The Politics of Potholes: Service Quality and Retrospective Voting in Local Elections

Craig M. Burnett, Hofstra University
Vladimir Kogan, Ohio State University

By conditioning their support for political incumbents on observed performance outcomes, voters can motivate elected officials to represent their interests faithfully while in office. Whether elections serve this function in subnational US government remains unclear, however, because much of the existing research on retrospective voting in these contexts focuses on outcomes that are not obviously salient to voters or over which the relevant government officials have limited influence. In this study, we examine one outcome—the quality of local roads—that is both salient and unquestionably under the control of city government. Our analysis leverages within-city variation in the number of pothole complaints in one of America’s largest cities and shows that such variation can explain neighborhood-level differences in support for incumbents in two political offices—mayor and city council—across several electoral cycles.

Representative democracy works best when voters can use the ballot to motivate policy makers to represent their interests faithfully while in office. Given evidence that voters remain woefully ignorant about the operation of American political institutions and the policy positions of competing candidates on Election Day (e.g., Campbell et al. 1960; Delli-Carpini and Keeter 1996), whether ordinary citizens possess the ability to do so remains a contested question at the center of the study of political behavior.

Retrospective voting provides one potential mechanism to resolve this democratic dilemma, allowing for effective representation with minimal voter effort. By conditioning their support for political incumbents on one piece of easily accessible information—the observed government performance—voters can use elections to create the political equivalent of Darwinian “natural selection,” ensuring that only faithful agents retain their offices over the long haul (e.g., Barro 1973; Ferejohn 1986). An extended literature examines whether voters behave in this way in national elections, providing ample support for the retrospective voting hypothesis (e.g., Lenz 2012).

But a great deal of policy making takes place outside of Washington, DC, as many political issues remain under the purview of state and local governments. Only recently, however, have scholars shifted their attention to examine how these dynamics play out in subnational elections. To date, the research has produced mixed findings. While there is evidence that voters’ subjective satisfaction with government performance influences their support for elected incumbents (Oliver and Ha 2007), it is far from clear whether these subjective evaluations accurately reflect the objective reality on the ground (Healy and Malhotra 2013). For example, while economic conditions appear to affect both gubernatorial approval and elections, voters do not always correctly apportion responsibility for economic performance between different levels (Brown 2010) and branches (Rudolph 2003a) of government. Moreover, national political considerations—such as approval of the incumbent president—appear to matter much more for state elections than state-level outcomes, such as homicide rates, SAT scores, and the burdens of state taxation (Rogers 2013).

We suspect that the mixed nature of these findings reflect two serious limitations in the existing literature. First, data availability rather than clear theory has generally driven the choice of performance outcomes examined in published analyses. In their study of mayoral approval in New York, for example, Arnold and Carnes (2012) rely on the number of city employees as a proxy for municipal service quality. When such measures yield null effects—as was the case in their...
study—it is impossible to know whether one should interpret the results as evidence against retrospective voting or simply as a failure to measure the dimensions of government performance that voters find most salient.

On the other hand, studies that have found evidence of retrospection in subnational voter behavior largely focus on outcomes over which these governments have limited, if any, direct control. When constituents condition their votes on perceived service quality, we argue it is important to consider whether they correctly attribute credit or blame for observed outcomes to the proper agents in government. When they fail to do so—and hold the wrong office accountable or blame public officials for outcomes beyond their control—retrospective voting might create perverse incentives that lead to worse democratic outcomes, a possibility that the existing literature on local accountability largely overlooks.  

While a number of recent studies note that voters sometimes misattribute responsibility for performance, particularly in federal or multilevel contexts (e.g., Healy and Malhotra 2013; Rudolph 2003b), the consequences of this insight—and the importance of choosing the appropriate policy domain on which to focus one’s inquiry—have not been adequately incorporated into research on local retrospective voting.

In this study, we seek to address both of these limitations using novel administrative data on pothole complaints in one of America’s largest cities. Our analysis focuses on an outcome—the quality of local streets—that we show is both highly salient to voters and clearly under the control of the local government. Aside from the substantive focus on this dimension of performance, another key difference between our approach and other recent studies is that we specifically examine within-city variation in both service quality and electoral behavior. Rather than comparing election outcomes between cities or within cities over time, we leverage neighborhood-level differences in the number of citizen pothole complaints and examine how these differences affect the vote share won by the same incumbents across precincts.

Our approach recognizes that government services and outputs likely vary significantly within cities, and we demonstrate that such variation can explain neighborhood-level differences in support for incumbents from both major political parties in two political offices—mayor and city council—across several electoral cycles.

In the remainder of the paper, we proceed by providing an overview of the existing theoretical and empirical literature on voter behavior in city elections. In doing so, we situate the available research on local retrospective voting within the broader debate about the constraints facing local governments and the dynamics of local democracy. Next, we introduce our case—the city of San Diego—and describe how we combine geocoded municipal records on street repair requests with a panel of precinct-level data on election outcomes, allowing us to examine how road quality influences voter behavior in San Diego city elections. After describing the unusual political context during the years included in our analysis—a context that provides a particularly close fit between theory and our data—we present the results of our analysis. The concluding section examines the implications of our findings for democratic accountability in local government. We argue that, in normative terms, the dynamics we document could encourage two sets of behaviors among public officials. On the one hand, accountability pressures may lead these officials to prioritize and invest in the services that citizens value the most, encouraging faithful democratic responsiveness exactly as envisioned by normative theory. On the other hand, strategic politicians may also respond to the resulting electoral incentives through a variety of gaming behaviors—such as shifting funding from less to more salient services and linking budgetary outlays to electoral calendars—that, over the long term, may reduce social welfare.

**RETROSPECTIVE VOTING IN LOCAL ELECTIONS**

Applying theoretical models built to explain behavior in national elections to local democracy requires acknowledging important differences between these two levels of government. For example, while all federal elections are partisan, approximately 80% of local government positions are filled through nonpartisan contests (Kaufmann 2004). In addition, local governments generally attract much less media coverage and candidate spending, among many other important differences. Are these differences meaningful enough to affect voting behavior and, in turn, the nature and quality of democratic representation?

A large literature in urban politics argues that cities are indeed unique. Local governments face a number of constraints—including competitive pressures from nearby jurisdictions, interference from higher-level governments, and binding tax-and-expenditure limitations—that greatly limit local discretion and the menu of available policy options. As a result of these con-

---

1. Consider Arceneaux (2006). In examining unemployment, education, and traffic congestion, the author finds that voters attribute causal responsibility for performance to the level of government they consider accountable for each area. Worryingly, however, there is no consensus among the respondents in the study about which level of government is actually responsible for each policy area (see Arceneaux 2006, 742, table 2).

2. A growing empirical literature challenges this view, showing that city governments have the capacity to tailor policies to respond to the ideological
constraints, political competition at the city level often focuses on the question of how to divide a fixed amount of resources, pitting broad groups of citizens against one another. As Kauffman (2004, 18–19) summarizes:

Local politics—and the kinds of issues that dominate local elections—are often more proximate and more discrete than are the larger symbolic issues of national elections. Local governments, while sometimes lawmakers, are principally service providers. . . . To most Americans, the role of their local government is to maintain or enhance their immediate quality of life, to provide necessary services, and at times to ameliorate intercity conflict. Thus the citizen's expectations of local government are inherently more connected to daily life than are his or her expectations of other governmental bodies. And as such, how people view their local leadership and why they vote for them will likely reflect the proximity of their interests.

One implication is that constraints on local government power should influence the dimensions of government performance voters focus on when forming their retrospective evaluations. If municipal government influence is limited to a small set of basic housekeeping services, it makes little sense for voters in local elections to hold city officials accountable for outcomes that are largely outside of their direct control.

Surprisingly, however, much of the empirical literature on retrospective voting in the local context focuses on just such issues. For example, Berry and Howell (2007) analyze how student achievement affects the reelection prospects of school board members, although more than 80% of the variation in academic achievement is explained by student-level factors, such as socioeconomic status (Chingos, Whitehurst, and Galbraith 2015; Kogan, Lavertu, and Peskowitz 2016b) rather than local educational policies. In addition, the largest remaining determinant of achievement that is plausibly under the control of local officeholders—teacher quality—varies far more within schools (and districts) than between them (see Chetty, Friedman, and Rockoff 2014), providing few clear expectations for electoral behavior.

In a recent study, Hopkins and Pettingill (2015) examine the electoral performance of mayoral incumbents in big US cities. Although their model includes a number of policies plausibly under local government control, none of these factors appear to correlate with incumbents' vote share. The only variable that consistently produces substantively large effects is the local unemployment rate. While mayors can certainly affect the local economy, factors such as the spatial distribution of industries, occupations, and natural resource endowments likely explain much more of the variation in subnational economic performance. Indeed, Miller (2013) finds that the win-loss records of local professional sports teams—a statistic over which government officials clearly have no control—frequently influence local mayoral elections. He interprets this result to mean that voters follow a "Prosperity Model" of voting, holding incumbents accountable for their overall well-being, regardless of whether government policies played any actual role in determining their level of happiness.

The existing literature rarely acknowledges the disconnect between what local governments actually do and where analysts search for electoral accountability in local elections. When scholars do note the apparent discrepancy, they nevertheless draw sanguine conclusions about the normative implications for local democracy. Miller's (2013, 16–17) discussion of the likely welfare implications of voters relying on sports records when deciding whether to reelect their incumbent mayors offers a case in point:

The flip side to how voters choose their leaders is how this incentivizes politicians in office looking for reelection. Under the Prosperity Model, incumbent politicians are motivated to maximize voter happiness at the time of the election. This is hardly the gravest of threats, even if motivating politicians to maximize well-being in the long run is preferable. . . .

It is fair to say that voters are imperfect and occasionally irrational. It does not follow that democracy is seriously questioned. In fact, voting for incumbents when well-being is high is a sensible rule of thumb for voters who cannot reliably connect the dots between political actions and outcomes. Although irrelevant events may seep into this calculation, voting remains an effective method for selecting and disciplining leaders.

We contend that such conclusions may not always be warranted and certainly do not follow logically from existing research. Consider another finding from Arnold and Carnes (2012), who use several decades of panel data on New York City mayoral approval to show that voters appear to blame local elected executives when crime levels rise and credit them when crime declines. In reality, however, there is little evidence that local governments have much control over crime rates. In a recent exhaustive analysis, Roeder, Eisen, and Bowling (2015) show that much of the decline in
crime that occurred between 1990 and 2013 was a result of broad national trends, such as falling alcohol consumption, rising income, an aging population, and the legalization of abortion. While local policies such as the introduction of data-driven law enforcement approaches (i.e., CompStat) and the growing size of local police forces played a small role on the margin, forces over which local governments had no direct control dwarfed these effects. Yet voters—at least in New York City—appear to have credited local officials anyway.4

If “retrospective voting” is merely a reflection of irrelevant considerations, the implications for local democracy are potentially far more troubling and the incentives for local officials more perverse than Miller (2013) suggests.5 If mayors know that voters will hold them accountable for crime, but also realize that there are few tools at their disposal to influence crime levels directly, they may be tempted to shape the few things that are under their control, namely, how crime rates are measured and reported. Indeed, recent scandals in both Chicago (Bernstein and Isackson 2014) and Los Angeles (Poston and Rubin 2014) provide reason to worry. In both cities, mayoral administrations faced growing political pressure to reverse rising violent crime rates. In the absence of policy tools to do so and with elections approaching, mayors in both cities put pressure on their police departments to misclassify serious crimes to make the statistics appear less dire.

Thus, in determining whether local democracy “works,” we argue it is insufficient to measure the extent to which observed policy and performance outcomes shape incumbent electoral prospects. It is also necessary to consider whether voters allocate credit and blame appropriately, correctly linking these outcomes to the government officials responsible. Our study moves in this direction by shifting the focus to precisely the kinds of services controlled by city government—in our case, the quality of local roads.

Our second contention is that, in evaluating voter behavior, it is important to pay attention to within-city heterogeneity in service quality. As Trounstine (2016, 720) notes, “Because local government decisions often concern spatial allocation, neighborhoods are important municipal actors in local politics.” Voters often evaluate the performance of local government on the basis of the quality of life in their neighborhood rather than that in the city as a whole. If voters indeed vote retrospectively, we expect their behavior to reflect the quality of the services that they receive, which we can rarely infer from average citywide outcomes that are the typical focus of analysis in the existing literature.6 In the sections that follow, we describe how our study incorporates both of these key theoretical insights.

FISCAL CRISIS AND GOVERNMENT SERVICES IN “ENRON-BY-THE-SEA”

As is the case with most of the research on local politics, our study examines a single city, San Diego. One disadvantage is that doing so potentially limits the generalizability of the findings. Fortunately, this is much less of a concern in our case because the municipality we examine is similar in important respects to most other major and mid-size US cities (see app. A; apps. A–K are available online). In particular, San Diego is ethnically and racially diverse, overwhelmingly Democratic, utilizes a “strong mayor” form of government,7 and elects city officials through nonpartisan ballots. Across these and many other dimensions, America’s eighth-largest city is a representative case, so the political dynamics we document are likely to provide useful insights about city politics writ large.

The key advantage of a single-city case study, however, is that we can utilize detailed data on both service quality and electoral outcomes at a very low level of aggregation, an empirical strategy that is simply not practical to implement across multiple jurisdictions. Since our unit of analysis is the electoral precinct and we focus on variation across neighborhoods, the effective N of our study is still significantly higher than in a typical analysis of local government election outcomes. Most important for our purposes, San Diego is a useful context in which to examine voting behavior because the city faces—and, in some ways, exemplifies—many of the competitive and legal constraints confronted by many other local governments.8

Much of San Diego’s population growth occurred in the three decades after World War II, an era that coincided with emerging fiscal pressures. In 1978, California voters passed Proposition 13, a constitutional initiative that permanently

4. Such policies may explain why crime declined faster in some cities than in others. Arnold and Carnes (2012), however, show that New York voters responded to the aggregate changes in crime over time, not just to the portion of that change attributable to local policy choices.

5. Arnold and Carnes (2012) take a similarly glass-half-full view, concluding, “The electoral connection provides powerful incentives for mayors to do their best in tackling whatever problems occur on their watches, just as it does for elected executives around the world” (962).

6. Berry and Howell’s (2007) finding that voters appear to place greater weight on the achievement levels in their neighborhood schools, rather than the average achievement in each school district, is consistent with this argument.

7. This is technically known as a mayor-council, as opposed to council-city manager, form of government.

8. The political evolution of the city in the second half of the twentieth century is examined in detail by Erie, Kogan, and MacKenzie (2011), so we summarize only the most relevant facts in this section.
reduced and froze property taxes and limited the reassessment of property values. Although the measure produced a sizable negative revenue shock for many local governments in California, the effect was particularly large in San Diego, which had unusually low property taxes prior to the change. As a result, Proposition 13 permanently locked in the city’s low revenue (Erie et al. 2011, chap. 3).

As San Diego continued to grow, local officials found creative ways to supplement local tax revenues to balance the annual budget. Starting in the early 1980s, for example, the city began diverting “surplus” earnings from its municipal pension system to pay day-to-day operating expenses. When several ambitious projects further strained the budget in the mid-1990s, the city struck a secret deal with its independent pension board to underfund the pension system. In exchange for their support to allow the city to pay less than its actuarially required pension contribution, city employees received a significant increase in pension benefits. City finances faced additional pressures after the 2001 recession, prompting the pension deal to be expanded, with the city continuing to underfund the pension system in exchange for yet more benefit increases for city workers.

After journalists brought to light these underfunding schemes in the early 2000s, the city was left on the brink of municipal bankruptcy. It significantly increased its pension contributions to help close a now multi-billion-dollar unfunded pension liability. Because San Diego government officials did not adequately disclose the pension underfunding on official financial statements, the city’s external auditor refused to sign off on its financial statements, which effectively banned San Diego from the municipal bond markets. With the city poised at a political and fiscal precipice, San Diego attracted growing national attention, with a New York Times headline describing the city as “Enron-by-the-Sea.” Unable to borrow and facing a ever-increasing pension bill, the city made dramatic reductions to public services and allowed basic infrastructure repairs to go unaddressed. By 2010, the bill for unfunded capital maintenance exceeded $1 billion, more than the city’s annual day-to-day operating budget.

Road quality and citizen satisfaction
As documented in Erie et al. (2011), San Diego’s fiscal and pension crisis resulted in a substantial reduction in service levels and quality. The city cut its library and recreation center hours, downsized its work force, and reduced essential services. One of the most striking changes, however, was the continued deterioration of city streets and sidewalks. With no bonding capacity and little spare operating revenues with which to carry out even the most basic street maintenance, road conditions deteriorated rapidly. One of the city’s major news organizations created a website on the photo-sharing service Tumblr, which it named “The Stumblr,” for readers to submit pictures of the worst streets and sidewalks. Even in 2015, years after the city’s fiscal crisis abated and revenues rebounded, a national assessment found that over half of San Diego’s streets were in poor condition, ranking it the eighth worst in the country (“Bumpy Roads Ahead: America’s Roughest Rides and Strategies to Make Our Roads Smoother” [TRIP 2015]).

When San Diego officials commissioned a representative survey of city residents in 2010, satisfaction with the quality of city streets was ranked second-to-last among the 25 services evaluated by respondents, just ahead of downtown parking availability. In a Priority Spending Index, which combined citizen satisfaction with existing services and their willingness to pay higher taxes or fees to fund maintenance or avoid further cuts, city streets beat out all other priorities by a large margin (see fig. 1).

EMPIRICAL STRATEGY
Given the salience of street quality, as documented in the city’s survey, San Diego provides an excellent test case to examine whether voters hold local government officials accountable at the ballot box for observed performance. If we find little relationship between local street conditions and voting behavior in San Diego, we can be confident that retrospective voting is unlikely to operate for less salient dimensions of performance or in other locales. On the other hand, if we find that voters do hold incumbents accountable, the magnitude of these effects can speak to the likely incentives that local officials face while in office under conditions most likely to produce accountability pressures. In this section, we provide an overview of our empirical strategy for carrying out just such an analysis. It describes the variables we use and the statistical models we estimate to measure the impact of road quality on voter support for local incumbents.

Independent variable: Citizen pothole complaints
To identify the effect of municipal service quality on incumbent political prospects, we leverage geographic variation in road conditions between neighborhoods in San Diego. In doing so, we assume that voters care most about the quality of government services that they personally receive rather than
the average service quality citywide. Although there are a number of potential justifications for this assumption, the most plausible one is that voters are likely to form their retrospective judgments based on casual observations from their own day-to-day experiences (Popkin 1994). It seems much more realistic that voters will notice whether a pothole has developed on their city block than to assume that they calculate the average asphalt quality index for all city-maintained streets.

Neighborhoods in San Diego differ in the quality of their roads and the number of potholes that develop on them for a variety of reasons. The city repaved or ressealed individual streets in different years, for example, which means that asphalt in some areas is simply older than in others. There is also variation in the material used to construct the roadway and the level of daily traffic flows, both of which can affect the rate of deterioration. Directly or indirectly, almost all of these factors reflect policy decisions made by city officials.10

The key advantage in focusing on within-city variation in road quality is that such differences between neighborhoods are plausibly independent of incumbents’ expected electoral performance. This is particularly true in our case, because we explicitly condition on each incumbent candidates’ lagged level of support from the previous election. Conditional on previous electoral performance, there should be no significant differences between neighborhoods where potholes develop (or where they are reported) and those where no such reports are made, allowing us to credibly identify the effect of service quality on voting behavior. As we discuss in the next section, controlling for lagged vote share significantly increases the internal validity of our design by allowing us to rule out plausible omitted variables. As an additional robustness check, however, we present results from a placebo-style specification that further strengthens our ability to make causal claims about the relationship between road quality complaints and voting behavior.

Our key independent variable is the number of citizen pothole complaints, aggregated at the level of the electoral precinct. The data, covering the years 2008–11, were first obtained by the Voice of San Diego, an online investigative journalism organization, through a public records request.11 The database contains more than 52,000 complaints and service requests involving city-maintained streets recorded by the San Diego’s street maintenance department during these years.12 Appendix B describes the geocoding protocols 11. Voice of San Diego generously agreed to share these data with the authors.
12. Constituents have a number of ways to transmit these complaints to the city. First, they can contact their local city council member, whose staff pass the information on to the Streets Department. Alternatively, residents can send the information directly to the department using an online form or a phone-based web application. Each pothole is logged only once into the city’s database, based on the earliest complaint made, and we do not observe the total number of complaints received about any given pothole in the data. After a complaint is registered in the city’s database, a crew of either city employees or contractors is dispatched to evaluate the problem and carry out any repairs that are necessary. Each crew is responsible for an assigned geographic area of the city. Because more complaints may come in from some neighborhoods than others on any given day, the response time may vary as well, but city officials assert that “complaints are treated equally regardless of source or area or origin” (Kyle 2009, 2012).
we used to map the potholes to voting precincts. Focusing on citizen complaints, rather than aggregate measures of road condition, also ensures that we measure service quality using the information that is clearly visible and salient to local residents.13

Table 1 provides a breakdown of service calls recorded by the Streets Department. Overall, more than two-thirds of the recorded complaints focused on potholes, although the data include a number of other, more serious road repair requests. For convenience, we refer to all of these service requests as pothole complaints in the text below.

One indicator that citizen complaints reflect actual problems with service quality is that the frequency of reported potholes is positively correlated with a pavement “overall condition index,” an objective measure of road quality (see app. C). Another indicator is the record of the city’s responses to the complaints. As we show in appendix D, crews dispatched to assess the roadway in response to the complaints ended up carrying out some kind of repairs in almost 98% of cases. This does not mean that each pothole is equally likely to be reported to the city, but it does suggest that the reports made reflect legitimate grievances.

Our empirical strategy accounts for the fact that such grievances may not always translate perfectly into votes. Both theoretical and empirical literature on voting decisions illustrate that voters’ use of performance information is subject to behavioral biases, such as myopia. Recent performance, such as macroeconomic conditions in the six months before an election, has a disproportionate effect on electoral outcomes (e.g., Healy and Lenz 2014; Huber, Hill, and Lenz 2012, Kogan et al. 2016b). We expect voters to exhibit similar myopia in our local context. As a result, we calculate a count of total complaints filed in each precinct in the six-month window immediately preceding each election using the notification dates recorded in the city’s database.14 As we discuss below, we also use these date stamps to construct a count of pothole complaints made in the six-month window after the election. Because potholes reported after Election Day should logically have little impact on incumbent performance, this post-election measure provides us with a useful “placebo” test, which we will describe in more detail below.

Table 1. Summary of Service Calls

<table>
<thead>
<tr>
<th>Problem Code</th>
<th>Total Complaints (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major asphalt repair</td>
<td>22.7</td>
</tr>
<tr>
<td>Minor asphalt repair</td>
<td>8.9</td>
</tr>
<tr>
<td>Pothole</td>
<td>68.4</td>
</tr>
</tbody>
</table>

**Dependent variable: incumbent electoral performance**

The dependent variable used in our analysis is the share of votes won by incumbent candidates in each election. In particular, we focus on two elected offices most responsible for making policy and setting funding levels for city services, the mayor and the city council. During our period of study, San Diego held two citywide elections, in 2008 and 2010. In each year, local candidates competed in a June primary. If no candidate received more than 50% of all votes cast for each office, the top two vote-getters proceeded to a runoff in November.

Although San Diego had eight city council districts during this period, term limits mean that we observed incumbent council members running for reelection in just two districts in these years. Thus, our analysis focuses on three office holders: Mayor Jerry Sanders and Councilmen Kevin Faulconer and Tony Young.15 Sanders stood for reelection in 2008, while the two councilmen both ran in 2010. All three candidates won more than 50% of the vote in the June primaries, securing their reelection without a November runoff. It is worth emphasizing that our data set includes incumbents from both major political parties—both Sanders and Faulconer are Republicans, while Young is a Democrat—although elections are technically nonpartisan in San Diego, and party labels do not appear directly on the ballot. Appendix E describes how we handle changing precinct boundaries over time, to make the election data comparable across electoral cycles.

**Statistical models**

Our baseline specification of each incumbent’s vote share in precinct $i$ at the time of his reelection, $\text{Incumbent}_{i,t}$ is

$$\text{Incumbent}_{i,t} = \alpha + \tau \text{Potholes}_{i,t} + \beta \text{Incumbent}_{i,t-1} + \epsilon_i.$$  

13. Like the proverbial tree in the forest, if no one noticed or reported a pothole to the city, there is little reason to expect it to affect electoral outcomes.

14. The choice of a six-month window also reflects, to a certain extent, data availability. The earliest complaints we observe are from January 2008, and the first election during the period covered in our data took place in early June 2008.

15. We exclude open city council races for two reasons. First, it is less plausible that responsibility follows the incumbents’ co-partisans in these races, because partisan labels do not appear on the ballot, making partisanship less salient. Second, most of these races feature multiple candidates from the same party, and there are few theoretical expectations about which co-partisans voters should hold responsible for the incumbents’ performance.
We include each candidates’ lagged vote share from the previous election, Incumbent\(_{it-1}\), to address several important concerns about omitted variable bias and reverse causality. In particular, there are three possible threats to internal validity in our cross-sectional setting. First, San Diego exhibits significant racial residential segregation. Because race and ethnicity are strongly correlated with partisan identity, the residential patterns affect the political geography of the city. The heavily Democratic areas are concentrated in the oldest, poorest, and most industrial parts of the city south of Interstate 8. This introduces the potential for a spurious correlation between local road quality and partisan political dynamics, creating the appearance that voters are acting in a retrospective manner when in fact they are not. Indeed, if we simply regress Mayor Sanders’s lagged vote share from 2005 on pothole complaints three years later, from the first half of 2008, we find a significant correlation, providing clear evidence that potholes are not “as-if randomly” assigned between neighborhoods.

Second, it is also possible that that local officials strategically allocate city resources and capital investments to certain constituents based in part on their expected or previous level of support in each neighborhood (e.g., Cox and McCubbins 1986). If policy makers direct road resurfacing funds to neighborhoods containing their strongest supporters or fix potholes quicker in these neighborhoods before voters have a chance to notice them and complain to the city, we would see a correlation between voting behavior and road quality, but the direction of the causal arrow would be reversed.

Finally, one additional complication is that our independent variable is based on pothole complaints rather than an objective measure of road conditions, although we show in appendix C that there is a strong empirical relationship between the two. Theoretically, this is an advantage, since we can focus only on considerations that are clearly salient to at least some constituents. There is, however, also a possibility that voters who already dislike the incumbent are simply more vigilant about noticing or reporting street maintenance issues, inducing a correlation between pothole counts and precinct-level voting behavior.

Fortunately, conditioning on lagged vote share helps address all three of these potential problems. This point is worth emphasizing: although it is improbable that pothole complaints are “as-if randomly” assigned, controlling for each incumbent’s previous vote share in the model goes a long way toward ensuring that this treatment is conditionally independent of our outcome of interest. Indeed, the only remaining omitted variables would need to (i) be correlated with pothole complaints, (ii) be correlated with candidate vote share at time \(t\), but (iii) not be similarly correlated with the same candidates’ vote share in the same precinct at time \(t-1\).

In addition, these relationships must be present in both the 2008 mayoral race and the 2010 city council contests for both Democratic and Republican incumbents. Although we cannot completely rule out the possibility that such time-varying variables exist, it is difficult to think of any that satisfy all of these conditions. We discuss two other estimation strategies—the fixed effects and first-difference specifications—in appendix F and explain why our modeling strategy makes the most sense in our context.

In addition to including a lagged dependent variable in our model, we can also use the post-election complaints as a useful placebo test. If there are omitted variables that are correlated with voting behavior and the propensity to file pothole complaints (even after controlling for voting behavior in the previous election), these variables should affect complaints filed in the six months after the election just as much as they effect complaints in the six months beforehand. Clearly, however, post-election complaints are significantly less likely to affect the retrospective judgments of voters on Election Day. In our setting, it is not possible to carry out a traditional placebo test, in which we replace the pre-election complaints variable with the post-election complaints because the number of potholes within neighborhoods is strongly correlated over time. This is true because potholes are most likely to be reported during both periods in areas with the worst-maintained asphalt, and the condition of the pavement does not change significantly in the six months after the election compared to the six months prior. Instead, we include the post-election complaints count as an additional control in our model. Conditioning on post-election complaints should account for any other omitted variable that is not already captured by our lagged dependent variable.

**Spatial interdependence**

The nature of the data presents another potential econometric complication. Since service levels are likely to be similar between nearby neighborhoods—for example, voters...
residing on two adjacent streets are likely to patronize the same local library branch or recreation center—a simple ordinary least squares (OLS) specification does not account for the spatial relationships between precincts. To account for the fact that service quality is likely to be more similar among nearby precincts than between distant locations, resulting in spatial clustering in the error terms, we also estimate models that add a spatial error component to our baseline specification:

$$\text{Incumbent}_{it} = \alpha + \tau \text{Potholes}_{it} + \beta \text{Incumbent}_{it-1} + \lambda w_i \xi_{i1} + \epsilon_i,$$

where $w_i$ is a spatial weights matrix specifying the relationship between precincts and $\xi_i$ is a vector of error terms.

As a final robustness check, we also estimate a spatial lag model, which allows voting behavior in one precinct to be affected by its proximate neighbors. The motivation for the spatial lag is that voters likely notice and care about road quality not only in their neighborhood but also the condition of the other streets they use regularly, perhaps during their daily commutes. This specification thus allows vote share in precinct $i$ to be affected by potholes in the same precinct but also in other nearby areas. The precise specification is

$$\text{Incumbent}_{it} = \alpha + \tau \text{Potholes}_{it} + \beta \text{Incumbent}_{it-1} + \rho w_i \text{Incumbent}_{it} + \epsilon_i,$$

where $w_i$ is again the spatial weights matrix and $\text{Incumbent}_{it}$ is a vector of election outcomes.\(^{19}\)

Theoretically, it is most plausible that the interrelationships between precincts are inversely proportional to their distance from one another. To reflect this, our preferred specifications use spatial weights that are an inverse of the quadratic distance between precincts. This corresponds to spatial effects that decay exponentially as the distance between precincts grows.\(^{20}\)

For the city council elections, we pool the results from the two districts, but we modify the models above by adding the subscript $d$ to the intercept term, $\alpha_d$, representing the district-specific fixed effects. Because we observe city council elections with both Democratic and Republican incumbents, this ensures that the effects we document are not the result of partisan-motivated attribution of responsibility (e.g., Rudolph 2006).\(^{21}\)

### RESULTS

To ease substantive interpretation and allow for straightforward comparison across models and variables, we present our results visually in a series of plots, with full regression output available in appendix H. Figure 2 presents the results of our initial specifications. The top part of the figure corresponds to the baseline OLS model, the middle to the spatial error specification, and the bottom portion contains the spatial-lag model.

Across all specifications and both office types, we find that pothole complaints have a significant negative effect on incumbents’ vote share. When constituents submit more road work requests in the six months before the election, incumbent officeholders suffer at the polls. The coefficients are similar in size for both the mayoral and council elections, with each additional pothole complaint reducing incumbent vote share by roughly 0.2 percentage points. Increasing the number of potholes from the 25th percentile in 2008 (zero potholes) to the 75th percentile (six potholes) is thus expected to reduce the incumbents’ share of the vote by about 1.2 percentage points. Although this effect is not particularly large, it is arguably meaningful in substantive terms. In a close election, local road quality could prove pivotal to whether the incumbent wins another term or loses the election. Note that the coefficients on the spatial-lag specifications are somewhat smaller because they include only the short-run rather than the full-equilibrium effects that account for the dynamics of the spatial spillovers between precincts, and thus they are not directly comparable to the other two estimates (see Ward and Gleditsch 2008, 44–49).

We present our placebo specifications, which include separate variables for both pre- and post-election pothole complaints, in figure 3. Including the complaints made after the election does not have any measurable impact on our estimates. Across the models, the coefficients on the pre-election variable are actually somewhat larger. By contrast, the coefficients on the post-election variable are statistically indistinguishable from zero, and they have the wrong sign.

---

18. For example, using the spatial weights we describe below, the Moran’s $I$ for the residuals from the OLS specification of the mayoral vote share model is .19 ($p < .001$). This indicates significant positive spatial autocorrelation, with the residuals in nearby precincts being more similar than those in precincts that are more distant from one another. Moran’s $I$ can range from $-1$ to $1$, so while the spatial correlation is significant in our case, it is relatively modest in magnitude.

19. As an additional robustness check, we have also estimated a spatial Durbin model. The results were identical to those below, so we present them in appendix G.

20. Our results are not sensitive to this choice. We find the same substantive effects if we instead use an inverse of the linear distance or apply equal weights to all observations within a mile of each precinct.

21. This might occur if voters blame the mayor’s co-partisans for potholes, regardless of whether they are the incumbents or challengers in the race.
across the board. These results significantly strengthen our ability to provide a causal interpretation to the documented effects of citizen pothole complaints made before the election on incumbent political support.

In addition to our main specifications, we carried out a number of robustness tests. In appendix I, we re-estimate all models after dropping the more serious asphalt repair requests, limiting our analysis only to pothole complaints as classified by the city. The results are substantively identical, and indeed the coefficient on the pothole count variable becomes somewhat larger. In appendix J, we also estimate separate models for each of the two council districts. The coefficients are nearly indistinguishable from those in the pooled models, although they are less precisely estimated in the case of Councilman Tony Young’s District 4 due to the much lower sample size in the disaggregated analysis. We have run several additional checks, including transforming the pothole count variable into a relative quartile ranking to address concern about outliers, re-estimating the mayoral election analysis in the subset of precincts that are used in our city council election results, and including President Obama’s 2008 vote share as an additional control. Overall, the results are highly consistent across different specifications, elections, and subsets of the data, increasing our confidence in their validity.

It is also important to note that, across all elections and specifications, we find a strong relationship between current electoral performance and lagged vote share. That should not be surprising—incumbents who perform well in a neighborhood in the previous election are likely to do well there again next time they run for office. The coefficient in the mayoral election, however, is not particularly large, with Mayor Sanders gaining roughly 0.3 votes in 2008 for each vote he won during the 2005 special election that first brought him to office. This relatively modest relationship likely reflects important differences in dynamics between these two elections. In the 2005 mayoral runoff, Sanders’s opponent was Democratic City Councilwoman Donna Frye, a self-styled progressive and environmental activist. In 2008, by contrast, local Democrats failed to recruit a viable opponent. Instead, Sanders’s leading challenger was Republican businessman Steve...
Francis. Although Francis tried to recast himself as an independent populist in 2008, he was widely considered to be more conservative than Sanders, a moderate Republican who publicly embraced same-sex marriage long before most other elected officials—and certainly before nearly all Republican elites—did so. Given these important differences, it is not surprising that Sanders’s winning coalition would change significantly between these two elections.

**Mechanisms: Changing minds or changing voters?**
Overall, we find consistent evidence that more potholes in individual neighborhoods result in fewer votes for incumbent San Diego officials. But the findings presented above do not speak to the precise mechanisms producing these effects. Incumbents could receive fewer votes because voters change their behavior—after observing the quality of their local roads, voters may update their beliefs about the competency of current officeholders and their qualifications vis-à-vis the challengers. Alternatively, incumbents could also suffer if potholes alter the composition of the electorate. For example, poor-quality streets might mobilize their opponents, increasing turnout among voters predisposed to vote against the incumbent, or, alternatively, they may depress participation among their political base.

We can offer some insights about the mechanisms involved by examining the impact of pothole complaints on voter turnout at the precinct level. As we report in appendix K, there appears to be little relationship between complaints filed before the election and the rate of voter participation on Election Day. We should stress that these results are merely suggestive: it is possible that the decrease in participation among the incumbents’ supporters is nearly perfectly offset by higher turnout among their opponents, producing no net change in the number of voters casting ballots. Nevertheless, these results suggest that the accountability effects we observe are likely driven by voters changing their minds rather than by potholes altering who participates in the elections.

**Substantive significance**
It is also useful to put the effects we document in context by comparing them to the findings on economic voting in US presidential elections. In these contests, incumbent presidents gain roughly 4 percentage points of vote share for each percentage point of election-year income growth (Healy and Lenz 2014). Thus, the electoral impact of a single pothole on city contests is roughly one-twentieth the magnitude of a single percentage point of income growth in national presidential elections. Of course, the average election for city office in San Diego is considerably less competitive than is the case for a typical presidential election, so the political importance of the effect may be more modest in terms of its capacity to motivate officeholders and alter their incentives while in office. Nevertheless, if the electoral consequences we document for road quality extend to voter satisfaction with other city services, retrospective considerations may help shape the electoral fortunes of some city officials and thus potentially serve as a mechanism of accountability.

**DISCUSSION AND CONCLUSION**
Overall, we find strong evidence that government performance in the provision of a particularly salient public service has a meaningful, albeit modest, impact on city elections in San Diego. That the results are consistent across three different modeling approaches, each of which makes different statistical assumptions about voter behavior, gives further credence to their validity. Compared to previous studies, our analysis offers a much closer fit between theory and empirics by focusing our inquiry on outcomes that voters report are salient to them and that are actually under the control of local government. In addition, our analysis relies on a more credible identification strategy by exploiting plausibly exogenous variation in service quality between neighborhoods rather than differences in average outcomes between cities.

Nevertheless, our analysis is limited to one dimension of performance during two elections in a single city, so it is important to consider the generalizability of our findings. Although San Diego’s municipal government structure and election rules are typical of other large and mid-size cities, suggesting that we should expect the electoral dynamics to play out similarly in other locales, there were several attributes unique to these elections worth highlighting. As we noted previously, service quality featured prominently in civic dialogues and political campaigns in San Diego during this period due to the cutbacks that occurred in the wake of the city’s pension crisis, and our analysis further focuses on the service that voters themselves said they cared about the most. If salience and magnitude of the failures (or successes) moderate the extent to which voters weigh government performance in service delivery when casting their ballots, we might expect to see smaller effects for other dimensions of performance and in other contexts. Similarly, there may be an asymmetry in voter response if they punish incumbents more for service reductions than they reward them for expansions. On the other hand, most of the potholes in our data set did ultimately get fixed in response to citizen requests. This might have helped blunt the elec-

---

24. Since 2000, the average margin of victory in all city contests has been more than 20% (the median was 15.3%).
toral consequences of individual potholes—both because individuals who filed the complaints may have rewarded the city government for its responsiveness to their requests and because making repairs in a timely manner reduced the overall number of drivers who had a chance to notice them. In contexts where government officials are less responsive, we might expect similar complaints to exact a larger electoral toll. All of these possibilities deserve close attention in future research on retrospective voting in local elections, which should further explore such potential heterogeneity.

While our findings are hopeful in normative terms, we believe they are not sufficient to reach definitive conclusions about the quality of American democracy. Whether the type of retrospective voting we document improves democratic accountability or perversely leads to less optimal policies ultimately depends on how government officials respond to the resulting electoral incentives, something we do not explore in our analysis. Two sets of responses are possible. First, knowing that their electoral fortunes depend in part on voter satisfaction with government performance, incumbents may work to maximize the quantity and quality of services their constituents value the most. This would be a desirable outcome and one envisioned by the standard sanctioning models (Ferejohn 1986). Alternatively, however, incumbents could respond strategically to the electoral incentives we document in ways that may ultimately reduce social welfare. For example, elected officials might leverage the temporal biases in voter behavior and systematically cut investment on preventative street maintenance early in their terms to build savings that might later be used to ramp up spending on repairs only in the months leading up to each election (Nordhaus 1975), increasing their credit claiming opportunities but also potentially the long-term costs. Alternatively, strategic politicians might respond to voter demands for well-maintained streets by simply redirecting resources from other essential but perhaps less visible or salient services, such as sewer and storm water systems.

We believe it is also critical to consider whether voters accurately attribute credit and blame for observed performance outcomes to the right individuals in government. Indeed, all of the incumbents we examine in this study were elected to office in the aftermath of San Diego’s pension crisis, and none of them were arguably to blame for the dire service declines the city has experienced over the past two decades. It is unclear (at least to us) how holding Mayor Sanders and Councilman Faulconer and Councilman Young accountable for the sins of their predecessors contributes to better democratic performance or responsiveness. Ultimately, the relationship between democracy and social welfare depends crucially on the types of retrospective evaluations voters bring to the ballot box and on how elected officials respond to the resulting political incentives. While our study provides a useful set of findings that speak to the former question, we do not view it as the final word on the topic.

ACKNOWLEDGMENTS

We thank Scott Lewis and Keegan Kyle at the Voice of San Diego for providing us with the pothole complaint data and Shaun Fontanella, Morton O’Kelly, and Chen Xiang at Ohio State University’s Center for Urban and Regional Analysis for their geocoding assistance. Dan Hopkins, Stéphane Lavertu, and Jessica Trounstine provided us with invaluable feedback and comments on earlier drafts. The authors’ names are in alphabetical order.

REFERENCES


Behavior Research Center. 2010. “City of San Diego Residents’ Opinions on City Services.” Technical report, April, Behavior Research Center, Phoenix, AZ.


