Do Elections Improve Constituency Responsiveness? Evidence from U.S. Cities*

Darin Christensen† and Simon Ejdemyr‡

January 24, 2017

Abstract

Do elections motivate incumbent politicians to serve their voters? In this paper we use millions of service requests placed by residents in U.S. cities to measure constituency responsiveness. We then test whether an unusual policy change in New York City, which enabled city councilors to run for three rather than two terms in office, improved constituency responsiveness in previously term-limited councilors' districts. Using difference-in-differences, we find robust evidence for this. Taking advantage of differential timing of local election races in New York City and San Francisco, we also find late-term improvements to responsiveness in districts represented by reelection seeking incumbents. Elections improve municipal services, but also create cycles in constituency responsiveness. These findings have implications for theories of representative democracy.

*We are grateful to Michael Bechtel, Gary Cox, Nick Eubank, Francisco Garfias, Justin Grimmer, Jens Hainmueller, David Hausmann, Clayton Nall, Julia Payson, and participants of Stanford's American Politics Workshop for comments on earlier drafts.

†Assistant Professor of Public Policy, UCLA Luskin School of Public Affairs, 337 Charles E. Young Dr. East, Los Angeles, CA 90095-1656. UCLA. Email: darinc@luskin.ucla.edu. Phone: (310) 825-7196. Darin acknowledges the support of the Stanford Graduate Fellowship.

‡Corresponding author. Ph.D. Candidate, Department of Political Science, Stanford University, 616 Serra Street Room 100, Stanford CA 94305. Email: ejdemyr@stanford.edu. Phone: 650-714-8905.

A prominent conception of representative democracy, dating back to at least James Madison, holds that periodic voting promotes political accountability (Madison [1788] 1966). This "electoral connection" (Mayhew 1987) encourages representatives to serve their constituents for fear of being ousted on election day. Prior empirical research on whether such a connection exists falls into two categories. First, do representatives shirk in their final term in office (when the electoral connection is absent)? Second, do representatives shirk when elections are distant in time (when the electoral connection is weakened)?

In this paper, we shine new light on these questions using a new measure of constituency services as well as a new quasi-experimental research design. Addressing constituency requests is a central activity for most elected officials (Cain, Ferejohn and Fiorina 1987; Fiorina 1989, ch. 7; King 1991; Mayhew 1987).[1] Yet, relatively few studies assess whether electoral incentives improve constituency responsiveness — or, conversely, whether a weaker electoral connection causes politicians to shirk on this activity.[2] We therefore collect data on more than 15 million service requests placed by residents in New York City (NYC) and San Francisco (SF). We then link these data with the election districts of local city councilors to study how response times to service requests are shaped by councilors' electoral incentives.

To do this, we first take advantage of an unusual policy change in NYC, where city councilors voted in 2008 to extend their own term limits from two to three terms. This policy change allows us to implement a difference-in-differences strategy, comparing changes to response times in districts with newly eligible councilors to changes in districts represented by first-term councilors, who were always eligible for reelection. This strategy eliminates many confounders that could bias the relationship between electoral incentives and incumbent effort, such as cross-sectional quality differences between politicians (due to skill or experience) and time shocks that affect responsiveness among all representatives. The results from this analysis indicate that elections substantially improve constituency responsiveness.

To further assess the importance of elections, and to extend our analysis beyond NYC, we also analyze whether incumbents are less responsive earlier in their terms, when voters direct less attention to politics

[1]For example, Mayhew (1987, 22) observes, "For the average congressman the staple way [to claim credit] is to traffic in what may be called 'particularized benefits,' [the bulk of which] come under the heading of 'casework' — the thousands of favors congressional offices perform for supplicants in ways that normally do not require legislative action. Each office has skilled professionals who can play the bureaucracy like an organ — pushing the right pedals to produce the desired effects."

[2]In Table A.1, we summarize the empirical literature on last-term shirking in the United States. Of the 26 empirical papers we surveyed, only two analyzed a measure of constituency services, and in both cases this measure was self-reported by state legislators (Carey, Niemi and Powell 1998; Carey et al. 2006).

(Lenz and Healy 2014; Huber, Hill and Lenz 2012). Term limits and staggered elections mean that incumbents within the same city run for reelection at different times. Thus, we compare changes to responsiveness among reelection-seeking incumbents to changes among incumbents who either have a reelection bid at a later time or are ineligible to seek reelection. We find that, while constituency responsiveness improves in all districts as elections approach, it improves much more rapidly in districts represented by reelection-seeking incumbents. The flip side of this finding, of course, is that incumbents exert relatively less effort earlier in their terms.

In addition to providing placebo tests to shore up the validity of our research design, we address two alternative interpretations of these results. First, we ensure that the results are not driven by *effort reallocation* from or to legislative representation. For example, NYC councilors that became eligible for a third term in 2008 may have heightened their efforts with constituency services at the expense of reduced legislative activity. Our analysis shows, however, that incumbents' efforts with legislation remained constant even as they were becoming more responsive to constituency concerns. Second, we do not find that constituents submitted more (or fewer) requests in districts where councilors became eligible for a third term. Our results are driven by the supply of constituency service, not changes in demand for councilors' time.

Our findings support the conception of representative democracy articulated by Madison. They also bolster prominent political economy models on elections (Alt, Bueno de Mesquita and Rose 2011; Besley 2006; Dewan and Shepsle 2011; Nordhaus 1975; Rogoff 1990; Shepsle et al. 2009; Tufte 1978; Przeworski, Stokes and Manin 1999). Despite elegant predictions, these models have been refuted (e.g., Besley 2006; Carson et al. 2004; Lott and Bronars 1993; Poole and Romer 1993; Keele, Malhotra and McCubbins 2013) almost as many times as they have been supported (e.g., Alt, Bueno de Mesquita and Rose 2011; Besley and Case 1995; Cummins 2012; Figlio 1995; Rothenberg and Sanders 2000; Snyder and Ting 2003).[3] Our results indicate that, in the context of two major U.S. cities, these models are correct in predicting both that elections discipline politicians and that they create cycles in incumbent responsiveness.

Elections and Shirking: Theoretical and Empirical Background

A large literature in political science shows how the electoral connection — which normally forces politicians to exert costly efforts on behalf of constituents — can be severed or attenuated. First, incumbents entering their

[3]Franzese (2002, 378) reviews the literature on electoral cycles and concludes, "On balance, then, the empirical literature uncovers some possible, but inconsistent and weak, evidence for electoral cycles in macroeconomic outcomes, with evidence for cycles in real variables generally weakest (but not wholly absent)." See Canes-Wrone and Park (2012), Grier (2008), and Krause (2005) for more recent contributions to this literature.

last terms in office no longer need to worry about voters punishing them at the polls for their (in)actions (e.g., Besley and Case 1995, Figlio 1995, Rothenberg and Sanders 2000). Second, when elections are distant in time, voters pay less attention to their representative's activities. It is in the immediate run up to elections that voters direct their attention to politics and, in so doing, discipline politicians (e.g., Huber and Gordon 2004; Nordhaus 1975; Shepsle et al. 2009; Tufte 1978).[4]

To clarify these ideas, consider a simple maximization problem in which incumbents choose how much effort $e_t \in [0, \bar{e}]$ to exert at a cost c in each of T periods (e.g., months) in their term. If eligible, effort increases their probability of reelection, represented by the function $\gamma(\cdot)$.

$$\max_{e_1,\ldots,e_T} \left\{ \sum_{t=1}^{T} \mathbb{1}(\text{Eligible})\gamma(e_t, t) - ce_t \right\}$$

Incumbents maximize their payoff by selecting the effort level that equalizes the marginal benefit and cost of effort in each period (i.e., $e_t^* = \arg\{\mathbb{1}(\text{Eligible})\ \gamma_e(e,t) = c\}$). While ineligible incumbents would never want to exert themselves, incumbents seeking reelection make some effort in every period, as the returns to doing so are always positive, even if sometimes minuscule.[5] Thus, *term limits reduce incumbent effort* — the first prediction we test in this paper.[6]

Now suppose that the returns to effort increase as the next election approaches.[7] Past research has offered two reasons for this. First, if voters suffer from recency bias (e.g., Lenz and Healy 2014; Huber, Hill and Lenz 2012), incumbents concentrate efforts just before their reelection contests — the period that weighs most heavily on voters' minds when they cast their votes (Nordhaus 1975; Shepsle et al. 2009; Tufte 1978). Second, even if voting is prospective rather than retrospective, reelection-seeking incumbents may ramp up their efforts as elections approach to signal their superior competence (Rogoff 1990).[8] These two strands of the literature both imply that the optimal level of effort for eligible incumbents increases as elections approach (i.e.,

[4]In addition to term limits and electoral proximity, simply having elections at all is frequently used to measure the impact of electoral incentives. In particular, past research has studied the impact of elections versus appointments, finding that elected representatives are more responsive (Grossman et al. 2014) and serve a broader set of constituents (Sances 2016) than appointed leaders.

[5]$\gamma_e(e,t) > 0$ for every $t \in \{1, \ldots, T\}$

[6]For a review of models making this and related claims, see Dewan and Shepsle (2011).

[7]Mathematically, $\gamma(e,t)$ is a continuous function with increasing differences.

[8]See Besley (2006) and Canes-Wrone, Herron and Shotts (2001) for other signaling models in this tradition.

$e_t^* > e_{\underline{t}}^*$ for any $t > \underline{t}$). Thus, *eligible incumbents should increase their effort levels over the course of their terms (while effort among ineligible incumbents should remain constant).*[9]

Empirical Challenges

Despite the clarity of these two predictions, one can find empirical studies that claim to support and refute both of them. Table A.1 provides evidence for this. This table summarizes 26 studies of last-term shirking in the United States. Of these studies, half find evidence of shirking while the other half find no or inconclusive evidence.[10]

Table A.1 highlights three ways in which empirical studies can be extended to potentially resolve or clarify these mixed results. First, past research has focused on "ideological shirking," analyzing politicians' voting records and policy outcomes while in office.[11] Constituency services — one of the most common activities in the daily lives of representatives (Cain, Ferejohn and Fiorina 1987; Fiorina 1989, ch. 7; Mayhew 1987) — have received less attention. Our review revealed two studies of this activity (Carey, Niemi and Powell 1998; Carey et al. 2006), both of which use measures of constituency services that were self-reported by state legislators.

Second, most studies of shirking focus on only one incumbent activity. Doing so makes it difficult to distinguish between two different outcomes: a shirking incumbent and an incumbent who is reallocating effort across activities. For example, if a last-term incumbent decides to devote ten fewer hours per week to legislation but allocates twenty additional hours per week to casework, a study focusing on legislation may wrongly conclude that the incumbent shirked in her last term. To reduce the possibility that such reallocation could be driving our results, this paper analyzes legislative activity in addition to constituency services. We acknowledge that these activities capture far from all ways in which representatives serve constituents; as Lott (1990, 133) points out, there are "as many [potential measures of effort] as there are outputs that a politician produces." We present the analysis of legislative action as a suggestive test of reallocation across two incumbent activities.

---

[9] This aligns with a large literature on electoral cycles in incumbents' behavior (Schumpeter 1939; Nordhaus 1975; Tufte 1978; Rogoff 1990; Schultz 1995; Franzese 2002; Canes-Wrone and Park 2012).

[10] Franzese (2002) and Canes-Wrone and Park (2012) review empirical results in the literature on electoral cycles in incumbent effort and reach the same conclusion.

[11] Studies of ideological shirking can in turn be divided into three categories, analyzing whether last-term incumbents (1) vote differently than they have previously (e.g., Lott 1987; Lott and Bronars 1993; Snyder and Ting 2003), (2) vote in opposition to their constituents' preferences (e.g., Besley 2006; Wright 2007; Tien 2001), or (3) favor a different set of fiscal policies (e.g., Erler 2007; Keele, Malhotra and McCubbins 2013).

Third, some research on shirking has relied on cross-sectional comparisons of legislators that are or are not in their last terms in office. Omitted variables that are difficult to measure, such as motivation or quality, may threaten causal inference in such studies. Our design extends more recent work that exploits panel data on officials' behavior (e.g., Alt, Bueno de Mesquita and Rose 2011; Bails and Tieslau 2000; Besley and Case 1995; 2003; Besley 2006; Erler 2007; Keele, Malhotra and McCubbins 2013; Snyder and Ting 2003). These studies compare changes to behavior over time, reducing confounding due to fixed incumbent characteristics. They also avoid selection concerns associated with retirement decisions by restricting comparisons to incumbents whose election eligibility is mandated by law rather than chosen.[12] We attempt to build on such studies by studying councilors who serve within the same city and deal with similar constituency requests.

Measuring Constituency Responsiveness in U.S. Cities

We collect detailed data on service requests in NYC and SF. These cities log information on each service request filed by residents via 3-1-1, a system recently implemented in many major U.S. cities to redirect non-emergency requests from 9-1-1 and to centralize hotlines maintained by individual city agencies. The NYC database has around 14 million observations going back to 2005; the SF database, around 1 million observations going back to 2008.[13]

Three aspects of these data allow us to measure local responsiveness to constituency concerns. First, we use the dates a request was opened and closed to measure *response times*.[14] We discuss what it means for a request to be opened and closed below. Second, we use the reported *location* of each request to match it with a council district boundary. Both NYC and SF have single-member districts (51 in NYC and 11 in SF), meaning that requests and council members are uniquely matched. Third, the data contain information about *request type* (e.g., public housing request, pothole, abandoned vehicle), which allow us to account for different response times to different types of requests in our analysis. The most frequent request types in the NYC database (2004-2013), alongside response time statistics, are displayed in Table 1.[15]

---

[12] One concern in our study may be that NYC councilors amended the law governing their election eligibility. We discuss the implication of this for our results below, noting that it likely biases against finding an effect of the term limit extension (see p. 10). We also show that our results are not driven by newly eligible councilors who supported the reform (SI 3).

[13] The NYC data are available at `https://nycopendata.socrata.com/`, and the SF data at `https://data.sfgov.org/City-Infrastructure/Case-Data-from-San-Francisco-311-SF311-/vw6y-z8j6/data` (as of winter 2016).

[14] We use the date a request was opened to determine when it occurred relative to either the term extension or election.

[15] In the subsequent empirical analysis, we trim the top 0.1% (for NYC 2008) or 1% (other analyses) of observations in terms of response times to eliminate large, potentially influential outliers. These different trimming rules are based on

Table 1 Summary Statistics for the Most Frequent Request Types in NYC (2004-2013)

Complaint Type	Frequency (in millions)	Percent of all requests	Cumulative percent	**Response times in days** Mean	Median	Trimmed mean†
Construction/Plumbing	2.3	16.4	16.4	29.3	12.0	24.0
Heating	2.1	14.7	31.1	5.5	4.0	5.2
Bridge/Highway/Street	1.2	8.7	39.8	4.4	1.0	3.1
Noise	1.2	8.2	48.0	4.3	0.0	4.1
Sanitation/Cleaning	1.0	6.8	54.8	3.2	1.0	3.1
Paint/Graffiti	0.8	5.7	60.5	31.4	13.0	27.2
Sidewalk/Sewer	0.7	4.9	65.4	44.7	1.0	11.1
Water	0.7	4.7	70.1	6.2	0.1	5.1
Construction-related	0.5	3.9	74.0	50.7	16.0	30.7
Street Light Condition	0.5	3.3	77.3	13.3	1.0	9.7
Other‡	3.2	22.9	100	39.3	5.0	21.1

†Excludes response times above the 99th percentile.
‡Includes 88 complaint types, which we keep as separate categories in the analysis.

How City Councilors Impact Service Responsiveness

We interviewed 3-1-1 representatives, heads of city agencies, and councilor staff to better understand how service requests are handled by city bureaucracies and to what extent councilors intervene to impact service responsiveness.[16] Based on these interviews, Table A.2 provides an outline of how requests are handled from the time they are submitted by a resident until they are closed by agency staff. A request has been *opened* when all the intake information about the request has been logged and it has been assigned to an agency. An agency will *close* the request after resolving it (which may require rerouting it to a different agency) or after determining that no action is necessary. In all cases, agency workers will physically inspect the issue to determine what type of work is needed.[17] More than anything, Table A.2 highlights the important role individual agencies play in

the number of large outliers in each of the samples. These rules result in response time distributions that are quite similar across the samples used for analysis. In SI 6, we demonstrate that our results are robust to different decisions about whether and how much to trim the data. We also dichotomize response times (e.g., more or less than five days) and run linear probability models, confirming that our conclusions are not driven by outliers.

[16]The interviewees included NYC's 3-1-1 director of communication, representatives from the Departments of Housing and Transportation familiar with the 3-1-1 process, and staff members from nine city council offices.

[17]There is no information in the database about the actual action taken by the agency — that is, whether an agency closed the issue after actually resolving it (4a-b in Table A.2) or after determining that the request was "non-warranted" (4c). It is unlikely, however, that differences in non-warranted requests could impact our results. First, the 3-1-1 director of communication and agency representatives indicated that a very small proportion of requests are non-warranted (e.g., prank calls). Second, even if that were not true, difference-in-differences, which we use below, account for baseline

the 3-1-1 system: 3-1-1 provides a centralized and standardized way for requests to be submitted to various agencies. Once there, agencies are responsible for resolving requests and can prioritize across different requests as they see fit.

Given the central role agencies play in resolving requests, it is not surprising that council staff say they regularly turn to agencies to address concerns about service responsiveness in their district. This happens both at a small and a large scale. At a smaller scale, all city council offices we talked to help residents with individual service requests. (Well-staffed offices have a "constituency services" team devoted just to this.) Often, this involves helping the resident file a 3-1-1 request. The councilor office will then follow up with agency intergovernmental liaisons or other agency staff to make sure city workers respond to the request as quickly as possible. Council staff said they are in contact with agencies on a daily basis. They also spoke about the effectiveness of these efforts. For example, a council staff member said that contacting agencies about a specific constituency concern "really smooths things along." Another council member concerned about over-development regularly monitors and responds to constituency concerns filed via 3-1-1 or the Department of Buildings, and is known for his success in limiting new housing development in the district. This type of action shows that councilors are highly motivated and able to help residents with service issues.

Our interviews also revealed two ways in which council offices can — and do — impact response times at a larger scale. The first and most common way is to inform agencies about broader issues within the district. Several district offices said the council member or an office representative meets frequently with agency commissioners or intergovernmental liaisons. For example, a staff member, speaking of sanitation and transportation issues in the district, told us "[our district representative] discusses the specific issues that constituents have so that they [i.e., the agencies] are aware of them and so that they can take appropriate measures." Office staff also send letters to agencies. For example, a staff member said her office often compiles issues their constituencies have and sends a letter addressed to the agency commissioner (e.g., they recently sent a letter to the Department of Transportation regarding potholes in the district).

The second way in which council districts affect constituency responsiveness at a larger scale is by working with other council members, forming task forces or taking legislative action. For example, a representative from a city council office said that, when they notice that an issue is prevalent, they have a meeting with other councilors, especially councilors that represent similar districts. Then if enough agree, they launch a task force

---

cross-sectional differences in non-warranted requests as well as for over-time changes that affect both treatment and control districts. Our inferences would only be threatened in the unlikely scenario that *changes* to response times to non-warranted requests were different in treatment and control districts *and* these changes were such that they improved overall response times more in treatment districts.

that consists of central city council administrators, agency representatives, and city council members. The representative gave an example concerning a request regarding special education. Noticing that there were not enough resources for special education within the district, the council office formed a task force that implemented a program to better integrate special education children into the public school system.

Although we cannot directly quantify the effectiveness of these particular actions in terms of response times, the interviews highlight plausible mechanisms for how elections influence local responsiveness. When councilors are no longer eligible to seek reelection, or when elections are distant in time, councilors are less motivated — and they allocate less time and staff resources — to pressure agencies to impact response times in their district.

Lastly, this discussion may raise the question of why agency commissioners and staff heed the demands of city politicians in the first place. First, the monitoring problems found in many principal-agent relationships are limited in our case. This is because the data collected by the 3-1-1 system can be used to monitor response times at the council district level. City councilors told us they use these data, as well as reports released by the city, to track responsiveness in their district.[18] Second, council staff report building professional and personal relationship with agency staff. Whether because of social or quid pro quo benefits, such relationships could be used to get agencies to reallocate resources when necessary. Third, if agency commissioners care about their budgets, then they should strive to do well by elected officials, who approve the city budget, including funding for both the 3-1-1 program and city agencies. Allocations to these agencies are not guaranteed year-to-year. For example, between fiscal year 2007 and 2010, the Department of Public Works (DPW) in SF saw its annual general fund allocation drop from nearly $27.9 million to $13.4 million (Dept. of Public Works 2010).

Effects of Reelection Eligibility on Constituency Responsiveness

To estimate a causal effect of reelection eligibility on politicians' efforts, we take advantage of the term-limit extension instituted by Mayor Bloomberg and the New York City Council on October 23, 2008. The extension enabled the mayor and city councilors to run for three rather than two four-year terms in office. But it did not affect every city councilor equally.[19] A subset of councilors (14 of 51) were in their first term of office at the time of the decision, and would have been eligible for another term regardless. Another group of incumbents suddenly went from being term-limited to eligible for reelection.

[18] Local Law 47 in NYC requires the 3-1-1 service to make periodic public reports with call data aggregated by city council district. In SF, the 3-1-1 data include information on the supervisor district in which the service request is located.

[19] We consider only city councilors, not the mayor, in our analyses.

Empirical Strategy

The term-limit change allows us to implement a difference-in-differences (DiD) design. Our treatment group consists of incumbents who were termed out before the October 23, 2008 decision but ran for a third term after the decision. This group has 29 incumbents, as not all of the 37 newly eligible councilors took advantage of the extension.[20] Our control group is incumbents who were allowed to seek reelection both before and after the decision. We estimate the DiD using a linear model with council-district and period fixed effects:

$$y_{idt} = \alpha_d + \delta_t + \beta D_{dt} + \gamma_{type} + \phi_{day} + \varepsilon_{idt} \tag{1}$$

where i indexes complaints, d city council district, and t time period (i.e., before or after the term limit extension). The outcome variable is the number of days it took to resolve the complaint. D_{dt} is an indicator equal to 1 for treated city council districts after the term-limit extension and 0 otherwise. The parameter associated with this variable, β, is of key interest. A negative estimate of β would indicate that response times dropped — improved — in treatment districts relative to control districts after the term limit extension. We discuss under what conditions this estimate can be interpreted causally below. Note that the model also includes fixed effects for council district (α_d) and time period (δ_t). The latter is simply an indicator for the period after the extension. Because response times vary by the type of request — as Table 1 makes clear — we also include fixed effects for complaint type (γ_{type}) and the day of the week on which the complaint was lodged (ϕ_{day}). In all analyses, we cluster the standard errors on councilor.

For β to provide a causal estimate of the effect of the term limit extension on response times, the parallel trends assumption must hold.[21] That is, in the absence of the term limit extension, treatment and control districts must have followed the same trend in responsiveness. Note that this assumption does not imply that

[20] The other eight incumbents left politics or ran for different positions (e.g., four ran for comptroller). We do not include these individuals in the main analysis, but show that our results hold including all incumbents in SI 2.

[21] We recognize that to recover the average treatment effect on the treated (ATT), two additional assumptions — the stable unit treatment value (SUTVA) and constant treatment effects — are required. Spillovers could result from city-wide improvements to response times, e.g., due to Mayor Bloomberg's reelection bid, in line with Levitt (1997). However, this would bias $\hat{\beta}$ toward 0. In addition, we note that ATT does not generalize to the control group (by definition). In our case, features of the treatment group — for example, their additional experience in office — may impact the size of the treatment effect. This does not violate the identifying assumption. ATT is a relevant quantity to the extent that most elected officials can stay in office for more than one term before term limits are imposed.

treatment and control districts must be balanced on levels.[22] For example, the design accounts for the fact that second-term councilors may have better average response times due to their longer tenure in office. The unit fixed effects in Equation 1, represented by α_d, account for all fixed differences (whether observed or unobserved) across councilors and council districts.

What, then, could call into question the parallel trends assumption? First, treated and control councilors could have different pre-treatment trends due to the upcoming election, which took place in November 2009. In fact, if our second hypothesis regarding election timing is correct, reelection eligible incumbents (the control group) should be ramping up their efforts in anticipation of the election, relative to ineligible incumbents. Fortunately, if this is true, it would bias against finding an effect of the term limit extension, meaning that our estimate of β provides a lower bound on the effect of the extension.

Second, given that city councilors approved the extension, it is possible that they could have anticipated that it would pass. If so, treated incumbents may have ramped up their efforts with constituency services before October 23, when the extension was formally approved in City Hall. Again, this would bias the estimate of β toward zero, making it more difficult to find an effect of the policy change. Furthermore, evidence from the time the extension was debated suggests that anticipatory effects were limited.[23] Lastly, we carry out a set of "placebo" tests below, substituting the actual date of passage with a set of earlier, fake dates. We find no evidence of diverging pre-treatment trends, suggesting again that anticipatory effects were limited.

Results

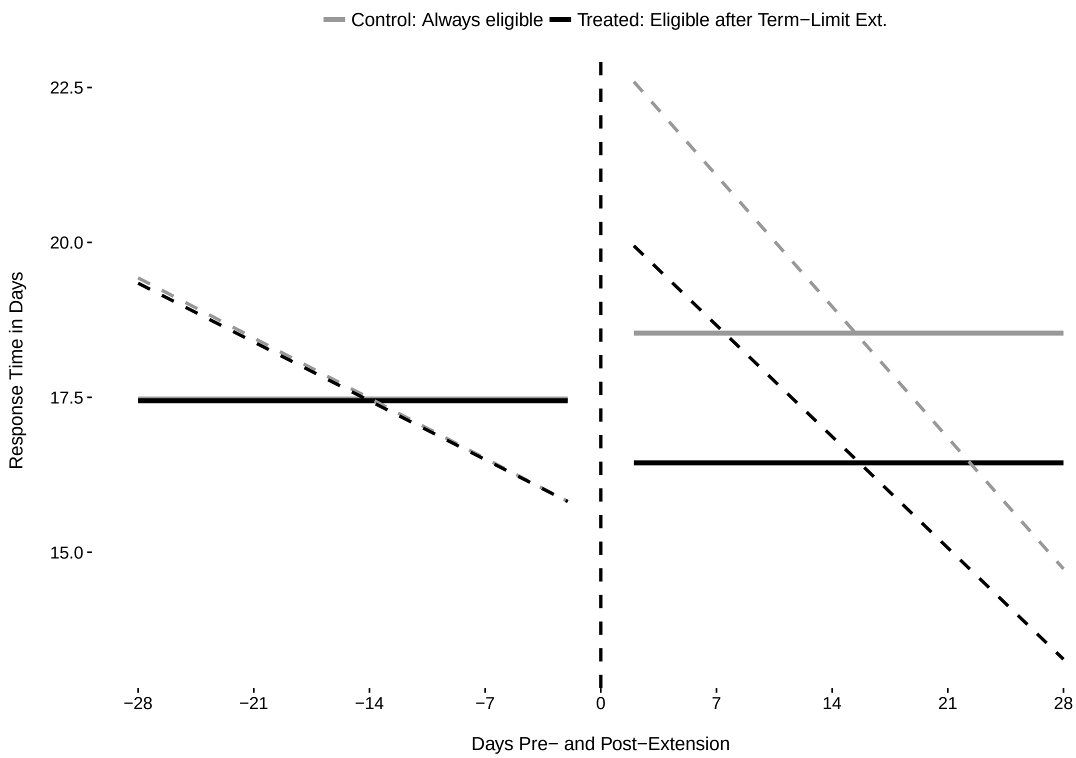
We begin by presenting visual evidence that reelection eligibility improves responsiveness to service requests. Figure 1 compares response times in the four weeks before and after the term-limit extension for districts that were (in black) and were not (in gray) affected by the term-limit extension. The solid lines show that, relative to control districts, average response times improved in treated districts after the extension, consistent with the idea that elections motivate responsiveness. Slower post-extension response times immediately following

---

[22] We do not assume that the treatment is as-if randomly assigned. That is, we are not making the exogeneity assumption ($E[\varepsilon|D_i] = 0$) that is common in a selection-on-observables setting.

[23] The term limit extension was catalyzed by Mayor Bloomberg positioning himself as the city's most capable leader in the face of the 2008 financial crisis. Given the uncertain economic climate and falling city revenues, Bloomberg was successful in convincing a majority of council members (and, in the 2009 election, voters) that his financial experience would be necessary in the tough times ahead (Honan 2008). The bill passed, 29-22, just two weeks after Bloomberg had decided that he wanted to run again. The final vote was preceded by 20 hours of public hearings and a full day of floor debate in what was described as a divided City Hall (Chan and Hicks 2008).

Figure 1 Non-parametric Estimates of Response Time by Treatment Status
Average response times increase (decrease) by ∼1 day in control (treated) districts.



The flat lines plot the average response times before and after the term-limit extension for councilors whose election eligibility was (in black) and was not (in gray) affected by the policy change. The dashed lines are linear trends. The figure excludes the day of the reform and one day on either side of the policy change.

the extension (as shown by the dashed lines) likely are due to two major storms. These storms slowed service delivery throughout the city in the days after the extension, although less so in treated districts.[24] Without a well-specified counter-factual, a simple before-after comparison of treated councilors would have understated the effect of the extension. Also note that treated and control districts had almost identical response times before the extension. This is not inconsistent with the idea that elections matter: many fixed characteristics among treated councilors — e.g., their additional experience — may have contributed to relatively quick response times in these districts before the extension even in the absence of a reelection incentive.

Table 2 presents our difference-in-difference estimates ($\hat{\beta}$), confirming that responsiveness in treated districts improved significantly relative to control areas after the term-limit extension. We use different time

[24]The National Weather Service archives (`http://www.erh.noaa.gov/okx/stormtotals.html`) recorded "significant weather events" on October 25 and again on October 28-29. The second storm dumped over two inches of rain on Laguardia Airport and generated wind gusts of 45 miles per hour.

Table 2 Effect of Term Extension on Constituency Services
The average response time fell by more than 1 day in treated districts.

	Dependent variable:			
	Response Time			
Time frame†	2	3	4	12
$\hat{\beta}$‡	-1.216	-1.478	-1.276	-0.645
	(0.733)	(0.665)	(0.553)	(0.848)
	p = 0.097	p = 0.027	p = 0.022	p = 0.447
	$\mathbb{1}$(Response Time $<$ 5 Days)			
Time frame†	2	3	4	12
$\hat{\beta}$‡	0.033	0.017	0.020	0.004
	(0.011)	(0.009)	(0.010)	(0.010)
	p = 0.004	p = 0.060	p = 0.045	p = 0.660
Observations	117,276	171,419	230,067	663,880

†Weeks on either side of the extension used to estimate Eq. 1
‡Difference-in-differences estimator (see Eq. 1)
Standard errors clustered on districts in parentheses

windows on either side of the extension (2-4 weeks); $\hat{\beta}$ is negative and of similar magnitude in all cases.[25] Our estimates are significant at the 5%-level when we use the three or four week time windows. When we restrict attention to the two weeks immediately before and after the policy change, our estimate is slightly less precise ($p = 0.097$).[26]

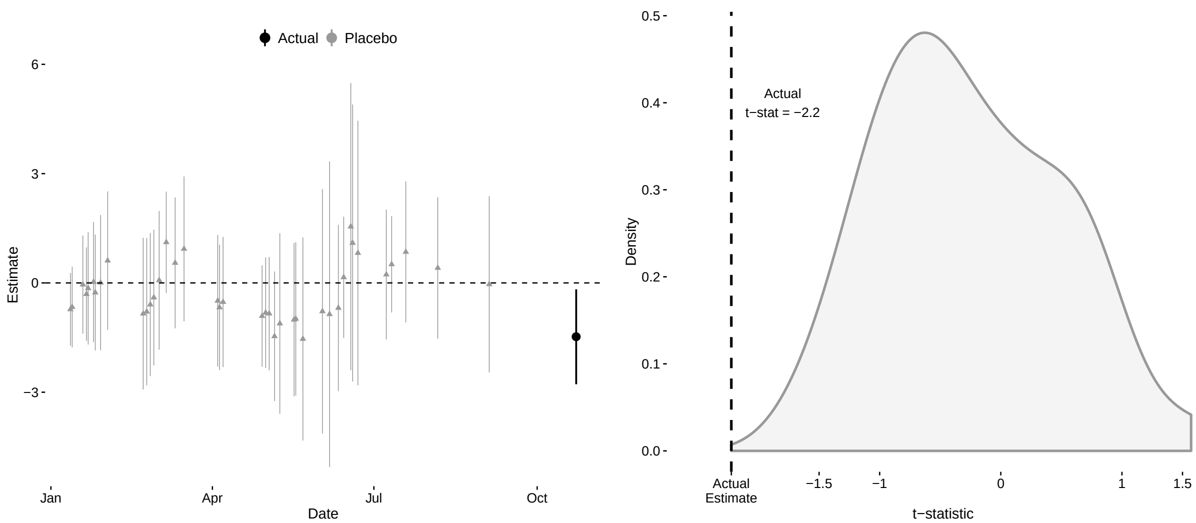
Our results are robust to an alternative modeling strategy. We transform our dependent variable, coding a new binary outcome equal to 1 if a complaint was resolved within five days and 0 otherwise. We then substitute this new outcome variable on the left-hand-side of Equation 1 and estimate linear probability models. The bottom-half of Table 2 includes the results from this specification. The probability that a complaint was resolved within five days increased by two to three percentage points in those districts affected by the term-limit extension, as compared to control areas. This should alleviate the concern that large outliers — requests resolved long after the policy change — are unduly influencing our results.

In the final column of Table 2, we include the results using a 12-week window around the policy change. As expected, the effect attenuates: the coefficient declines by a factor of between two and five. Both treated

[25]Summary statistics for the key variables included in Equation 1 are shown in Table A.3.

[26]Our inferences are unchanged if we employ a version of the block bootstrap, in which we randomly draw 51 districts with replacement to form our bootstrap sample.

Figure 2 Placebo Estimates of Response Time by Treatment Status
Placebo tests indicate that divergent pre-treatment trends do not explain the effect.



(a) Estimates and 95% CIs (b) Distribution of t-statistics

We draw 40 dates at random from January 1, 2008 to September 11, 2008 (six weeks prior to the actual term-limit extension). We then re-estimate Equation 1 using these "placebo" dates to define the treatment event. Displayed on left are the estimates and 95% confidence intervals for $\hat{\beta}$ from Equation 1 using three weeks of data on either side of each date. The estimate from the actual term-limit extension is the right-most, black point. On the right, we show the distribution of t-statistics from these estimates; our actual result is indicated by the dashed line.

and control councilors are eligible for reelection after the policy change, and it appears that their response time trends become more similar as the 2009 election approaches.[27]

We also conduct a series of "placebo" tests to demonstrate that the differences we find cannot be attributed to differential trends prior to the policy change. We randomly draw 40 dates from January 1, 2008 to September 11, 2008 (six weeks prior to the actual term-limit extension) and then estimate Equation 1 using these placebo dates to define the treatment. Figure 2(a) displays these estimates and their 95% confidence intervals. The estimate from the actual term-limit extension is the right-most, black point. As is apparent in Figure 2(b), the t-statistic of our actual estimate is more negative than all of the placebo estimates; our actual result is the only coefficient significant at the 5%-level.[28]

Based on these placebo tests, there is no evidence that changes in responsiveness across groups prior to the term-limit extension account for our findings. This speaks to a final concern: whether treated councilors foresaw the policy change and changed their behavior prior to its passage. We find no evidence of this. We also

[27] We test this expectation more fully in SI 4.

[28] We also use a year before and after the actual policy change as placebo dates (October 23, 2007 and 2009). In both cases the results are positive and cannot be statistically distinguished from 0.

do not find that the effects were larger among treated councilors who voted in favor of the term-limit extension (SI 3).[29]

Do the improvements we detect matter? On average, response times decreased by 4%, or just over a day, in treated districts relative to control districts following the term limit extension. We know of no other studies that employ comparable data, so we cannot rely on the literature to benchmark our effect sizes. However, our estimates can be compared to the effects of events known to severely hamper city services. For example, the January 20-23, 2005 blizzard, which dropped over a foot of snow in NYC, resulted in a 7% increase in response times to service requests opened in the time window of the blizzard, and labor day weekends on average result in an 8% increase in response times. Using these events as a benchmark, our results indicate that the term limit extension induced substantively meaningful improvements in responsiveness in districts held by newly eligible incumbents.

Effects of Election Timing on Constituency Responsiveness

Do elections also affect *when* incumbents exert effort? We use data from two elections to answer this question: the New York City Council Elections of 2005, and the analogous San Francisco Board of Supervisor Elections of 2010.

Empirical Strategy

In each set of elections, we compare incumbents who are seeking reelection (treated) with incumbents who cannot seek reelection due to term limits or staggered elections (control). In NYC, which has term-limits but no staggered elections, 44 of 51 incumbents ran for reelection in 2005. In SF, which has term-limits *and* staggered elections (half of the Board is elected every two years in alternating elections), our treatment group consists of a sole incumbent seeking reelection: Carmen Chu of District 4.[30] We recognize the limitations

[29]Nor do we find clear evidence that councilors in more competitive districts (based on their margin of victory in the 2005 primary) are driving our effects.

[30]All of our models include district (i.e., councilor) fixed effects, so we are not simply reporting a level difference in effort between Chu and other incumbents.

of having a single treated incumbent and interpret these results cautiously.[31] Pooling the data across the two elections only strengthens our results.

We evaluate whether response times fall more precipitously as elections approach in districts where incumbents are eligible to stand for reelection, relative to districts represented by an ineligible councilor. To do this, we estimate the time-trend in response times for both groups, after accounting for level differences across districts, the nature of the complaint, and the day of the week on which a complaint was filed. We then test whether response times are falling faster where incumbents are eligible for reelection (i.e., whether the time-trend is more negative in treated districts). Our empirical model is

$$y_{idt} = \alpha_d + \delta t + \beta(D_d \cdot t) + \gamma_{type} + \phi_{day} + \varepsilon_{idt} \quad (2)$$

where i indexes complaints, d city council district, and t represents days before the general elections. D_d is an indicator equal to 1 for treated city council districts (those with an eligible incumbent). As before, we include fixed effects for districts, request type, and the day of the week on which the complaint was made, and cluster the standard errors at the council district-level. For the SF elections, we use term-limited incumbents and incumbents not up for election as control groups in separate regressions. To analyze legislative efforts, we use the same specification without the fixed effects for complaint type (γ_{type}) or the day of the week (ϕ_{day}).

The parameter of interest again is β.[32] The key identifying assumption is that in the absence of the election, treated and control incumbents would have followed parallel trends in effort levels. A negative estimate indicates improving effort (i.e., more sharply declining response times) among treated incumbents relative to control incumbents.

Results

In Figure 3, we explore whether the timing of elections affects responsiveness to 3-1-1 requests using a non-parametric approach (i.e., we allow for non-linear time-trends). In both NYC and SF, it appears that response times to service requests declined more rapidly in treated districts than in control districts. (Before creating

---

[31] We also note that careful comparisons can yield meaningful inferences even when there is a single treated cluster; see, for example, Abadie, Diamond and Hainmueller (2012). Unfortunately, without pre-treatment data, we cannot employ the synthetic control method developed by Abadie, Diamond and Hainmueller (2012).

[32] There may be level differences in responsiveness across our treatment and control groups. However, this comparison is confounded by differences between councilors — for example, in their experience — so we do not devote attention to the intercepts, α_d.

this plot, we first partial out the variation in response times explained by the complaint type, council district, and day of the week on which the complaint was made.) The SF figure (right) is particularly striking: roughly one year prior to the election, response times fell off sharply in the treated district, while they continued to increase in districts with ineligible incumbents. In NYC (left), response times appear to be declining almost monotonically in treated districts, while in control districts, response times continue to increase through the winter months of 2005. This figure demonstrates that our findings persist, even if we allow for flexibly estimated time-trends across our groups.[33]

In Table 3, we present the results from estimating Equation 2. The results are presented for three separate samples, split by city and type of control group: (1) NYC using term-limited incumbents as control, (2) SF using incumbents not yet up for reelection as control, and (3) SF using term-limited incumbents as control. By splitting the control group in SF, we are able to evaluate whether our findings are comparable for term-limited lame ducks and off-cycle incumbents. The results are also split by the number of days before the election we use to estimate Equation 2, corresponding to 2, 1.5, 1, and 0.5 years.[34]

These results confirm that response times declined more rapidly in districts with an eligible incumbent than in districts without an eligible incumbent. The estimates of β are negative in 11 of 12 models, and can be distinguished from 0 at conventional levels of confidence in 8. Our estimates get more uncertain as we shrink the number of days before the election we include to estimate Equation 2. One likely explanation for this is that the number of observations declines with shorter time frames, decreasing our power to reject the null hypothesis of no effect.

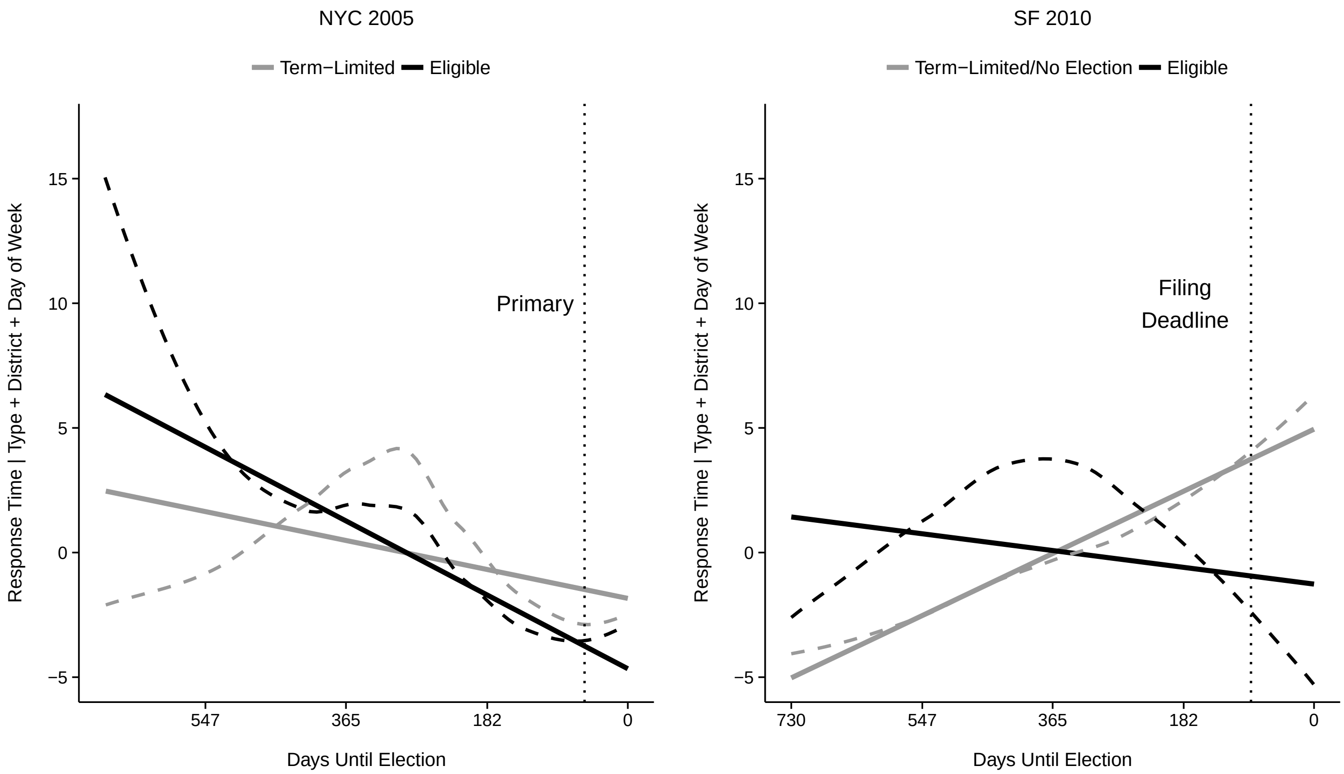
If the differences we discover are driven by reelection incentives, then they should disappear after the election. That is, we should not find differential trends in responsiveness among our treated and control groups in the post-election period. To assess this, we recreate Figure 3 — estimating both flexible and linear time trends in response times — using data from one year *after* the elections in NYC and SF. As Figure 4 illustrates, the trends in our treated and control districts are very similar in the post-election period.[35] This provides further

[33]Interestingly, response times in NYC appear to increase slightly following the primary election. City Council elections in NYC are partisan, and — in all but a few districts — the Democratic nominee has an overwhelming advantage in the general election. The primary election, on the other hand, tends to be competitive. After weathering the primaries, incumbents may therefore be unconcerned that shirking will be punished by partisan voters. Estimating Equation 2 for both the general and primary elections (in separate regressions) leads to substantively similar results.

[34]Our data from NYC go back only to January 1, 2004, so the two-year window corresponds to 677 rather than 730 days.

[35]In SI 7 we perform a series of placebo tests using data from the post-election period. These tests and Figure 4 both suggest that response times do not follow different time trends following the elections.

Figure 3 Non-parametric and Least Squares Estimates of Response Time by Treatment Status
In both NYC and SF, response times fall faster in districts with electorally eligible incumbents.



The straight, solid lines are least squares estimates of response times in the two years preceding the election for councilors that are (in black) and are not (in gray) eligible to stand for reelection. The dashed lines are loess fits with a span of 0.9. The dependent variable for this figure has been "centered" to partial out the variation due to the type of response, district in which the response is made, and day of the week on which the response is lodged. The dashed, vertical black lines indicate, respectively, the date of the primary election in NYC and the filing deadline for candidates in SF. The figure on the left is based on a random sample of 200,000 observations, or just under 10% of the full sample.

evidence that, when election contests are not imminent, response times follow similar trends in these treated and control districts.

To interpret the substantive effect of the estimates, note that they represent the implied effect for *one service request* as we move *one day* closer to the election. Taking the estimates from column 1, the results from NYC suggest that moving six months closer to the election corresponded to a two day reduction in response times in treated districts relative to control. In SF, the results were twice as large. These effects are roughly the same magnitude as (or slightly larger than) the change in response times induced by the January 2005 blizzard in NYC or by labor day weekends, events that substantially affect service delivery.

One aspect of the SF results is worth highlighting. The estimates are more negative when the control group is term-limited incumbents rather than incumbents facing no election in 2010. Our simple decision-theoretic model from above predicts that incumbents facing reelection — even if that contest will not occur for

Table 3 Estimates of Differential Time-Trends in Constituency Services (β in Equation 2)
The linear trend in responsiveness falls significantly faster where incumbents can run for reelection.

	Dependent variable:			
	Response Time			
Time frame†	730	547	365	182
	NYC 2005			
$\hat{\beta}$	−0.012	−0.006	0.005	−0.022
	(0.004)	(0.004)	(0.008)	(0.019)
	$p = 0.003$	$p = 0.097$	$p = 0.533$	$p = 0.249$
Observations	2,376,717	2,238,742	1,578,977	774,375
	SF 2010 (Control: No Election)			
$\hat{\beta}$	−0.014	−0.025	−0.039	−0.002
	(0.004)	(0.006)	(0.009)	(0.022)
	$p = 0.001$	$p < 0.001$	$p < 0.001$	$p = 0.943$
Observations	160,678	120,549	81,801	42,571
	SF 2010 (Control: Term-Limited)			
$\hat{\beta}$	−0.022	−0.039	−0.071	−0.047
	(0.011)	(0.020)	(0.036)	(0.040)
	$p = 0.050$	$p = 0.055$	$p = 0.049$	$p = 0.237$
Observations	135,393	101,458	68,933	35,531

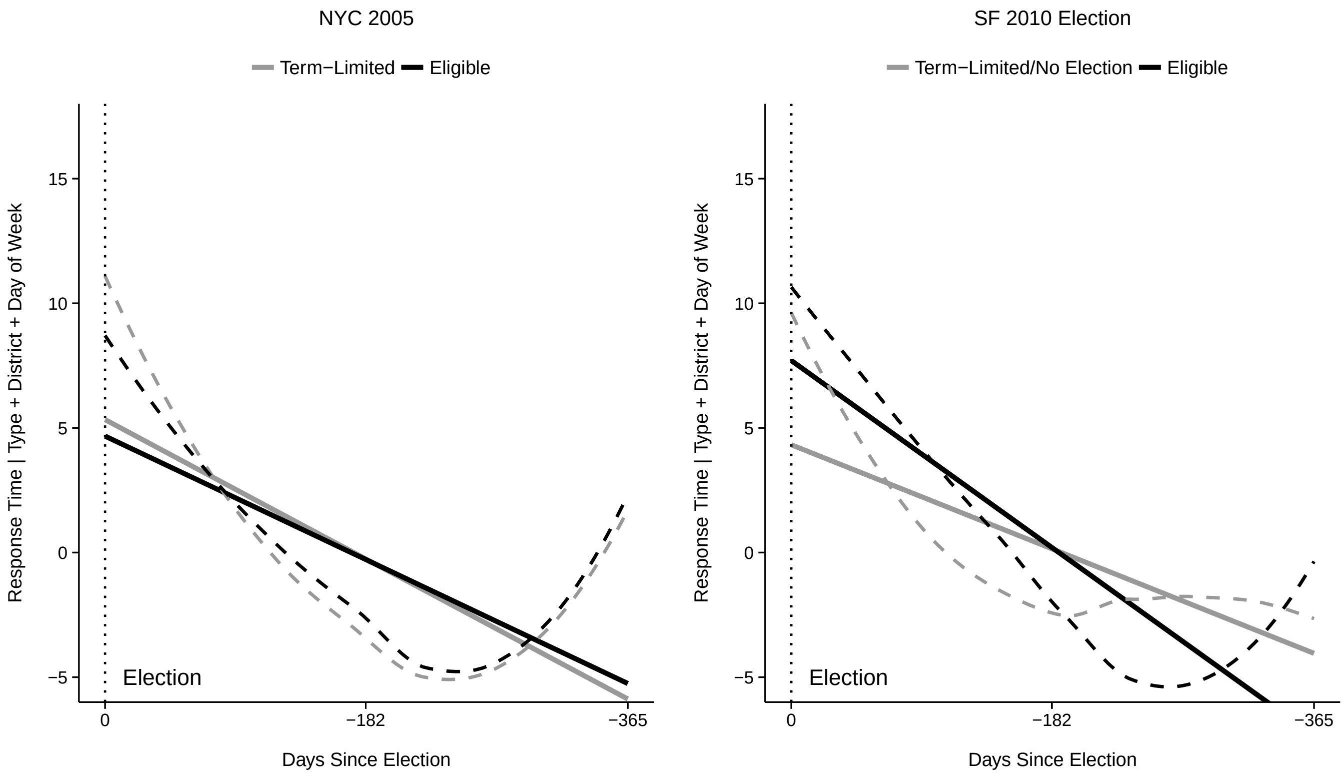
†Days before election used to estimate Eq. 2
Standard errors clustered on districts in parentheses

two more years — should be more concerned about public service responsiveness than term-limited incumbents. Voters may be particularly attuned to the responsiveness of their elected officials during election times, whether or not their supervisor is seeking reelection. Thus, from the perspective of an incumbent seeking reelection in two years, improved performance during this time period may be an opportunity to persuade future supporters at a moment when supervisors' efforts are particularly salient.[36]

Our election timing results are robust to an alternative coding of our dependent variable. In SI 5, we define indicator variables for whether a service request was resolved within five days. We then estimate linear probability models that are otherwise identical to the specifications employed above. We find that the proba-

[36]An alternative interpretation of our results is that bureaucrats in city agencies are less responsive to lame-duck councilors. However, this interpretation cannot explain why our reelection seeking incumbent outperforms her off-cycle colleagues, who are not lame ducks.

Figure 4 Response Times by Treatment Status *After* the Election
After the election, there is no evidence differential trends in responsiveness.



The straight, solid lines are least squares estimates of response times in the year following the election for councilors that are (in black) and are not (in gray) eligible to stand for reelection. The dashed lines are loess fits with a span of 0.9. The dependent variable for this figure has been "centered" to partial out the variation due to the type of response, district in which the response is made, and day of the week on which the response is lodged. The dashed, vertical lines indicate the election dates. The figure on the left is based on a random sample of 200,000 observations.

bility that a service request was resolved within this short time-span increased at a faster rate in districts with eligible incumbents. This analysis and our use of different trimming rules (see SI 6) should ameliorate concerns that our results are driven by outlying response times.

Reallocation from Legislative Activity?

To assess whether improved response times to constituency services come at the expense of legislative action, we also collected data on city councilors' legislative activity. The data for NYC come from the city's Legislative Research Center, which compiles all of the legislation introduced in each city council meeting, including information on which councilors sponsored or co-sponsored the actions.[37] We collect similar data for SF. In

[37] Available at `http://legistar.council.nyc.gov/`. To extract data from this site, we amended scripts from Legistar Scraper, a Python library from Gregg and Poe (2013).

Table 4 Estimates of Changes in Legislative Activity
Eligible incumbents either maintain or increase their legislative effort, suggesting no reallocation.

	Dependent variable:						
	Total Legislative Actions Introduced						
	NYC 2008	NYC 2005		SF 2010 (Control: No Election)		SF 2010 (Control: Term-Limited)	
Time frame	42	365	182	365	182	365	182
$\hat{\beta}$	−0.393	−0.0004	−0.0002	0.006***	0.016***	0.002	0.008***
	(1.108)	(0.003)	(0.006)	(0.0004)	(0.002)	(0.002)	(0.001)
Observations	160	1,224	612	287	168	205	120

†Days before election used to estimate model (excluding γ_{type}); ***p<0.01
Standard errors clustered on districts in parentheses

November 2009, the SF Board of Supervisors started to publish information about which supervisors sponsored particular ordinances, resolutions, and requests for hearings at each Supervisor meeting.[38]

These data sources allow us to generate panel data on legislative activity for every city councilor and supervisor in NYC and SF. For both cities, we code our outcome variable as the number of local laws and resolutions sponsored or co-sponsored by a council member at each meeting. As councilors are better able to control when a bill is introduced than when it is eventually passed, we use the date of the council meeting in which the legislation was introduced and not the date of its eventual passage or dismissal.

We find no evidence that the term limit extension induced changes in legislative efforts. Using Equation 1 but omitting fixed effects for type and the day of the week on which the request was made, the DiD estimates are noisy zeros, never beginning to approach conventional levels of statistical significance. Table 4 (column 2, "NYC 2008") presents one of these estimates, using the total number of local laws and resolutions (co)sponsored in the six weeks before and after the policy change as the outcome variable. These findings suggest that incumbents were not cutting back on legislative effort as they ramped up their work on constituency services.

We also find no evidence that local politicians reallocate their efforts from legislative activity to constituency services in the run-up to elections (Table 4, last six columns). We again use the total number of legislative actions taken by a city councilor in any one meeting as our outcome variable. We have also tried specifications in which we split the outcome variable by type of legislative action, with substantively similar conclusions. We use only two time frames (1 and 0.5 years) when estimating these models due to the availability of these data.

[38]The legislation introduced in each supervisor meeting in 2009 can be found at `http://www.sfbos.org/index.aspx?page=1589`.

The table indicates that election timing does not appear to affect legislative activity in NYC. The standard errors on the interaction term are several times larger than the (substantively small) coefficient estimates. On the other hand, we find evidence that the eligible incumbent's legislative efforts improved leading up to the 2010 elections in SF. These estimates are positive and distinguishable from 0 at the 1% level in 3 of 4 specifications. Non-parametric estimates similar to those in Figure 3 indicate that while the eligible incumbent maintained a relatively constant level of legislative activity in the run-up to her reelection contest, activity levels dropped among both term-limited and off-cycle incumbents. This drop-off in the control group is driving the sign of the estimates. To interpret the results substantively, we would predict eligible incumbents to sponsor roughly two more actions than non-eligible colleagues in the six months prior to elections. Taken together, these results suggest that the motivating effects of elections on constituency services do not come at the expense of legislative action.

Discussion and Conclusion

This paper considers an age-old question in political science about how elections shape the work that representatives do while in office. Analyzing 15 million service requests from NYC and SF, we find that city councilors' electoral incentives are a robust predictor of responsiveness to constituency concerns. Elections encourage overall improvements to constituency responsiveness, consistent with many models of representative democracy. Elections also induce cycles in responsiveness: incumbents ramp up their efforts as elections approach, suggesting that they take into account voter myopia when designing their reelection strategies.

These findings may inform our understanding of representation in U.S. cities. Despite progressive era municipal reforms that reined in the ability of local party machines to carry out municipal services (Anzia 2012; Bernard and Rice 1975; Bridges 1997; Trounstine 2008), scholars have identified many ways in which political mechanisms still shape local service provision. For example, polarized political preferences (Alesina, Baqir and Easterly 1999; Trounstine 2015), the racial identities of politicians and constituents (Hajnal and Trounstine 2005; Schumaker and Getter 1977), and variation in local institutions (Hajnal 2009; Hajnal and Trounstine 2010) all influence how local governments distribute services. We extend this scholarship by showing that local elections also play an important role in shaping within-city allocation of such services. As more cities adopt open data initiatives that store large-scale and multifaceted information, scholars will be able to advance this line of research further.

Our study may also have implications for studies of shirking at the state and federal levels. While these studies on the whole report mixed results from elections, those that have focused on outcomes other than ideological shirking — constituency services, legislative attendance, and oversight — all have found that elections discipline politicians (Table A.1). This suggests that the results we report here may extend to state and federal public offices. However, there are also reasons to believe that our results would be different at the federal level. This is because Congress and the cities in our study have very different electoral institutions — most obviously, Congress does not employ term limits. This difference has implications for selection mechanisms: all else equal, the existence of term limits means that a different type of candidate will select into running for public office (Besley 2006). It may therefore be that Members of Congress, who may have planned on a long career in office, would react more adversely to term limits than the city councilors in our sample.

Lastly, what are the normative implications of our findings? In Madison's ([1788] 1966) conception of representative democracy, elections discipline politicians. However, our results may be indicative of pandering (Canes-Wrone, Herron and Shotts 2001) or selective responsiveness to a vocal subset of constituents (Sances 2016) rather than diligent effort to serve all residents. While fully disentangling this issue would require a systematic classification of how local representatives best should spend their time, this concern may highlight the importance of studying multiple incumbent activities. Although we have only analyzed two, it is reassuring that improved constituency services, a "visible" good that could be used for pandering (Mani and Mukand 2007), do not come at the expense of reduced legislative activity. In SI 8, we have also attempted to test whether the impact of elections is different in neighborhoods with different racial demographics, which could indicate that elections only improve the outcomes of a select few. We find no consistent evidence that our results are moderated by race. We emphasize, however, that this does not mean that race is inconsequential for service provision in NYC and SF, as we only test for heterogeneous — not direct — effects. We leave this important line of inquiry for future research.

References

Abadie, Alberto, Alexis Diamond and Jens Hainmueller. 2012. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* .

Alesina, Alberto, Reza Baqir and William Easterly. 1999. "Public Goods and Ethnic Divisions." *The Quarterly Journal of Economics* 114(4):1243–1284.

Alt, James, Ethan Bueno de Mesquita and Shanna Rose. 2011. "Disentangling Accountability and Competence in Elections: Evidence from US Term Limits." *The Journal of Politics* 73(01):171–186.

Anzia, Sarah F. 2012. "Partisan Power Play: The Origins of Local Election Timing as an American Political Institution." *Studies in American Political Development* 26(01):24–49.

Bails, Dale and Margie A Tieslau. 2000. "The Impact of Fiscal Constitutions on State and Local Expenditures." *Cato J.* 20:255.

Bernard, Richard M and Bradley R Rice. 1975. "Political Environment and the Adoption of Progressive Municipal Reform." *Journal of Urban History* 1(2):149–174.

Besley, Timothy. 2006. *Principled Agents?: The Political Economy of Good Government*. Oxford: Oxford University Press.

Besley, Timothy and Anne Case. 1995. "Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits." *The Quarterly Journal of Economics* 110(3):769–798.

Besley, Timothy and Anne Case. 2003. "Political Institutions and Policy Choices: Evidence from the United States." *Journal of Economic Literature* pp. 7–73.

Bridges, Amy. 1997. "Textbook Municipal Reform." *Urban Affairs Review* 33(1):97–119.

Cain, Bruce E and Thad Kousser. 2004. *Adapting to Term Limits: Recent Experiences and New Directions*. Public Policy Institute of California San Francisco, CA.

Cain, Bruce, John Ferejohn and Morris Fiorina. 1987. *The Personal Vote: Constituency Service and Electoral Independence*. Harvard University Press.

Canes-Wrone, Brandice and Jee-Kwang Park. 2012. "Electoral Business Cycles in OECD Countries." *American Political Science Review* 106(01):103–122.

Canes-Wrone, Brandice, Michael Herron and Kenneth Shotts. 2001. "Leadership and Pandering: A Theory of Executive Policymaking." *American Journal of Political Science* 45(3):532–550.

Carey, John M, Richard G Niemi and Lynda W Powell. 1998. "The Effects of Term Limits on State Legislatures." *Legislative Studies Quarterly* pp. 271–300.

Carey, John M, Richard G Niemi, Lynda W Powell and Gary F Moncrief. 2006. "The effects of term limits on state legislatures: a new survey of the 50 states." *Legislative Studies Quarterly* 31(1):105–134.

Carson, Jamie L, Michael H Crespin, Jeffery A Jenkins and Ryan J Vander Wielen. 2004. "Shirking in the Contemporary Congress: A Reappraisal." *Political Analysis* 12(2):176–179.

Chan, Sewell and Jonathan P. Hicks. 2008. "Council Votes, 29 to 22, to Extend Term Limits.". *The New York Times*. Available at `http://cityroom.blogs.nytimes.com/2008/10/23/council-to-debate-term-limits-change/?_r=0`.

Clark, Jennifer Hayes and R Lucas Williams. 2013. "Parties, Term Limits, and Representation in the US States." *American Politics Research* pp. 1–23.

Crain, W. Mark and Lisa K. Oakley. 1995. "The Politics of Infrastructure." *Journal of Law and Economics* 38:1–17.

Crain, W. Mark and Robert D. Tollison. 1993. "Time inconsistency and fiscal policy: Empirical analysis of US states, 1969–89." *Journal of Public Economics* 51(2):153–159.

Cummins, Jeff. 2012. "The Effects of Legislative Term Limits on State Fiscal Conditions." *American Politics Research* pp. 1–26.

Dewan, Torun and Kenneth A. Shepsle. 2011. "Political Economy Models of Elections." *Annual Review of Political Science* 14:311–331.

Erler, H Abbie. 2007. "Legislative Term Limits and State Spending." *Public Choice* 133(3-4):479–494.

Figlio, David N. 1995. "The Effect of Retirement on Political Shirking: Evidence from Congressional Voting." *Public Finance Review* 23(2):226–241.

Fiorina, Morris P. 1989. *Congress, keystone of the Washington establishment*. Yale University Press.

Franzese, Robert J. 2002. "Electoral and Partisan Cycles in Economic Policies and Outcomes." *Annual Review of Political Science* 5(1):369–421.

Gregg, Forest and Mujumbe Poe. 2013. "legistar-scrape.".
URL: *https://github.com/fgregg/legistar-scrape*

Grier, Kevin. 2008. "US Presidential Elections and Real GDP Growth, 1961–2004." *Public Choice* 135(3-4):337–352.

Grossman, Guy et al. 2014. "Do Selection Rules Affect Leader Responsiveness? Evidence from Rural Uganda." *Quarterly Journal of Political Science* 9(1):1–44.

Hajnal, Zoltan and Jessica Trounstine. 2005. "Where Turnout Matters: The Consequences of Uneven Turnout in City Politics." *Journal of Politics* 67(2):515–535.

Hajnal, Zoltan L. 2009. *America's Uneven Democracy: Race, Turnout, and Representation in City Politics*. Cambridge University Press.

Hajnal, Zoltan L and Jessica Trounstine. 2010. "Who or What Governs?: The Effects of Economics, Politics, Institutions, and Needs on Local Spending." *American Politics Research* 38(6):1130–1163.

Honan, Edith. 2008. "NY Council Extends Term Limit so Bloomberg Can Run.". *Reuters.* Available at `http://www.reuters.com/article/2008/10/23/us-newyork-bloomberg-idUSTRE49M70J20081023`.

Huber, Gregory A and Sanford C Gordon. 2004. "Accountability and Coercion: Is Justice Blind When It Runs for Office?" *American Journal of Political Science* 48(2):247–263.

Huber, Gregory A., Seth J. Hill and Gabriel S. Lenz. 2012. "Sources of Bias in Retrospective Decision Making: Experimental Evidence on Voters' Limitations in Controlling Incumbents." *American Political Science Review* 106(04):720–741.

Keele, Luke, Neil Malhotra and Colin H McCubbins. 2013. "Do Term Limits Restrain State Fiscal Policy? Approaches for Causal Inference in Assessing the Effects of Legislative Institutions." *Legislative Studies Quarterly* 38(3):291–326.

King, Gary. 1991. "Constituency Service and Incumbency Advantage." *British Journal of Political Science* 21(01):119–128.

Krause, George A. 2005. "Electoral Incentives, Political Business Cycles and Macroeconomic Performance: Empirical Evidence from Post-War US Personal Income Growth." *British Journal of Political Science* 35(01):77–101.

Lenz, Gabriel S. and Andrew Healy. 2014. "Substituting the End for the Whole: Why Voters Respond Primarily to the Election-Year Economy." *American Journal of Political Science* 58(1):31–47.

Levitt, Steven D. 1997. "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime." *The American Economic Review* 87(3):270–290.

Lewis, Daniel C. 2012. "Legislative Term Limits and Fiscal Policy Performance." *Legislative Studies Quarterly* 37(3):305–328.

Lott, John. 1990. "Attendance Rates, Political Shirking, and the Effect of Post-Elective Office Employment." *Economic Inquiry* 28(1):133–150.

Lott, John R. 1987. "Political Cheating." *Public Choice* 52(2):169–186.

Lott, John R. and Stephen G. Bronars. 1993. "Time Series Evidence on Shirking in the U.S. House of Representatives." *Public Choice* 76:125–49.

Madison, James. 1966. *Federalist No. 57.* Garden City, NY: Anchor Books.

Mani, Anandi and Sharun Mukand. 2007. "Democracy, Visibility and Public Good Provision." *Journal of Development Economics* 83(2):506–529.

Mayhew, David R. 1987. The Electoral Connection and the Congress. In *Congress: Structure and Policy*, ed. Matthew D McCubbins and Terry Sullivan. New York: Cambridge University Press Archive.

Nordhaus, William D. 1975. "The Political Business Cycle." *The Review of Economic Studies* pp. 169–190.

Poole, Keith T and Thomas Romer. 1993. Ideology, "Shirking", and representation. In *The Next Twenty-five Years of Public Choice*. Springer pp. 185–196.

Przeworski, Adam, Susan Stokes and Bernard Manin. 1999. *Democracy, accountability, and representation*. Cambridge: Cambridge University Press.

Rogoff, Kenneth S. 1990. "Equilibrium Political Budget Cycles." *The American Economic Review* pp. 21–36.

Rothenberg, Lawrence S. and Mitchell S. Sanders. 2000. "Severing the Electoral Connection: Shirking in the Contemporary Congress." *American Journal of Political Science* 44(2):316–325.

Sances, Michael W. 2016. "The Distributional Impact of Greater Responsiveness: Evidence from New York Towns." *The Journal of Politics* 78(1):105–119.

San Francisco Department of Public Works, FY 2010-11 Proposed Budget. 2010.

Schultz, Kenneth A. 1995. "The Politics of the Political Business Cycle." *British Journal of Political Science* 25(01):79–99.

Schumaker, Paul D and Russell W Getter. 1977. "Responsiveness Bias in 51 American Communities." *American Journal of Political Science* pp. 247–281.

Schumpeter, Joseph. 1939. *Business Cycles: A Theoretical, Historical, and Statistical Analysis*. New York: McGraw Hill.

Shepsle, Kenneth A., Robert P. Van Houweling, Samuel J. Abrams and Peter C. Hanson. 2009. "The Senate Electoral Cycle and Bicameral Appropriations Politics." *American Journal of Political Science* 53(2):343–359.

Snyder, James M and Michael M Ting. 2003. "Roll Calls, Party Labels, and Elections." *Political Analysis* 11(4):419–444.

Tien, Charles. 2001. "Representation, Voluntary Retirement, and Shirking in the Last Term." *Public Choice* 106(1-2):117–130.

Trounstine, Jessica. 2008. *Political Monopolies in American Cities: The Rise and Fall of Bosses and Reformers*. University of Chicago Press.

Trounstine, Jessica. 2015. "Segregation and Inequality in Public Goods." *American Journal of Political Science* .

Tufte, Edward R. 1978. *Political Control of the Economy*. Princeton: Princeton University Press.

Vanbeek, James R. 1991. "Does the Decision to Retire Increase the Amount of Political Shirking?" *Public Finance Review* 19(4):444–456.

Wright, Gerald C. 2007. "Do Term Limits Affect Legislative Roll Call Voting? Representation, Polarization, and Participation." *State Politics & Policy Quarterly* 7(3):256–280.

Appendix 1 Summary of Existing Literature on Last-Term Shirking

We conducted a survey of work on last-term shirking in the United States by collecting articles from 1990 and onward from Google Scholar. We started with results returned from key word searches (e.g., "shirking term limits" and "shirking retirement congress"), choosing articles that clearly studied the effect of term limits or retirement on an outcome that captures incumbent effort (or, conversely, shirking) in some way. We followed the current literature in defining incumbent effort broadly, e.g., including fiscal outcomes that may not be directly attributable to politicians. We identified additional studies by following up on citations in the articles that were initially returned in our search. This procedure resulted in a total of 26 articles on the topic.

Table A.1 summarizes these articles. As can be seen, they study governors, state legislators, and members of congress. Studies of governors and state legislators analyze the effect of term limits on shirking, while studies of members of congress analyze the effect of retirement. The studies in the table also differ with respect to their dependent variable and, hence, the type of shirking they consider. Most studies analyze ideological shirking — that is, whether the voting record or fiscal policies put in place by last-term incumbents differ from their previous records or their constituents' preferences. Six studies analyze legislative attendance rates, two studies consider casework/constituency services, and one study looks at agency oversight.

Table A.1 Recent empirical investigations of last-term shirking in the United States

Paper	Sample	Dependent variable(s)	Evidence of shirking?			
			Ideo.	Attend.	Casework	Oversight
Term limits						
Alt et al. (2011)	Governors	Fiscal variables	✓			
Bails and Tieslau (2000)	State legislatures	Fiscal variables	×			
Besley (2006)	Governors	Fiscal variables; congruence	×			
Besley and Case (1995)	Governors	Fiscal variables	✓			
Besley and Case (2003)	Governors	Fiscal variables	✓			
Cain and Kousser (2004)	California	Vote deviation; oversight	×			✓
Carey et. al. (1998)	State legislatures	Legislation; casework	×		✓	
Carey et al. (2006)	State legislatures	Legislation; casework	×		✓	
Clark and Williams (2013)	State legislatures	Vote deviation; attendance	✓†	✓		
Crain and Oakley (1995)	Governors	Capital investments	✓			
Crain and Tollison (1993)	Governors	Fiscal volatility	✓			
Cummins (2012)	State legislatures	Budget balance	✓			
Erler (2007)	State legislatures	Fiscal variables	×			
Keele et al. (2013)	State legislatures	Fiscal variables	×			
Lewis (2012)	State legislatures	Fiscal variables	✓			
Wright (2007)	State legislatures	Congruence; attendance	×	✓		
Retirement						
Carson et al. (2004)	Congress	Vote deviation	×			
Figlio (1995)	Congress	Vote deviation; attendance	✓	✓		
Lott (1987)	Congress	Vote deviation; attendance	×	✓		
Lott (1990)	Congress	Attendance		✓†		
Lott and Bronars (1993)	Congress	Vote deviation	×			
Poole and Romer (1993)	Congress	Vote deviation	×			
Rothenberg and Sanders (2000)	Congress	Vote deviation; attendance	✓	✓		
Snyder and Ting (2003)	Congress	Vote deviation	✓†			
Tien (2001)	Congress	Congruence	✓			
Vanbeek (1991)	Congress	Vote deviation	×			

Notes: ✓ = results in study can be interpreted as evidence of shirking, X = no evidence of shirking, † = conclusion applies only to a subset of states or legislators. The four types of shirking are with respect to vote content (ideology), legislative attendance rates, constituency services (casework), and agency oversight. Cells are left blank if a study did not consider a given type of shirking. “Fiscal variables” include per capita state government expenditure and taxation (and sometimes borrowing costs and economic growth).

Appendix 2 Information about Service Protocol and Time Line

Table A.2 provides an overview of how 3-1-1 service requests are handled from when they are opened until they are closed. A request has been *opened* when all the intake information about the request has been logged and it has been assigned to an agency. The agency will *close* the request when the issue has been resolved (which may require rerouting it to a different agency) or when it has determined that no action is necessary. In all cases, agency workers will physically check on the issue to determine what type of work is needed. In the paper, we code our dependent variable as the number of days it took from the time a request was opened until it was closed.

Table A.2 Service Request Protocol and Timeline

1. A request is called in to a 3-1-1 response center or submitted online.
2. A 3-1-1 representative (or online system, if the request was submitted online) will ask for and determine the type of request, and an actioning agency will be assigned. Based on a pre-assigned workflow, the service representative will then do the intake of required information (as requested by the assigned agency) and submit the request. The request has been *opened*, and the date is recorded in the 3-1-1 database.
3. The request appears in the acting agency's queue for action, and will have a service level agreement (SLA) specifying a due date based on the request type. The agency may prioritize among different requests as they see fit.
4. Agency staff physically inspect the reported issue, with one of three potential outcomes, each of which results in the request being *closed*:
   a. Agency staff resolve the issue (may require revisits). Some issues are easily verified as resolved (e.g., graffiti removed), while others may require following up with the constituent (e.g., calling the next day to verify that heating works). Once the agency has ensured that the issue is resolved, the agency will report it as closed.
   b. The issue is rerouted to a different agency. This may happen if the agency determines the issue falls outside its jurisdiction. For example, NYC Department of Transportation may reroute a highway issue to NY State, or a misreported issue may be rerouted to a different city agency. In this case the issue will be marked as closed once it has been resolved by the new agency.
   c. Agency staff determine that no work is warranted (e.g., trash was already picked up, prank call) and mark it as closed.

Appendix 3 Summary Statistics

Table A.3 Summary Statistics for NYC Service Request Data, 10/23/2007 - 10/23/2009

Statistic	N	Mean	St. Dev.	Min	Max
y: Response Time (Days)	3,420,140	17.48	56.57	0	1,043
t: $\mathbb{1}$(Post-Ext.)	3,420,140	0.50	0.50	0	1
D: $\mathbb{1}$(Compliers)	2,742,258	0.73	0.44	0	1

Note: Trimmed Top 0.1% of Response Times

Table A.4 Summary Statistics for NYC, 1/1/2004 - 11/08/2005

Statistic	N	Mean	St. Dev.	Min	Max
y: Response Time (Days)	2,378,172	22.65	60.10	0	785
t: Days Before Election	2,378,172	281.31	168.30	0	677
D: $\mathbb{1}$(Treated)	2,376,717	0.87	0.34	0	1

Note: Trimmed Top 1% of Response Times

Table A.5 Summary Statistics for SF, 11/02/2008 - 11/02/2010

Statistic	N	Mean	St. Dev.	Min	Max
y: Response Time (Days)	612,338	24.08	57.86	0.00	524.97
t: Days Before Election	323,105	414.52	244.81	0	853
D: $\mathbb{1}$(Treated)	612,338	0.05	0.21	0	1

Note': Trimmed Top 1% of Response Times

Supporting Information:
Do Elections Improve Constituency Responsiveness? Evidence from U.S. Cities

To be published online.

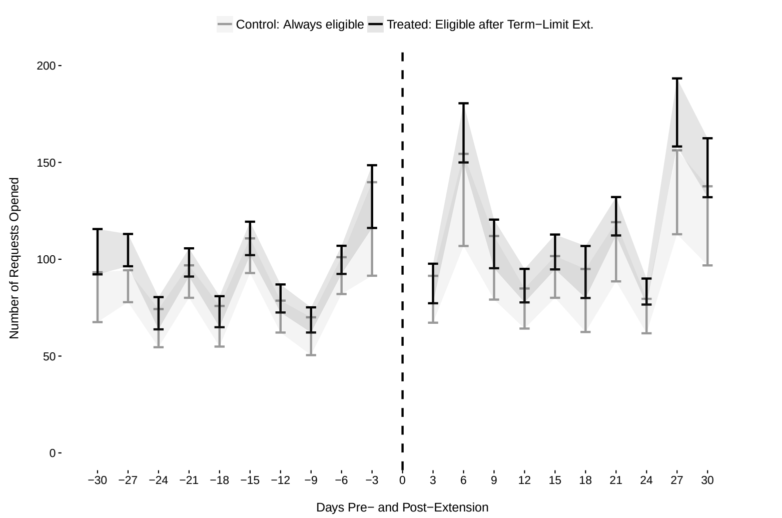
SI 1 Changes in Request Volume Following Term Extension (NYC 2008)

We explore whether changes in response times are driven by changes in the demand for constituency services. Concretely, we estimate whether the number of requests received by treated councilors before and after the term extension is different from the number of requests received by councilors in the control group. Figure SI.1 and Table SI.1 indicate that this is not the case. The number of requests received by treated councilors may have increased by 3 to 8 requests per day per councilor, a small increase (between six and fourteen percent of the inter-quartile range) that cannot usually be distinguished from 0 at traditional levels of confidence.

Table SI.1 Effect of Term Extension of Request Volume

	Dependent variable:		
	Number of Requests (District-Day)		
	2	3	4
$\hat{\beta}$	3.424	4.451	7.611
	(2.881)	(2.941)	(4.284)
	p = 0.235	p = 0.131	p = 0.076
Avg. IQR	56.77	51.71	52.46
Observations	1,160	1,720	2,280
Adjusted R^2	0.438	0.476	0.503
Note:	Standard errors clustered on districts in parentheses		

Figure SI.1 Average Number of Requests around the Term Extension
No differential change in request volume around the term extension.



We total the number of requests in every district-day and then compute the average volume for treatment and control districts. The figure displays the 95% confidence interval around those averages. We group days into three-day intervals to avoid over-plotting.

SI 2 Robustness to Including Incumbents Not Seeking Third Term (NYC 2008)

The treatment group for our main analysis consists of incumbents who were termed out before the October 23, 2008 decision but ran for a third term after the decision. Not all of the 37 newly eligible councilors took advantage of the extension; eight incumbents left politics or ran for different positions (e.g., four ran for comptroller). In Table SI.2, we show that our results hold but decline in magnitude when we include all term-limited incumbents in our treatment group. The results attenuate as expected: we would not expect councilors intending to leave office to exert more effort following the term limit extension.

Table SI.2 Effect of Term Extension on Constituency Services, All Term-Limited Incumbents
The average response time fell by ∼1 day in treated districts.

	Dependent variables:			
	Response Time			
Time frame†	2	3	4	12
$\hat{\beta}$‡	−0.998	−1.210	−1.093	−0.429
	(0.688)	(0.626)	(0.491)	(0.803)
	p = 0.148	p = 0.054	p = 0.027	p = 0.593
	$\mathbb{1}$(Response Time < 5 Days)			
Time frame†	2	3	4	12
$\hat{\beta}$‡	0.026	0.013	0.016	0.001
	(0.010)	(0.008)	(0.009)	(0.009)
	p = 0.011	p = 0.121	p = 0.082	p = 0.931
Observations	145,229	211,947	284,927	822,431

† Weeks on either side of the extension used to estimate Eq. 1
‡ Difference-in-differences estimator (see Eq. 1)
Standard errors clustered on districts in parentheses

SI 3 Heterogeneous Effects by Vote on the Extension (NYC 2008)

Table SI.3 indicates councilors' votes on the term extension based on their treatment status and decision to run for a third-term in office.

Table SI.3 Voting for Term Extension by Treatment Status

Group	Freq.	Prop. Voting Yes
Control	12	0.33
Second-Term, Contesting	28	0.79
Second-Term, Retiring	11	0.27

In Table SI.4, we estimate whether treated incumbents that supported the term extension on October 23, 2008 saw larger reductions in response time. This analysis employs a four-week window around the policy change and does not indicate heterogeneous effects based on councilors' support for the reform.

Table SI.4

	Dependent variable:	
	Response Time	
	Contesting Incumbents	All Incumbents
D_{dt}	-1.654	-1.121
	(0.945)	(0.630)
	p = 0.081	p = 0.075
$D_{dt} \times \mathbb{1}$(Voted Yes)	0.460	0.041
	(1.004)	(0.682)
	p = 0.647	p = 0.953
Observations	230,067	284,927
Adjusted R^2	0.170	0.175
Note:	Standard errors clustered on districts in parentheses	

SI 4 Post-Extension Trends as 2009 NYC Election Approaches

In Table SI.5, we explore post-extension trends in response times leading up to the November 3, 2009 election in NYC. Since all councilors were eligible to run for reelection after the term limit extension, we do not expect different response time *trends* in districts represented by previously term-limited councilors (treated) and districts represented by councilors who were always eligible for reelection (control).

We test this expectation using Equation 2. We use two time periods: a year and six months before the election. If there were no differential changes to response times in treated and control districts, then $\hat{\beta}$ from Equation 2 should be indistinguishable from 0. Table SI.5 provides strong evidence for this: the estimates of differential time trends are very close to 0 with p-values between 0.67 and 0.78.

We emphasize that this analysis is different from our tests of the impact of the term-limit extension, for which we use Equation 1. The extension analysis shows that response times dropped in districts represented by previously term-limited councilors in the weeks after the extension. However, as we also show, this effect attenuates with time. This makes sense given the analysis presented here: as the 2009 election approaches, response time trends in treated and control districts become indistinguishable.

Table SI.5 Tests for Differential Trends before the 2009 NYC Election

	Dependent variable:	
	Response Time	
Days until election:	365	182
$\hat{\beta}$	0.0001	−0.0001
	(0.0001)	(0.0004)
	p = 0.666	p = 0.778
Observations	1,296,798	582,707

Note: Standard errors clustered on districts in parentheses

SI 5 Linear Probability Models: NYC 2005 and SF 2010

We code an indicator variable for whether a public service request was resolved within five days of being opened. We then estimate linear probability models with these as our dependent variables. This recoding flips the interpretation of the coefficients: a positive coefficient now implies that responsiveness is improving and that a request is more likely to be resolved within several days after opening. The linear probability models for NYC 2005 and SF 2010 are included below; the linear probability models for NYC 2008 are included in Table 2 in the body of the paper.

Table SI.6 Estimates of Differential Time-Trends in Constituency Services (β in Equation 2)
DV: $\mathbb{1}$(3-1-1 Response Time $<$ 5 Days) $\times$ 100

Time frame†	730	547	365	182
	NYC 2005			
$\hat{\beta}$	0.0059	0.0047	0.0089	0.0066
	(0.0020)	(0.0020)	(0.0077)	(0.0090)
	$p = 0.004$	$p = 0.019$	$p = 0.247$	$p = 0.464$
Observations	2,376,717	2,238,742	1,578,977	774,375
	SF 2010 (Control: No Election)			
$\hat{\beta}$	0.0037	0.0143	0.0042	−0.0228
	(0.0018)	(0.0040)	(0.0043)	(0.0049)
	$p = 0.001$	$p < 0.001$	$p < 0.001$	$p = 0.943$
Observations	160,678	120,549	81,801	42,571
	SF 2010 (Control: Term-Limited)			
$\hat{\beta}$	0.0083	0.0245	0.0198	0.0032
	(0.0022)	(0.0116)	(0.0126)	(0.0206)
	$p < 0.001$	$p = 0.035$	$p = 0.117$	$p = 0.878$
Observations	135,393	101,458	68,933	35,531

† Days before election used to estimate Eq. 2
Standard errors clustered on districts in parentheses

SI 6 Robustness to Trimming

The tables below demonstrate the robustness of our results to different decisions about how to trim the dependent variable, which has a long right tail.

Table SI.7 Robustness of Election Eligibility Results to Trimming Decisions

	Dependent variable:		
		NYC 2008	
Quantile Trimmed	0.98	0.99	1
$\hat{\beta}$	−0.00860	−0.01219	−0.00884
	(0.510)	(0.551)	(0.572)
	$p = 0.086$	$p = 0.025$	$p = 0.038$
Observations	235,495	236,295	236,302
Note:	Std. errors clustered on districts.		

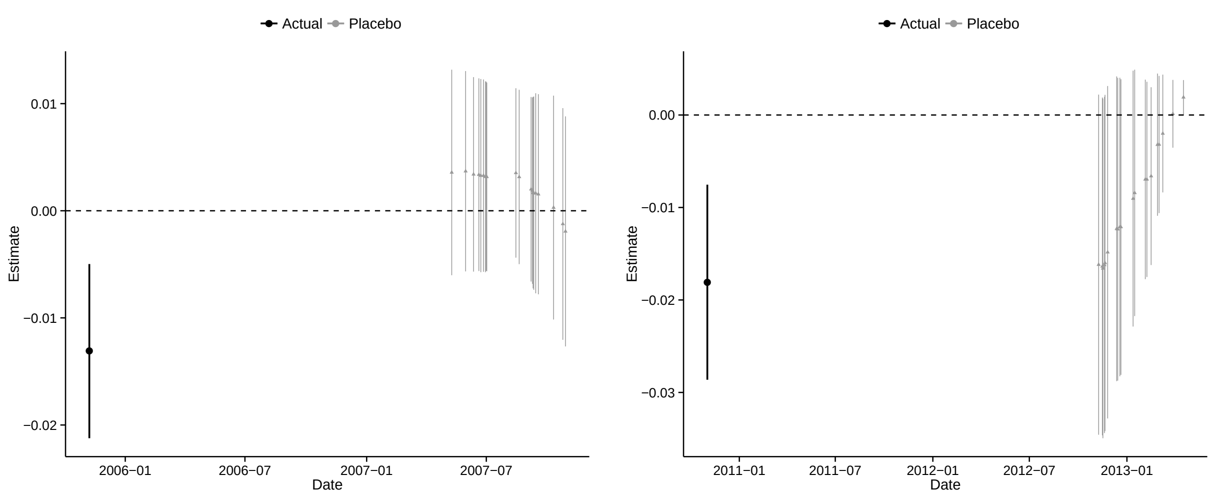
Table SI.8 Robustness of Election Timing Results to Trimming Decisions

	Dependent variable: Response Time (Days)		
		NYC 2005	
Quantile Trimmed	0.98	0.99	1
$\hat{\beta}$	−0.00860	−0.01219	−0.00884
	(0.00311)	(0.00410)	(0.01162)
	$p = 0.00575$	$p = 0.00293$	$p = 0.44678$
Observations	2,352,693	2,376,705	2,400,697
		SF 2010 (Control: No Election)	
Quantile Trimmed	0.98	0.99	1
$\hat{\beta}$	−0.00603	−0.00911	−0.01379
	(0.00211)	(0.00283)	(0.00415)
	$p = 0.00421$	$p = 0.00130$	$p = 0.00089$
Observations	157,464	159,071	160,678
		SF 2010 (Control: No Incumbent)	
Quantile Trimmed	0.98	0.99	1
$\hat{\beta}$	−0.01404	−0.01796	−0.02172
	(0.00950)	(0.00929)	(0.01107)
	$p = 0.13932$	$p =$0. 05319	$p = 0.04967$
Observations	132,685	134,039	135,393
Note:	Std. errors clustered on districts.		

SI 7 Placebo Tests: NYC 2005 and SF 2010

We randomly draw 20 dates from the period following the 2005 election in NYC and the 2010 election in SF. We then re-estimate β from Equation 2. We use 730 days to estimate these placebo regressions. Thus, our "placebo" election dates have to be drawn from an interval of 200 days that is at least two years after the actual election date. This explains why the placebo dates fall well to the right of the estimate associated with the actual election dates (the left-most, black points). To save space, we pool our two types of control districts in SF. These tests suggest that the differential time trends that we discover in our analysis do not persist after the election.

Figure SI.2 Placebo Results for 2005 NYC and SF 2010 Elections



(a) NYC 2005 (b) SF 2010

Displayed above are the estimates and 95% confidence intervals for $\hat{\beta}$ from Equation 2 using two years of data before each placebo date. The estimate from the actual term-limit extension is the left-most, black point.

SI 8 DiD Estimates by Racial Composition

Our findings suggest that public service responsiveness improves when public officials are eligible to seek re-election and as elections approach. We are also interested in whether some neighborhoods benefit more from these improvements than others. In particular, we are interested in whether the racial composition of neighborhoods is responsible for heterogeneous treatment effects in responsiveness. Public officials may, for example, favor coethnic constituents or constituents from a particular racial group.

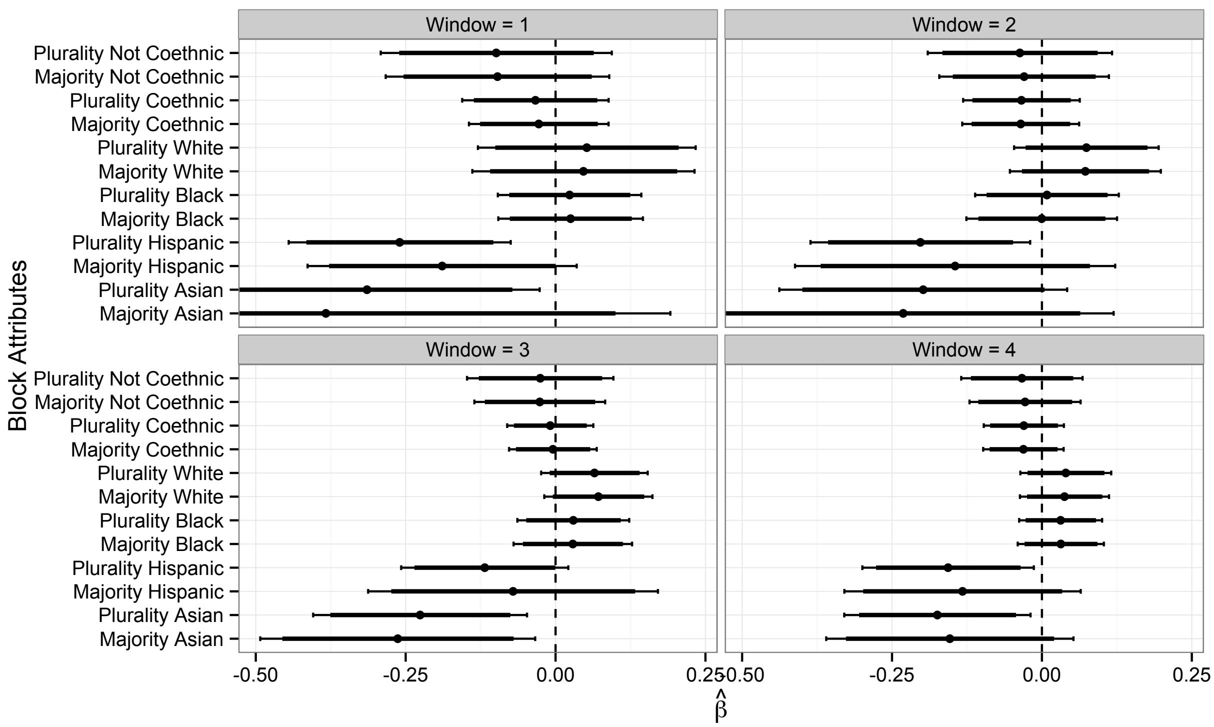
To carry out this analysis we begin by matching the 3-1-1 database with census data on the racial composition of every block in NYC. We then code two characteristics that indicate each block's relationship with the city council member representing the electoral district in which the block is located: whether or not its largest group is coethnic with the city council member; and whether or not its majority group (if any) is coethnic with the city council member. We also create variables for each block's plurality group and (if applicable) majority group without regard for its relationship with the city council member. We then re-run our analyses on different subsets of the data based on these variables.

The results are displayed in Figures SI.3 and SI.4. They suggest that the heightened responsiveness that followed the term-limit extension in 2008 is not driven by ethnic favoritism: neighborhoods in which the plurality (or majority) group is coethnic with the city council member representing the district do not see larger drops in response times than other neighborhoods. We find some evidence, however, that neighborhoods populated primarily by Hispanics or Asians see larger drops in response times, though this may not reflect the ethnicity of these neighborhoods but rather some unobserved characteristic of the neighborhoods that we are not controlling for here (e.g., location).

Moving to the 2005 council elections, response times dropped quite uniformly across neighborhoods two years to a year and a half before the elections. In general, there are few interesting heterogeneous treatment effects to report, perhaps with the exception that Asian neighborhoods saw larger drops closer to the election.

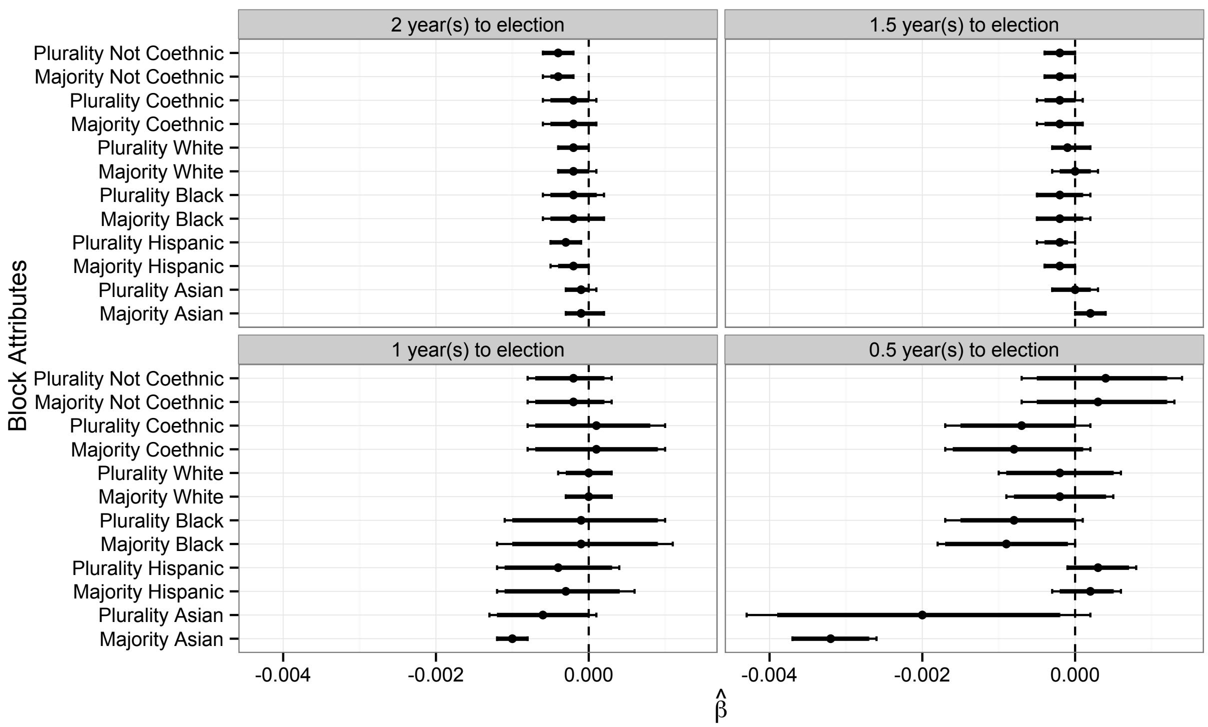
Our failure to uncover consistent heterogeneous effects related to neighborhoods' ethnic composition should not be taken to imply that there are no ethnic disparities in public service delivery. Our empirical strategy leverages changes in responsiveness and, thus, does not address level differences in service delivery across neighborhoods of varying composition.

Figure SI.3 The Term Limit Extension of 2008 and Heterogeneous Effects by Neighborhood Race



We run Equation 1 in subsets of neighborhoods defined by the attributes on the y-axis. "Window" gives the number of weeks on either side of the extension we use. The estimates of β (with 95% CIs) from these regressions are displayed above.

Figure SI.4 The 2005 NYC Election and Heterogeneous Effects by Neighborhood Race



We run Equation 2 in subsets of neighborhoods defined by the attributes on the y-axis, for different time windows before the election. The estimates of β (with 95% CIs) from these regressions are displayed above.