. Originally popularized in the 1970s and 1980s by the Indiana School, the initial wave of scholarship considered the normative criteria for and implications of coproduction in the face of shrinking fiscal capacity across all levels of government in the United States (Ferris, 1984; Kiser, 1984; Parks et al., 1981; Percy, 1984). Initial definitions of coproduction itself varied, as scholars grappled with the type of output and production functions that would be included or excluded (Ferris, 1984; Kiser, 1984; Rich, 1981). This paper uses definition of coproduction advanced by Ferris and used much of the subsequent literature, where coproduction is delineated as voluntary citizen participation in the production of publically provided goods and services (Ferris, 1984). More recent coproduction research focuses on empirical analysis of the motivations behind and consequences of coproduction (Bovaird, 2007; Chen, Tsou, & Ching, 2011; Clark & Guzman, 2014; Isett & Miranda, 2015; Jakobsen & Andersen, 2013). In transitioning from normative to empirical agendas, new-wave research employs data from specific coproduction services and systems to test theories of coproduction.

One such service type is the coproduction of public goods through 311 systems. The 311 system is a non-emergency contact system where citizens can report service disruptions, vandalism,
public property damage, or other issues. Originally conceived in the mid 1990s as a way to eliminate the misuse of the 911 emergency system in Baltimore, 311 systems have since diffused to municipalities across the United States. Individuals reporting problems are coproducers in the sense that the task of identifying problems in need of attention – information generation about the state of the world – is divested from government and provided instead by citizens-as-volunteers. Once a report is entered into the system it is assigned to the appropriate agency or department. Prioritization of particular reports is a function of administrative discretion within the municipal government; citizens notify the government that a problem exists but have no means directly to affect the level of attention any one report receives. In this sense 311-based coproduction comports to the typology developed by Bovaird (2007); citizens coproduce the service itself, but service delivery and approach are determined by professionals.

That coproduction of this type closely resembles traditional models of civic participation in general naturally raises distributional questions, as studies of civic engagement have consistently found socioeconomic biases in participation rates (Verba, Schlozman, Brady, & Brady, 1995). Some early coproduction scholars recognized the potential for distributional issues in coproduction. Jeffrey Brudney noted that the political economy concerns that could lead to cost offloading from governments to coproducing citizens, with disproportionate negative effects for low socioeconomic neighborhoods (Brudney, 1985). The data-rich nature of 311 systems and their potential reliance on information technology that similarly raises distributional issues makes them attractive targets for empirical exploration of such distributive biases. Minkoff (2013) finds no evidence of inequality across socioeconomic status in New York City’s 311 system. Clark et al. (2013) use data from Boston’s 311 system to test for distributive bias, finding that areas with lower incomes and higher proportions of Hispanic residents are less likely to engage in 311-based coproduction, but other

---

1 This leads, of course, to indirect attempts to draw attention to particular problems, e.g., multiple reports for the same problem or complaints to department administrators and/or elected officials.
racial and ethnic groups use the system in proportion to their share of the population. A subsequent study employed survey data from San Francisco to examine the system’s representativeness, again finding no substantive pattern of bias (Clark & Brudney, 2014).

The nature of how reports are made is another area of research on 311-based coproduction. Over time, 311 systems have transitioned from a phone-only system to one incorporating web- and smartphone-based technologies. These additions are attractive to governments as they reduce reliance on relatively expensive call centers, as well as to citizens that prefer the use of such technologies in lieu of the telephone. Both Clark & Brudney (2014) and Clark et al (2013) devote attention to the impact of different technologies available for submitting reports. In the case of Boston, Clark et al. (2013) find that low-income areas of the city utilize smartphone-based 311 applications for submitting reports at a higher rate than higher-income areas, and argue that this may be evidence of the application’s ability to bridge the participation gap across income strata. Similarly, Clark and Brudney (2014) finds that smartphone-based 311 applications serve to increase coproduction rates among minority groups in San Francisco. There is little, however, in the way of research that evaluates the impact of the e-government evolution in 311 systems on system performance. This paper contributes to the literature on coproduction by empirically assessing the extent to which differing digital technologies used to submit 311 reports appear associated with different times to resolution. As such, the paper links two important approaches to public service innovation: coproduction and e-government technology.

**E-government Technology and Coproduction**

The use of IT to reduce transaction costs in the public sector has been extensively studied (Berry, Berry, & Foster, 1998; Bozeman & Bretschneider, 1986; Lee & Perry, 2002; Moon & Bretschneider, 2002). By the mid 1990s the public sector had integrated many IT systems into its internal structuration and processes (Coursey & Norris, 2008; Moon & Norris, 2005). Around this
time, the rapid development and adoption of the World Wide Web (web) created new opportunities to develop innovative approaches to the provision of public information and services, leading to the deployment of e-government systems.

This particular application of e-government allows for the integration of what have in earlier research been considered separate models of political versus managerial implications of new technology. As Musso, Weare, and Hale (2000) argue, e-government can have managerial as distinct from political effects. The political approach focuses on the effect of e-government technology on civic participation and engagement among previously underrepresented groups within the polity, potentially leading to more democratic processes and outcomes (for an overview, see Musso, Weare, & Hale, 2000; also Moon, Lee, and Roh (2014)), as well as increasing public trust in government (Kim & Lee, 2012). Managerial-oriented research, in contrast, focuses on the managerial and corporate functions of local government, and argues that e-government can improve public sector performance. Work in this vein focuses on how e-government provides information and services to citizens over the internet (Jun & Weare, 2011; Moon, 2002; Musso et al., 2000; Weare, Musso, & Hale, 1999; West, 2004). In both cases, much of the early scholarship predicted both that these changes would occur and that they would manifest rapidly as disruptive innovations (Coursey & Norris, 2008). These expectations, however, failed to live up to empirical scrutiny: adoption of e-government followed a much more incremental path and evidence of their transformative nature was, at best, mixed (Coursey & Norris, 2008; Jun & Weare, 2011; Moon, 2002; Musso et al., 2000; Norris & Reddick, 2013; Weare et al., 1999). This study integrates political considerations regarding engagement and distributional implications of technology use with attention to the managerial implications of differing technology use in 311 systems.

More recent e-government research has focused on the use of “Government 2.0” systems that incorporate technological advances allowing for participatory, collaborative two-way
communication over the internet. This change in the structuration of content production was made possibly primarily through technological developments that drastically lowered the technical, knowledge, and resource barriers to application design and content production (Chang & Kannan, 2008; Chun, Shulman, Sandoval, & Hovy, 2010; Nam, 2012; Reddick & Aikins, 2012). These technologies are seen by some practitioners as essential for engaging the millennial generation as they continue to come of age (Mancini, 2012). Geographic Information Systems (GIS) technology is one example. Ganapati (2011) details the evolution of GIS from an ‘elitist’ system with a steep learning curve to a web-based platform that is seamlessly integrated with other applications. (Ganapati, 2011).

Social media constitute another suite of technologies that have figured heavily into recent e-government research. Adapting social media to existing institutional patterns and behaviors, though, represents a challenge long recognized by e-government scholars (Fountain, 2001). Linders (2012) identifies social media’s capacity to enable multiple forms of coproduction, ranging from pushing information to fostering dialogue between government and citizens. These innovations are consistent with the idea of “radical disintermediation,” wherein the traditional boundaries between government and the citizenry are blurred or eliminated outright (Dunleavy et al., 2006). Much of both the theoretical and empirical literature on specific social media platforms focuses on this dialogic capability of the technology (Ae Chun et al., 2012; Chadwick, 2008; Linders, 2012; Mossberger, Wu, & Crawford, 2013; Smith, 2010; Thomas, 2013). Less, though, is known about the effects of social media platforms on “citizen sourcing” coproduction activities like 311 (Linders, 2012). This paper contributes to this literature by examining how the explicit integration of Twitter, a core social media platform, affects both usage patterns and the time to resolution for 311 reports in San Francisco.
The 311 System in San Francisco

San Francisco’s 311 system offers a unique opportunity to study the effect of synthesizing e-government technologies with service coproduction, as the city has been at the leading edge of technology integration. San Francisco launched its 311 system in March of 2007 with both phone and email capabilities. A custom website for submitting reports followed shortly, with its first report submitted eight months later in November. In 2009, then-Mayor Gavin Newsom announced that Twitter would be incorporated as a medium for submitting 311 reports and following up with the city. The idea arose during an unrelated meeting between Newsom and executives from Twitter, and the Mayor directed the city’s technology department to begin scoping the project shortly thereafter (City of San Francisco, 2009). That same year, the city also began working with other municipalities and the private sector to develop an open standard for accessing 311 systems through Application Program Interfaces (APIs). This standard, eventually dubbed Open311, led to the city launching a smartphone application for use with 311 that went live to the public in 2012 (City of San Francisco, 2013). These applications (private developers also have two different applications available for free) use smartphones’ built-in geospatial system to automatically submit the geocoordinates associated with the reported issue. They also allow the user to take and attach a photograph of the issue with the phone’s integrated camera. San Francisco also rolled out an internal submission system in 2012, whereby city staff can enter reports, presumably in order to track all such issues within a single database for efficiency. Figure 1 shows how the introduction of these different technologies to the 311 system has changed the patterns of use over time through October 21, 2015. Open311-based reports have rapidly grown in overall number and as a proportion of total reports since its introduction, while Twitter usage has remained relatively low in absolute and relative terms and phone has consistently lost share of total reports over time.
Both Open311 and Twitter represent innovative technologies for citizen use in 311-based coproduction. However, they differ substantially in their implementation, with potentially important consequences for the reports made through them. In the case of Open311, development of the technology was driven by professional standards for software development via a public-private consortium of US and Canadian governments and private actors. Open311’s architecture is designed to minimize human error in every stage of the coproduction process. Automatic geocoding minimizes the risk of the citizen misreporting the location of the problem, while attached visual evidence of the problem reduces the risk that city staff will fail to locate it. Open311 also removes the human intermediary in the form of call center staff, eliminating this vector for mistakes. Furthermore, the mobile nature of Open311 may reduce the lag time between a citizen observing a problem and reporting it. Thus,

Hypothesis 1. Service reports made via Open311 should, on average, be resolved faster compared to other submission methods due to reductions in error generation and no lag time between submission and integration into the report queue.

The integration of Twitter with San Francisco’s 311 system, in contrast, appears to be largely motivated by political considerations. The potential marginal gains in efficiency are less clear-cut than Open311, though they are not entirely absent. Twitter’s mobile-oriented platform offers the same potential reduction in time-to-reporting as Open311. It also allows users to attach photographic evidence of the problem. However, Twitter was fundamentally designed as a short-length rapid communication system. The level of detail involved in producing 311 reports may or may not be consistent with a 140-character limit on messages. City staff members are also required to act as intermediaries, translating citizen’s tweets into 311 system-compliant reports and responding to follow-up reports and other general inquiries. The public nature of these conversations and reports represent a path for political considerations to affect response time.
Unlike any other 311 technology, reports made via Twitter are observable by anyone searching for the appropriate hashtag (#SF311). Failing to resolve these reports in a timely fashion could lead to public shaming and agitation for increased accountability within the 311 department and the city government overall. Thus,

Hypothesis 2a. Service reports made via Twitter should be, on average resolved faster because of the combination of observability of reports made and the negative valence associated with issues (especially with attached photos – e.g. human waste on sidewalk).

Hypothesis 2b. Service reports made via Twitter should be, on average resolved slower because of the compatibility issues between the technology behind Twitter and the level of detail needed by the 311 operator to properly input a request into the SF311 database.

**DATA AND METHODS**

This article leverages the City and County of San Francisco’s publically available 311 case database ([https://data.sfgov.org/City-Infrastructure/Case-Data-from-San-Francisco-311-SF311-vw6y-z8j6](https://data.sfgov.org/City-Infrastructure/Case-Data-from-San-Francisco-311-SF311-vw6y-z8j6)) as its primary data source. The database consists of all reports made via 311 that have resulted in a work ticket for the city from July 1, 2008 through October 31, 2015 (n = 1,381,057). It contains information on the nature of the report, the date and time that it was created and, if applicable, closed, its status as of the date the data were accessed (open or closed), the department assigned to resolve it, the technology used to submit the report, and geolocation data in both address and x,y point coordinate format.

Demographic data are drawn at the census tract level from the 2008-2012 five-year American Community Survey dataset provided by the Census Bureau. Geographic Information Systems (GIS) software is used to join these demographic data spatially with 311 reports using the geocoding data within the 311 dataset and TIGER/Line census tract boundaries.

---

2 The final date is an arbitrary function of when the data is accessed by the user; the database is updated daily.
I restrict my sample in several ways. First, I exclude cases involving graffiti on private property; resolution of these cases is the responsibility of the property owner and not the city. Observations are further restricted to reports made within the limits of the City of San Francisco. The County of San Francisco operates two parks outside of the city where reports were made in 2013: Sharp Park in Pacifica, and Camp Mather in Tuolume (Yosemite National Park). The County also operates San Francisco International Airport, which is located in San Bruno. Additionally, reports made outside of both the City and County of San Francisco’s jurisdictions are excluded. The final sample consists of 1,276,549 observations. Descriptive statistics for all continuous variables can be found in Table 1.

I employ survival analysis because my outcome of interest is the time-to-event for requests made through the 311 system. The 311 data can be conceptualized as a classic follow-up observational study, where observations may enter the study at any point between its start date (July 1, 2008) and the last follow-up date (October 31, 2015). Cases that were not closed by the last follow-up date are considered censored; of the 1,276,549 observations, 95,829 (7.5%) are censored. In every case this censoring consists of non-informative (i.e. the censoring mechanism was not conditional on properties of the case) right censoring (Hosmer, Lemeshow, and May, 2008; Klein and Moeschberger, 2003). I use the Cox proportional model because my research questions focus on the effect of a new ‘treatment’ (technology) on case survival, and not on the baseline hazard function itself. The Cox proportional hazard model relaxes assumptions about the hazard function through semiparametric estimation (Cox 1972; Hosmer, Lemeshow, and May, 2008; Klein and Moeschberger, 2003). The model is estimated using maximum likelihood, with the form:

\[ \ln \left( \frac{\lambda(t)}{\lambda_0(t)} \right) = \beta_1 T + \beta_2 D_j + A_j + C_j + Y_t + S_t + \varepsilon_i \]
Where $T$ is an indicator variable for the type of technology used; $D_j$ is a vector of demographic covariates for the census tract where the request was made; $A_j$ are department fixed effects, $C_j$ are request category fixed effects; $Y_t$ are calendar-year fixed effects; $S_t$ are climatological season fixed effects; and $\epsilon_t$ is a stochastic error term. However, postestimation diagnostics for model specification indicate that the fundamental assumption of the Cox hazard model, that baseline hazards remain consistently proportional over time, are violated both for multiple variables (including the treatment effects of interest) and the model as a whole. Estimates generated through exponential and Weibull parametric specifications using gamma-distributed frailties (analogous to random effects) are also included and discussed in turn.

The dependent variable is a dichotomous indicator for whether the observation was closed (i.e., experienced the ‘event’). The model also contains a continuous measure of the time in days between when a report was input into the system and when its status was changed to closed. To account for the fact that 29% of all cases are closed in less than 24 hours, the variable is constructed by dividing the time to closure in hours by 24. Days were chosen as the period for ease of interpretation. Because time-to-event data are almost always right-tailed and must always be positive, this variable is logged.\footnote{This is done automatically by the software package (in this case, STATA)}.

The independent variables of interest are dichotomous indicators for the technology used to submit a report, with phone-based reports serving as the base case. Place-based demographic control variables include the log of population, population density, the proportion of residents aged 65 or older, the proportion of college graduates, proportions of each race and ethnicity, median household income, and the proportion of vacant housing units. It is important to note that these demographic characteristics are features of the area in which reports are made, not of the individual submitting the report.
I include several fixed effects to control for unobservable characteristics that may promote or inhibit case resolution. These include fixed effects for the responsible department to control for unobservable departmental characteristics that may promote or inhibit case resolution, as well as the category of request as different types of service problems are likely to have varying unobservable capital and labor costs that will affect how long they take to resolve. For request categories, I stratify graffiti by whether it is reported as offensive or not offensive, and I create a ‘Noise Complaint’ category for cases with no assigned category but with request detail fields that specify noise as the motivating factor behind case generation (n = 453). I also use fixed effects for the calendar year in which the case was opened to control for variation across years, and for whether the case was opened during San Francisco’s ‘rainy’ season, which runs from November through April (NOAA, 2015). Future refinements of this paper will include stratifying the model by year instead of using a fixed effects approach as an additional robustness check; however using indicator variables for time of entry is supported in the methodological and empirical literature (Allison 1984, 2009; Cabral and Lazzarini, 2014). In all cases, the fixed effect specification soaks the unobservable influences of departmental and service factors, reducing the chance of specification error.

RESULTS

With regard to categories of report by technology used, there are several differences that suggest differential use across technologies. Figure 2 shows the percentage share for the top five categories of report made through each technology available to the public. Reports made by phone are predominantly for street and sidewalk cleaning (39%) followed by abandoned vehicles and general requests at a distant second (8% each). Three of the most reported categories via phone – abandoned vehicles, damaged property (8%) and housing authority requests (4%) – are associated with repeat interaction with a fixed location, suggesting that these reports may be made from a
residence or place of work. Graffiti-based reports are not a common report category for phone submissions (4%), whereas they are either the first or second most common category for all other technologies.

Use of the web submission technology is strongly associated with offensive graffiti – 28% of all web submissions. Street and sidewalk cleaning is the second most common request at 19%. Web reports also have the highest proportion of abandoned vehicle reports (16%) and illegal postings (9%) of any technology. Both web and Open311 have streetlight issues as their fifth-most common report type at 4%. While Open311 also closely resembles the web with regards to graffiti (27%), the type of graffiti is divided between offensive (16%) and non-offensive (11%). However, the dominant report type for Open311 is street and sidewalk cleaning at 41%. Twitter also has a large proportion of graffiti reports (29%), though they are all categorized as non-offensive. Street and sidewalk cleaning reports closely follow at 25%. Twitter also has the largest proportion of reports involving a broken or damaged sidewalk or curb (8%).

Place-based demographic differences in technology use for submitting 311 reports include differences across several socioeconomic strata. Figure 3 shows the mean demographic characteristics of the census tracts in which 311 reports are made by the publically-available technology used. Usage of the telephone call center is associated with places with higher proportions of Asian and Pacific Islander residents, as well as senior citizens. Both phone and web-based usage is associated with areas that have more homeowners. Reports generated through the web are associated with census tracts that have a greater proportion of non-Hispanic Whites and college graduates. Web usage is lowest in areas with higher proportions of Hispanic, foreign-born, and African American residents. Twitter is associated with census tracts with higher proportions of Hispanic and foreign-born residents, but has the lowest rate of association of any technology among non-Hispanic Whites, Asian and Pacific Islanders, homeowners, and college graduates. Open311
usage does not dominate in its association across census tracts with any one demographic group, but its use relative to other technologies is associated with increased proportions of Hispanic, Asian and Pacific Islander, and college graduate residents.

Turning to the model of time to completion for 311 reports, Table 2 summarizes the findings from the exponential (1), Weibull (2), and Cox proportional (3) hazard models. These results are both preliminary and, in the case of the Cox proportional model, likely biased due to misspecification based on diagnostic tests for proportionality. For the sake of brevity, the inclusion of fixed effects is simply noted in the table. Hazard ratios are reported in the table; the following discussion of the results uses percent change in the time to closure for clarity.

The results differ significantly based on whether or not the baseline hazard is fully parameterized. The exponential and Weibull frailty hazard models are parameterized, while the Cox proportional hazard model is not. The exponential and Weibull frailty models both suggest support for hypotheses 1 and 2b; reports submitted through Open311 and Twitter are associated with a shorter time to resolution compared to reports made via phone, while those made via Twitter have coefficients suggesting a longer time to resolution, though without statistical significance, when holding other variables constant. Both specifications suggest that, controlling for other variables, reports submitted through Open311 are associated with 4-5% reductions in time to resolution. In the case of Twitter, the exponential model indicates a 3% increase in time to resolution that is marginally significant (p < 0.10) while the Weibull model estimates for Twitter fail to achieve statistical significance. Both models report web-generated requests as having a closure time 14-15% faster than those made by phone, with internal requests having an 8% decrease in time to resolution, again compared to phone requests.

On the other hand, the Cox proportional hazard model supports the rejection of hypothesis 1 and support for hypothesis 2b. Controlling for demographics, type of report, responsible
department, and year and season submitted, reports made through Open311 have an average resolution time 12% longer than the average report made via phone (p < 0.001). Holding those same control variables constant, reports made via Twitter have an average time to closure 5% longer than the average phone report (p < 0.001). Reports made via the web interface have a time to closure 10% shorter than the mean phone time when controlling for responsible department, type of report, demographics, and time of request (p < 0.001). Interestingly, when holding all other variables constant, reports made by other City and County agencies or departments are not statistically different from those made by non-employees over the phone.

For both the exponential and Weibull models several demographic controls are statistically significant. Controlling for other variables, a 1% increase in population is associated with a 2% increase in the time to resolution (p < 0.001). A 1% increase in the proportion of college graduates in an area is associated with a 24% decrease in the time to resolution (p < 0.001). Compared to the reference category of non-Hispanic whites, a 1% increase in the proportion of non-Hispanic African-American residents is associated with a 50% (exponential) or 51% (Weibull) decrease in the time to resolution when holding other variables at their mean. Similarly, a 1% increase in Hispanic residents relative to the non-Hispanic White reference category is associated with 56% (exponential) or 57% (Weibull) decrease in the time to resolution when controlling for other factors (p < 0.001). Finally, as with non-Hispanic African-Americans and Hispanics, a 1% increase in the proportion of Asian and Pacific Islander relative to non-Hispanic Whites is associated with a 17% decrease (both models) in the time to resolution, holding other variables constant. No other demographic variables achieved statistical significance.

In the Cox proportional hazards model the demographic control variables achieve statistical significance with two exceptions: proportion foreign born and the homeownership rate. Holding other variables at their mean, a one percent increase in the population of the area associated with a
request is associated with a 1% increase in the time to resolution (p < 0.001). The log of population density is significant at p < 0.001 as well, but its effect on the time to completion is indistinguishable from the baseline. A one percent increase in educational (college degree) attainment is associated with a 5% decrease in time to closure, while a one percent increase in the proportion of senior citizens in the census tract where a report was generated is associated with a 6% decrease in the average time to closure.

Controlling for other variables, a one percent increase in the proportion of Hispanic residents in the census tract where a report is made is associated with a 20% decrease in the time to closure (p < 0.001). Holding the same control variables at their mean, a one percent increase in the proportion of African-American residents in the census tract where a report is made is associated with an 11% decrease in the average time to closure (p < 0.001). The effect for the proportion of Asian and Pacific Islander residents is also statistically significant at p < 0.001, but its marginal effect is less than 1/100th of 1% on the time to closure.

**DISCUSSION**

The stark differences in the estimates across parametric and semi-parametric models, combined with the rejection of the proportionality assumptions in postestimation diagnostics suggest that considerable additional work in model specification and data analysis are necessary before drawing conclusions about the effect of different submission technologies on the time to resolution for 311 cases. It may be that the effects of different technologies vary in ways that are effectively different across time. For example, Open311’s geolocation and image attachment features as well as standardized fields for report details reduce the possibility for human error on the part of the user, but it may also be the case that the disintermediation between user and database may lead to more misclassified or inappropriate requests that take longer for the city to investigate and close.
The difference in signs and magnitudes for technologies between both parametric models and the Cox proportional hazard may be attributable to the inclusion of the frailty; while Cox proportional models can include a ‘shared’ frailty it is computationally intensive to generate on such a large sample. The Cox model results have so far been shown to be robust to the inclusion of a shared frailty based on the calendar year as well as different stratification approaches; more tests are necessary. Another option to consider is using time-varying covariates.

It is also possible that the model specification is missing important parameters about the change in what sort of requests are made through the 311 system over time as different technologies are introduced. In the case of Twitter, which has been in operation as a submission technology since 2009, it is questionable whether it will ever see widespread given its consistently low usage rate. Open311, on the other hand, has seen rapid growth in overall usage in both absolute and percentage terms since its deployment in 2012. The fact that Open311 and Twitter dominate sidewalk and curb reports may be a product of the mobile nature of the submission technology, just as phone and web’s dominance of abandoned vehicles, illegal postings, and housing authority reports may be a product of their generally fixed nature. This is possibly suggestive of a shift in the nature of reports made through 311 as Open311 continues to gain usage share. The proportion of types of reports that are more likely to be discovered while in transit, e.g. distressed sidewalks and broken streetlights, may increase relative to other issues.

This may have attendant consequences for how the City and County allocate resources dedicated to fixing problems uncovered through coproduction, especially when considering the impact of performance management regimes. Assuming a finite budget earmarked for the services reported through 311, and a performance management system that structures rewards around absolute numbers and/or proportion of outstanding problems resolved, it is likely that reports
requiring less capital-intensive resolutions (e.g., sidewalk cleaning vs. towing and impounding an abandoned vehicle) will receive preferential treatment.

It is unclear as to why so few graffiti-oriented reports are made by phone. One possibility is that public property graffiti is more likely to be noticed by those traveling to or from work, during which time they could use either Twitter or Open311 to report the problem \textit{in situ}. For those observing graffiti in transit without smartphones, Twitter accounts, and/or the Open311 app, it may be that they are more inclined to use the 311 web interface if they are already using a computer. Similarly, it may be that illegal postings are considered more of a nuisance in commercial areas and are reported by workers or business owners using a desktop computer.

Generally speaking, the demographic findings appear to comport with prior research on the distributional question of 311 usage by citizens insofar as areas with minority and low socioeconomic groups do not appear to suffer from delays in servicing reports (Clark & Brudney, 2014; Clark et al., 2013). In fact, areas with higher proportions of African-American and Hispanic residents have correspondingly shorter times to resolution independent of the type of report or technology used to submit it. Of course, because demographic characteristics can only be attributed to places where reports are made rather than the individuals submitting them, any inferences drawn from demographic analysis in this study should be made with caution. A more systematic analysis of the distributional consequences associated with San Francisco’s 311 system in general and the adoption of new submission technologies in particular is worthy of additional research in a future study.

**CONCLUSION**

This study contributes to the literature on coproduction and e-government by empirically assessing the differences in the time it takes for the City and County of San Francisco to close 311
reports across multiple technologies using survival analysis. This study also addresses how the use of these technologies varies across types of reports made and the place-based demographics associated with system use. Results for the effect of technology on resolution time vary significantly based on model choice and should be interpreted with extreme caution. Analyses of patterns of use suggest that both Twitter and Open311 have higher reporting rates for issues associated with observations made on foot. Phone and web requests, meanwhile, had significantly higher reporting rates for abandoned vehicles, possibly suggestive of usage associated with longer periods spent in a particular location. Demographic analysis found that rates of Twitter usage are associated with increased proportions of Hispanic residents but less so with non-Hispanic Whites, college graduates, and homeowners. Open311 usage is moderately associated with higher proportions of both groups as well as African Americans, Asian and Pacific Islanders, and college graduates. The use of different technologies also differs across other demographic dimensions, including the concentration of senior citizens, college graduates, and homeowners. These demographic-related findings should be interpreted with caution, however, given that they are based on demographic characteristics of the places where reports are made rather than people making the reports.

Indeed, there are several limitations to this study that warrant attention. First, it only examines usage and outcomes in one jurisdiction. Moreover, while it is often useful as a case study for early adopter phenomena, San Francisco’s idiosyncrasies call for caution in generalizing findings to other major metropolitan areas, although it should be noted that it does support study of 311 systems in Boston. It would be helpful to replicate the results from Boston and San Francisco, both cities that have a relatively concentrated population of highly educated and higher income residents, with cities that have lower levels of education and higher concentrated poverty.

As mentioned earlier, another key limitation is that the demographic data are limited to the 2010 census tract level and not associated with discrete 311 reports. This is a common issue with
research on 311 systems, as they do not collect demographic information on users by design. Clark and Brudney (2014) circumvent this problem by using survey data on 311 users from the City of San Francisco, but this comes at the expense of both introducing issues of bias endemic to survey responses and not being able to employ the report data, making it inappropriate for answering the questions motivating this study. This analysis would also benefit from a distinction between phone reports made via landline and those made via cellular phone, but San Francisco’s 311 database does not support such granularity.

While this study analyzes the differences in resolution time associated with different reporting technologies, it does not address questions about the specific impact of departments and types of reports on these times. Future research will be devoted to examining whether there is variation in ‘buy in’ on coproduction across different types of municipal departments and agencies. Finally, the questions as well as the results in this study beg for contextualization via interviews with past and present City and County staff involved with the 311 system. Future work on this study will focus on obtaining this key qualitative data.
References


FIGURES AND TABLES

Figure 1. 311 Usage Rates By Different Reporting Technologies Over Time

Request Medium Usage Over Time

Thousands


Phone
Web
Twitter
Open311
Internal
Figure 2. Share of Reports by Technology (Top 5 Categories)

**Phone**
- Street and Sidewalk: 39%
- Abandoned Vehicle: 8%
- General Requests: 8%
- SFHA Requests: 7%
- Damaged Property: 6%

**Web**
- Graffiti - Offensive: 28%
- Street and Sidewalk: 19%
- Abandoned Vehicle: 16%
- Illegal Postings: 9%
- Streetlights: 4%

**Open311**
- Street and Sidewalk: 41%
- Graffiti - Offensive: 16%
- Graffiti - Not Offensive: 11%
- Illegal Postings: 6%
- Streetlights: 4%

**Twitter**
- Graffiti - Not Offensive: 29%
- Street and Sidewalk: 25%
- Sidewalk or Curb: 8%
- 311 External Request: 6%
- General Requests: 5%
Figure 3. Demographic Characteristics of Report Locations by Technology

**Mean % Non-Hispanic White**

- Open311
- Twitter
- Web
- Phone

**Mean % Hispanic**

- Open311
- Twitter
- Web
- Phone

**Mean % African American**

- Open311
- Twitter
- Web
- Phone

**Mean % Asian & Pacific Islander**

- Open311
- Twitter
- Web
- Phone

**Mean % Foreign Born**

- Open311
- Twitter
- Web
- Phone

**Mean % Seniors**

- Open311
- Twitter
- Web
- Phone

**Mean % College Graduates**

- Open311
- Twitter
- Web
- Phone

**Mean % Homeowners**

- Open311
- Twitter
- Web
- Phone
Table 1. Summary Statistics For All Continuous Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time to closure (days)</td>
<td>8.28</td>
<td>0.72</td>
<td>8.38</td>
<td>3.30</td>
<td>9.31</td>
</tr>
<tr>
<td>Population (log)</td>
<td>30084.30</td>
<td>21510.79</td>
<td>26509.49</td>
<td>16.09</td>
<td>164627</td>
</tr>
<tr>
<td>% Age 65+</td>
<td>0.13</td>
<td>0.07</td>
<td>0.12</td>
<td>0</td>
<td>0.61</td>
</tr>
<tr>
<td>% College Graduates</td>
<td>0.51</td>
<td>0.20</td>
<td>0.53</td>
<td>0.06</td>
<td>1</td>
</tr>
<tr>
<td>Median Household Income (log)</td>
<td>11.10</td>
<td>0.57</td>
<td>11.23</td>
<td>9.36</td>
<td>11.96</td>
</tr>
<tr>
<td>% Hispanic</td>
<td>0.17</td>
<td>0.14</td>
<td>0.12</td>
<td>0</td>
<td>0.59</td>
</tr>
<tr>
<td>% African-American</td>
<td>0.07</td>
<td>0.10</td>
<td>0.03</td>
<td>0</td>
<td>0.68</td>
</tr>
<tr>
<td>% Asian and Pacific Islander</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Foreign Born</td>
<td>0.35</td>
<td>0.14</td>
<td>0.35</td>
<td>0.06</td>
<td>0.83</td>
</tr>
<tr>
<td>% Vacant Units</td>
<td>0.10</td>
<td>0.06</td>
<td>0.09</td>
<td>0</td>
<td>0.32</td>
</tr>
</tbody>
</table>
### Table 2. Hazard Model Results on Time to Resolution for 311 Reports

<table>
<thead>
<tr>
<th></th>
<th>(1) Exponential</th>
<th>(2) Weibull</th>
<th>(3) Cox Proportional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open311</td>
<td>1.05*** (0.007)</td>
<td>1.05*** (0.007)</td>
<td>0.879*** (0.003)</td>
</tr>
<tr>
<td>Twitter</td>
<td>0.98 (0.015)</td>
<td>0.98 (0.016)</td>
<td>0.953*** (0.008)</td>
</tr>
<tr>
<td>Web</td>
<td>1.15*** (0.007)</td>
<td>1.15*** (0.007)</td>
<td>1.096*** (0.003)</td>
</tr>
<tr>
<td>Internal</td>
<td>1.08*** (0.014)</td>
<td>1.08*** (0.014)</td>
<td>1.007 (0.007)</td>
</tr>
<tr>
<td>Population (log)</td>
<td>0.98*** (0.005)</td>
<td>0.98*** (0.005)</td>
<td>0.986*** (0.002)</td>
</tr>
<tr>
<td>Population Density (log)</td>
<td>1.00 (0.003)</td>
<td>1.00 (0.003)</td>
<td>1.000*** (0.000)</td>
</tr>
<tr>
<td>% Age 65+</td>
<td>0.95 (0.036)</td>
<td>0.95 (0.036)</td>
<td>1.068*** (0.021)</td>
</tr>
<tr>
<td>% College Graduates</td>
<td>1.24*** (0.035)</td>
<td>1.25*** (0.035)</td>
<td>1.050*** (0.013)</td>
</tr>
<tr>
<td>Median Household Income (log)</td>
<td>1.00 (0.009)</td>
<td>1.00 (0.008)</td>
<td>1.015*** (0.004)</td>
</tr>
<tr>
<td>% Hispanic</td>
<td>1.56*** (0.052)</td>
<td>1.57*** (0.053)</td>
<td>1.204*** (0.015)</td>
</tr>
<tr>
<td>% African-American</td>
<td>1.50*** (0.048)</td>
<td>1.51*** (0.05)</td>
<td>1.106*** (0.017)</td>
</tr>
<tr>
<td>% Asian and Pacific Islander</td>
<td>1.18*** (0.038)</td>
<td>1.18*** (0.039)</td>
<td>1.000*** (0.000)</td>
</tr>
<tr>
<td>% Foreign Born</td>
<td>0.95 (0.032)</td>
<td>0.95 (0.033)</td>
<td>0.971 (0.016)</td>
</tr>
<tr>
<td>% Homeowner</td>
<td>0.97 (0.015)</td>
<td>0.97 (0.016)</td>
<td>0.993 (0.008)</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Likelihood Ratio Chi-Square</td>
<td>622886.25</td>
<td>603358.72</td>
<td>602769.4</td>
</tr>
<tr>
<td>N</td>
<td>1,276,549</td>
<td>1,276,549</td>
<td>1,276,549</td>
</tr>
</tbody>
</table>

Standard errors in parentheses:  * p<0.05  ** p<0.01  *** p<0.001

Income Reported in 2012 Dollars