Is it All Good in the Neighborhood? How Partisanship May Distort Evaluations of Municipal Services

Lucas Boschelli*

March 22, 2022

Abstract

Do voters retrospectively evaluate municipal services? Previous work within local politics would suggest that voters form their evaluations based on the quality of the service and their access to it. Instead, I argue that voters evaluate associated services through a partisan lens rather than objective performance due to the nationalization of a particular state and local political issues. This process occurs when local services become polarized at the national level, with the two parties being associated with distinct and opposing views on those services. I attempt to test this argument through a cross-sectional analysis of local school and police evaluations. The results confirm that for polarized services such as policing, individuals have a systematic bias in favor of their party's position regardless of the service's objective performance. Additionally, I find that this bias exists regardless of the partisan control of state and local governments. These findings provide insight as to how nationalization shapes retrospective evaluations of government performance and carry with them implications for the future of local accountability.

*Ph.D. Student Department of Political Science, Washington University in St. Louis, One Brookings Drive, St. Louis MO 63160. boschellil@wustl.edu

1 Introduction

While there has been extensive work examining how voters evaluate the performance of federal and state governments, comparably less attention has been paid to evaluations in the context of local governments. In what work has been done, researchers often focused on small geographical areas and examined how information availability shapes evaluations (e.g., Berry and Howell 2007; Payson 2017). We still know relatively little about how voters' individual-level characteristics shape how they evaluate their local governments and their respective services. Nor do we know how well these evaluations correlate to the actual performance of local governments. Traditional conceptions of local politics might suggest that these evaluations are primarily a result of voters' day to day interactions with these services(e.g., Kaufman 2004; Yoder 2020); however, this framework fails to take into account the growing role partisan identities play in voter's evaluations of government services at state and federal levels(Healy and Malhotra 2013). Furthermore, the increasing nationalization of politics means local political issues are no longer as insulated from traditional partisan discourse as previously conceived (Hopkins 2018).

In this article, I attempt to further our understanding of how voters evaluate their local governments' performance and the role partisan identities play in the construction of these identities. I argue that local services become polarized when local issues become nationalized, and distinct partisan cleavages form in the national discourse. For these polarized services, partisanship becomes a powerful identity through which individuals construct their local evaluations. This partisan lens then biases assessments in the direction of national party preferences.

To test this argument, I use a nationally representative and geographically diverse survey that asks individuals to grade two prominent local services, schools and police. While both are primarily under the purview of local governments and receive significant political attention, they differ regarding their partisan division. Through the use of administrative records from both the Census Bureau and National Weather Service, I then match all respondents to their respective local governments. This matching allows me to, in addition to individual characteristics, (1) account for locality-based effects that may influence a respondent's evaluation, and (2) assign each respondent an objective measure of both services to isolate any form of partisan bias.

Overall, I find a significant partisan bias in voters' evaluations of polarized services even after accounting for objective performance and prominent socio-economic features. Despite living in areas with similarly performing police agencies, Republicans systematically rate local police higher than Democrats and Independents. Notably, this bias is equivalent to a substantial change in objective outcome measures for police and the effects of income and race. While I find some differences in how partisans evaluate local schools, the results are comparably small and outweighed by school performance. To test the robustness of my results, I subset my data set to only large cities in order to incorporate additional city-specific measures, alternative local services, and city and state levels of partisan control. I find support for my theory, and interestingly my results suggest that the partisan bias is based on national partisan preferences rather than on party control of the government.

This research contributes to the growing literature on the role of partisanship within local politics and provides further evidence of the nationalization of American politics. Notably, my research is one of a relatively few to map a geographically diverse collection of specifically local evaluations to their respective municipalities. My results highlight how local services, much like national policy issues, can become polarized and susceptible to partisan biases. These partisan effects are especially problematic as they can have a more influential impact on an individual's evaluations than a local service's objective performance. Additionally, these results have potential implications for the already low levels of accountability found with local politics as voters may rely more on national opinion rather than local conditions when constructing their evaluations.

2 The Process of Polarization

I argue that some local services have become polarized due to the nationalization of politics; however, I feel it is essential to outline what I mean by a service becoming polarized and how this process occurs. Borrowing from the polarization literature—namely Levendusky (2010) and Sinclair 2006—I define a local service as polarized when party conversations surrounding a service have become ideologically distanced and distinctly homogeneous. As ideological distances of elite opinions or media coverage surrounding a local service increases, individuals likely sort into their party's respective position. When this sorting becomes severe enough—there is little overlap between party positions—a local service has become effectively polarized. After polarization, individuals have apparent partisan affiliations and party positions they can lean on when constructing their evaluations.

However, we know from Hopkins (2018), which examines the trend of nationalization, that not all local politics is nationally relevant or inherently political. While some services and issues may tie into national policy issues such as schooling and education or police and criminal justice, others are distinctly local such as road management or zoning. Thus we shouldn't expect local services to polarize uniformly. Instead, the process should occur on a service by service basis. Additionally, we know that how local politics connects to an individual's political identity and set of political preferences is not always readily apparent. Often local political outcomes require some form of catalyst for individuals to make political connections to local surroundings (Hopkins 2018). Previous work in local accountability suggests that it takes news coverage or a political spotlight on a service for individuals to make political connections to a service (Payson 2017; Berry and Howell 2007). Therefore, it is likely those services, which are broadly relevant to many individuals, gain sufficient media and political attention and eventually become polarized.

2.1 How Partisanship Shapes Evaluations

A fundamental assumption of my argument is that polarization and party identity shape individuals' evaluations of government performance. Broadly speaking, we know individuals craft their political opinions through their party identification (e.g., Campbell 1980, and Zaller 1992) and that given the increasing alignment of ideological and policy preferences, an individual's overall evaluation tends to correspond to either that of their party's elites or their party's average position (Lenz 2009; Sinclair 2006). Past research in retrospective voting provides us with greater insight into why this is and its mechanisms. First, when individuals interpret policy outcomes, they have to decide how they should allocate blame or credit. Voters tend to attribute positive outcomes to their own party's performance and blame adverse outcomes on the other party (e.g., Healy and Malhotra 2013). While police agencies may not belong to a specific party, Republican individuals may associate them more with their party. Thus any negative performance on their behalf may be attributed to other political actors.

Second, partisanship can be used as a powerful heuristic when it is unclear how to interpret policy outcomes or to establish lines of responsibility. In a study examining retrospective voting in the wake of Hurricane Katrina, Malholtra and Kuo (2008) found that individuals were more likely to blame Republican officials for the deaths and damages resulting from the storm. However, these effects disappeared when respondents received the Republican official's title. There is also evidence that when information environments become polarized, individuals are more likely to adopt partisan cues even in the presence of contradictory objective information (Druckman et al. 2013). Thus when tasked with evaluating police performance, individuals may rely more heavily on their partisan identity than objective performance.

2.2 Why Schools and Police

Ideally, to analyze how local services become polarized, I could use longitudinal data on individual evaluations of various local services. Then as services become polarized in the national sphere, I could measure the divergence between partisans. Unfortunately, no such data is sufficiently detailed to ask about local services or covers an extended time frame. An alternative way to measure polarization's effects on a service is to examine evaluations between an arguably polarized and non-polarized service and compare the differences between them. I argue that local schools and police provide a good candidate pair for this comparison.

First, both schools and police are predominantly administered and funded by local and county-level governments. In the case of schools, most public schools are governed primarily through school districts, which derive over half of their total funding through local property taxes (Urban Institute 2020). School districts do have to adhere to state and federal guidelines to secure additional financing, but hiring practices, setting the teaching curriculum, modifying school boundaries, and adjusting property taxes are primarily at their discretion.[1] While there exist police agencies at almost all governmental levels, over 80% of agencies are supervised directly by municipal governments. These local police agencies also employ roughly 67% of all full-time sworn officers (BJS 2019). Much as is the case for school districts, police agencies must adhere to state and federal guidelines to secure outside grants. Still, most of their administrative, procedural, and hiring practices are left to their governing municipalities' discretion.

Second, both local schools and police agencies are both politically salient and widely relevant services. Schools and police both receive relatively consistent coverage from local and national news outlets and are often addressed by political elites. Beyond coverage, both services also play a visible socio-economic role within their communities. Each service plays a key role in their community (i.e., through direct educational services or public safety) and

[1] I should note here, that while districts are allowed to develop their curriculums, the adoption of programs such as Common Core, means districts have to meet specific educational standards. This requirement often means many districts teach towards these standards to maintain or increase funding levels.

has downstream effects such as impacts on property values or income inequality. Thus both of these services are likely already politically salient for individuals, and they have already formed some degree of political opinion surrounding them.

Third, schools and police differ regarding the degree of polarization surrounding them. In the wake of highly publicized police violence against racial minorities, opinions towards police have become increasingly partisan and differentiated. Republican party figures have come out primarily in support of police agencies, whereas Democratic have taken a far more critical viewpoint. Recent surveys have highlighted how partisans fundamentally have different views over police agencies' roles within their communities. Democrats are more likely to view police officers as enforcers, whereas Republican individuals tend to view police more as protectors (Pew Research Center 2017). While schools and education are not free from partisan division, parties are not as clearly divided. Republicans and Democrats differ along dimensions of social values and federal oversight but are similar along other dimensions, such as access to education and accountability within schooling. Unlike the case for police agencies, there does not appear to be a clear partisan divide over all dimensions of schooling. Even amongst the most divisive aspects of schooling (e.g., sex education, and prayer within schools) partisan differences are relatively small compared to other issues such as abortion or gun rights (Shapiro et al. 2016). This difference in partisan polarization allows for a practical test of my proposed theory. If polarization does shape local evaluations, I should find that Republicans systematically evaluate police higher than Democrats. Given the lack of clear partisan sorting surrounding schools, I should see no significant difference between partisans.

3 Data and Methods

Overall, I argue that local services have become polarized due to the nationalization of local issues. My goal is to highlight this phenomenon through a cross-sectional comparison of local school and police evaluations. If my theory is supported, I should find that voters are biased towards their party's position in their assessments of police, but not schools. To empirically test this difference, I rely on nationally representative public opinion data and school and police performance measures. I also use voter's geographical information to account for municipal level influences in their service evaluations.

3.1 Local Opinions and Individual Characteristics:

One of the primary challenges associated with measuring opinions at a local level is a lack of geographically diverse survey(s) surrounding predominantly local issues. However, the 2018 Cooperative Congressional Election Study fielded a small battery of questions about local services. This battery of questions was fielded to the entire sample providing 48,245 individual evaluations of local communities.[2] Specifically, the survey asked respondents to grade their local communities regarding schools, policing, roads, and zoning. I focus on how respondents rated both their local school districts and police agencies. Respondents assigned grades on a five-point scale ranging from -2, for poor, to 2, for excellent, with 0 representing an average evaluation. [3] To account for each respondent's partisan identity, I collapsed a 7-point party ID into a 3-point party ID, which I recorded with two binary indicators representing Republican and Independent. I use this approach rather than the traditional 7-point scale to account for the possibility that partisanship has a non-linear effect. A growing body of work has suggested that independents as an identity are differentiated from traditional partisans, and thus their identity may uniquely shape evaluations. In the Appendix, I replicate the primary results using a conventional 7-point party ID and find no substantive differences in the results.

In addition to these evaluations, I also take several steps to account for other socio-economic characteristics that may influence both a respondent's interaction with local goods and their partisan identity. Previous works such as Troustine (2010) found that local services

[2] This sample size constitutes the total number of respondents who answered the question(s) of interest, as well as the required demographic information below.

[3] In the original survey, respondents gave an alphabetic rating ranging from A to F. I standardized this rating to a -2 to 2 scale to ease interpretability and modeling

and infrastructure are developed strategically along the lines of race and income. [4]. Many of these identities are also highly correlated with party ID. Therefore, to address concerns of potential confounding, I first include binary indicators for respondents' Gender, race, and Income.[5] Second, I have indicators for both Parenthood and Homeownership. These variables moderate an individual's interactions with schools and police and provide a comparison to contextualize partisan effects.

Finally, I attempt to account for geographic differences between partisans. Recent work has shown that conservative individuals live farther away from city centers and in more rural areas than their liberal counterparts (Gimpel et al. 2020). These decisions drastically shape each individual's interactions with both schools and police and the forms of the services themselves. To account for this potential problem, I include indicators for self-reported urban/rural residency, which I report as Rural. While this indicator may not capture the nuance in the strategic settlement, it attempts to address the geographical and partisan confounding. Descriptive statistics for these controls can all be found within the Appendix.

3.2 Matching Geographic and Objective Measures

To examine whether partisanship biases an individual's evaluation, I require some form of an objective measure of both schools and police performance. Additionally, there is still the possibility that some unobserved characteristics of an individual's municipality may influence their evaluations. To address both of these concerns, I first match each individual to their corresponding municipality and then assign an objective measure of each local service. While the CCES doesn't record an individual's city or town, it does record their zip code. These zip codes were initially used to match respondents to salient electoral races; however, using administrative records from both the National Weather Service and Census Bureau, I assign each respondent to their nearest city.[6] Matches use state, county, and zip codes to avoid

[4]For more extensive coverage of strategic development, I direct the reader to works such as Trounstine 2010, Einstein et al. 2018, and Anzia 2013

[5]I record race as a series of dummy variables including Black, Hispanic, and Asian.

[6]I should note that matches are technically probabilistic as both agencies use a zip code's population centroid to determine whether it falls within municipal boundaries. There is a chance that individuals living

incorrect assignments. All told, I matched all 48,245 respondents within my sample of the 2018 CCES to their nearest municipality. With each respondent matched to a city, my next task was finding an appropriate objective measure for each local service.

One standard measure of a school district's performance is its standardized test scores. Since 2010, school districts in all fifty states must report their Common Core test performance to state-level education departments. These scores are often widely accessible online and frequently picked up by media outlets. Previous work has found that voters use these scores to either reward or punish incumbents in school board elections. (Berry and Howell 2007; Payson 2017) Through the use of the popular service, SchoolDigger, I aggregate the 2018 performance of all school districts covered by my CCES sample. Test scores are standardized to an increasing 0 to 5 scale relative to their percentile performance within their respective state. [7] Given these scores, however, I still face the issue of correctly assigning respondents to their respective school districts. Cities vary immensely regarding the number and layout of their school districts. Some smaller towns share a school district with other municipalities in their county, while others contain several interlocked and competing districts. These district lines often don't follow municipal or other geographical boundaries. To address this issue, I opt to take the average test performance of all school districts within an individual's zip code. This process yields a single score, School Rating, which reflects the general performance of schools within each respondent's immediate area. The average school rating for each respondent was 2.651 with a standard deviation of 0.86. The next challenge I faced was finding a measure for local police agencies.

There exists a significant debate in the realms of sociology and economics over how best to quantify police performance. One prominent way the public may evaluate their local police agencies is along the lines of crime prevention. While fine-grain crime data is available for a subset of larger cities, my matched sample of respondents resides in towns with populations ranging from under 1,000 to over 2 million. Thus, to include objective measures for most of

on the outskirts of zip codes may be incorrectly assigned to a city; however these individuals should be in part accounted for controlling for rural residency

[7] For more information regarding these scores, I direct the reader to SchoolDigger.com's methodology section

my sample, I opt to use 2018 county-level violent and property crime rates provided by the Federal Bureau of Investigation's Uniform Crime Reporting Program (UCR). Specifically, I aggregate the rates of murder, nonnegligent manslaughter, rape, robbery and aggravated assault, and record them as one variable Violent Crime Rate. Similarly, I aggregate rates of burglary, larceny, and motor vehicle theft into one measure of Property Crime Rate. Both measures are standardized to incidents per 100,000 individuals to scale relative crime rates between areas of varying populations. I chose to use the current year's crime rate as past work in retrospective voting found that individuals tend to be myopic and overweight current conditions over long term change (e.g. Achen and Bartels 2004).[8] By using county-level measurements, I do risk losing within and between city-level variation in crime rates. In addition to this, county-level rates may misconstrue crime rates for unincorporated cities compared to their surrounding counties. Despite these potential pitfalls, I still chose to use these measures due to the coverage they provide to my sample.

Beyond graphic availability, these scores are potentially problematic for other reasons. First, traditional crime rates are highly sensitive to reporting practices. These measures can depend just as much on the number of police officers in an agency or an area's developed trust in its police force as the amounts of crime (Fielding and Innes 2006). Second, these measures only capture one dimension of police work, crime prevention. A significant portion of what police agencies do is to address non-criminal social problems surrounding mental health, poverty, and addiction; however, this work is almost entirely uncaptured by crime-based metrics (Hodgkinson et al. 2019). Despite these problems, crime rates are often picked up and publicized by public officials and media outlets alike. Thus individuals are likely to use these scores, to some degree, to inform their evaluations of their local police. While I acknowledge these scores are flawed and capture a minimal scope of police performance, they are available for the entirety of my sample and often publicly salient.

[8] I should note that almost all of this previous work examined retrospective voting primarily in an economic sense. There is a chance that crime rates differ from other economic conditions, and voters interpret them differently. To investigate this, I test several alternative specifications of crime rates, such as using a lagged measure of violent and property crimes and the annual difference in county crime rates. The results can be found in the Appendix.

In the public eye, police agencies are evaluated mainly along the lines of crime prevention. While fine-grain crime data is available for a subset of larger cities, my matched sample of respondents resides in towns with populations ranging from under 1,000 to over 2 million. Thus, to include objective measures for most of my sample, I opt to use 2018 county-level violent and property crime rates provided by the Federal Bureau of Investigation's (FBI) Uniform Crime Reporting Program (UCR). From the UCR, I specifically include county-level violent and property crimes per 100,000 individuals and record them as *Crime Rate*. Using these measures allows me to scale relative crime rates between areas of varying sizes. I chose to use the current year's crime rate as past work in retrospective voting found that individuals tend to be myopic and overweight current conditions over long term change (e.g. Achen and Bartels 2004).[9] By using county-level measurements, I do risk losing within and between city-level variation in crime rates. In addition to this, county-level rates may misconstrue crime rates for unincorporated cities compared to their surrounding counties. Despite these potential pitfalls, I still chose to use these measures due to the coverage they provide to my sample.

Beyond graphic availability, these scores are potentially problematic for other reasons. First, traditional crime rates are highly sensitive to reporting practices. These measures can depend just as much on the number of police officers in an agency or an area's developed trust in its police force as the amounts of crime (Fielding and Innes 2006). Second, these measures only capture one dimension of police work, crime prevention. A significant portion of what police agencies do is address non-criminal social problems surrounding mental health, poverty, and addiction; however, this work is almost entirely uncaptured by crime-based metrics (Hodgkinson et al. 2019). Despite these problems, crime rates are often picked up and publicized by public officials and media outlets alike. Thus individuals are likely to use these scores, to some degree, to inform their evaluations of their local police. While I

[9] I should note that almost all of this previous work examined retrospective voting primarily in an economic sense. There is a chance that crime rates differ from other economic conditions, and voters interpret them differently. To investigate this, I test several alternative specifications of crime rates, such as using a lagged measure of violent and property crimes and the annual difference in county crime rates. The results can be found in the Appendix.

acknowledge these scores are flawed and capture a minimal scope of police performance, they are available for the entirety of my sample and often publicly salient.

4 Results

Both school districts and police agencies are influenced by federal, state, and local policy decisions. To capture this process and test my theory, I employ a series of methodological tests examining which factors influence an individual's evaluations as well as the mechanism through which any partisan bias may emerge. The section proceeds as follows. First, using my entire dataset, I test the core of my theory and examine whether there exists a partisan bias in individuals' evaluations of local police as compared to schools. Then, I subset my data to only large cities in order to replicate my results with more city-specific measures and alternative modeling strategies. Next, I examine whether my choice of comparison is robust, by seeing if the pattern holds for another prominent local service, roads. Finally, by including measures of mayoral and gubernatorial partisanship I extend my work to see whether the partisan bias is one of elite valence or party positioning. Overall I find general support for my theory of local polarization.

4.1 Polarized Evaluations

To model partisanship's effect on local perceptions of school and police, I opt to use a simple linear model with state fixed effects and clustered robust standard errors. I include all of the controls mentioned in the data section within the modeling process. For the analysis below, I limit the scope of my analysis to only Democrats and Republicans for ease of intepretability.[10]

Table 1 below presents the results of my linear model. Results are subset to relevant covariates and full regression results are presented within the Appendix. From the first column of Table 1, we can see that the coefficient surrounding Republican party ID is both positive and significant at the $p < 0.05$ level. This effect appears robust to the inclusion

[10]Within the Appendix, I replicate the results including independents and find no substantive differences in the results.

of both objective measures of police performance and other salient identities such as race. Identifying as Republican translates to a one-fourth of a standard deviation increase in an individual's evaluation of local police. To contextualize this effect, it is the equivalent of moving from a county with approximately 90 violent crimes per 100,000 individuals to one with a rate of over 2735. This finding suggests for highly polarized local services, partisan biases can vastly outweigh differences in objective performance. Notably, the magnitude of this effect is also similar to that of race, an identity traditionally thought to dominate local politics and evaluations of police. Thus in the context of a polarized service party ID appears to become a salient and influential identity. I should, however, caution the above findings. While party ID appears to vastly outweigh objective police performance, this fact may be a product of using county-level estimates. Individuals do appear responsive to these measures, but it may be the case that they are more responsive to finer grain measurements of crime.

Table 1: OLS Regression of Local School and Police Evaluations on Partisan Identity

	Service Evaluation:	
	Police	School
Republican	0.264*	-0.035
	(0.018)	(0.028)
Property Crime Rate	0.00000	
	(0.00003)	
Violent Crime Rate	−0.0004*	
	(0.0001)	
School Rating		0.340*
		(0.014)
Age	0.007*	0.0002
	(0.0005)	(0.001)
Gender (M)	0.004	0.034*
	(0.010)	(0.011)
Black	−0.282*	-0.016
	(0.033)	(0.035)
Income	0.025*	0.016*
	(0.002)	(0.002)
State Fixed Effects	✓	✓
Additional Controls	✓	✓
Observations	29,020	34,812
R^2	0.078	0.103
Adjusted R^2	0.076	0.101

Note: The above table presents the results of a OLS regression with state fixed effects and clustered standard errors. Crime rates are standardized to incident per 100,000 citizens. Full regression results can be found within the Appendix. Standard errors are presented in parentheses.$^*p<0.05$

For my theory of polarized services to be fully supported, however, we should see that party ID has a smaller effect in the context of school evaluations. Examining the second column of Table 1, we can see that my theory is supported. The coefficient surrounding Republican party ID does appear to be negative but is not significant at the $p < 0.05$ level. Unlike in the case for police above, how well an individual's local schools perform on standardized tests appear to be the most influential driver of their evaluation. Thus for non-polarized local goods, individuals appear to rely more on the service's objective performance and relevant identities rather than their partisan identity when forming their evaluations. Overall these results appear to support my theory of local polarization; however, the relatively small effects of violent crime rates and lack of effect associated with property crime rates may lead some to worry about my use of county measures or the mechanisms behind the partisan bias. I aim to address some of these concerns in the following sections.

4.2 A Closer Analysis of Large Cities

While the full CCES data set provides a geographically diverse sample with which to test my hypothesis, it means my measurements for local services must be inherently broad and are limited. To address this I subset my data to only the cities for which I have over 75 respondents. In total, this leaves me with 10,613 survey respondents from 72 unique cities. As one might expect, these cities are all relatively large with the smallest city of the sample, Pensacola FL, still having a population of over 52,000. This larger size, allows me to gather finer grain data regarding each cities local services as they tend to have more robust data reporting practices.[11]

For each of the 72 cities in my subset, I collected the reported 2018 violent and property crime rates directly reported by the city to the UCR program. While these measures help to alleviate concerns associated with county-level estimates, they still suffer from problems associated with using crime rates to gauge police performance. To supplement these rates,

[11] I should note that by selecting on larger cities I am selecting on a subsample of CCES respondents who prefer to live in more urban environments. In the Appendix, I provide a comparison of all covariate measures between the original CCES data set and my large cities sample.

I also include the number of employees each city employs within its respective police department. I standardize these measures to the number of employees per 100,000 citizens in order to compare across municipalities of varying sizes. While the number of employees is not directly correlated to a cities ability to combat crime, it serves as a proxy for city-level spending on policing. For school performance, I use the same standardized test scores used in the primary analysis.

One of the key problems present in my primary analysis above is that an individual's partisanship is not assigned randomly. In fact, one's partisan identity is closely correlated with a variety of other identities that influence access to local services and selective residency. In an effort to combat concerns of endogenous confounding, I employ a procedure of covariate balanced propensity score weighting (CBPSW) as presented in Imai and Ratkovic (2014). In order to improve matching, I match on all of the previous controls included in the initial analysis in addition to city-specific objective performance measures. Using these weights I then regress the effect of partisanship on evaluations of both local schools and police using both random state and municipality effects. Due to the variability associated with choosing weights, I opt to bootstrap the standard errors for my estimates. For each bootstrapped sample, I recalculate the fits and fit the hierarchical model. The results of my analysis are presented in Table 2 below.

Table 2: Large City Analysis Using CBPS Weighting of Partisanship on Local Evaluations

	Service Evaluation:	
	School	Police
Republican	−0.043	0.262*
	(0.027)	(0.029)
Constant	0.240*	0.345*
	(0.063)	(0.042)
Observations	7,591	4,956
AIC	24,304.160	14,917.710

Note: The above table presents the results of a CBPS Weighting with bootstrapped standard errors for my large cities sample. Effects are modeled using a hierarchical model in which respondents are nested within cities and then states. Respondents were matched on *Gender, Income, Education, Rural Residency, Home Ownership, Race, School Ratings, Crime Rates*, and *Number of Employed Officers.* Bootstrapped standard errors are presented in parentheses. *p<0.05

From the table above, the coefficients surrounding *Republican* once again are statistically significant at a $p < 0.05$ level for police while a similar pattern is not seen for school. Additionally, the effect of partisanship in the case of police appears to be of a similar magnitude to my initial analysis despite the inclusion of additional city-specific measures. I should note, however, that CBPSW does not completely solve all concerns of endogeneity as the process is still sensitive to large misspecifications of the modeling process. Accounting for all possible confounders associated with one's partisan identity and interaction with local services is nearly impossible for this form of observational study; however, this procedure provides more probable estimates than those generated by traditional forms of regression alone.

4.3 An Alternative to Schools

There may be some concern that schools do not provide an adequate contrast to test my theory of polarization. Given trends following nationwide shutdowns in response to the COVID-19 pandemic, some may argue that schools are far more polarized than I posit earlier within this paper. Luckily given my focus on larger cities, I can examine another prominent local service, roads. While roads may not receive as much media or political attention, they are a salient service most individuals interact with to some degree on a daily basis. One object measure of roads is traffic congestion. Using historic traffic data provided by TomTom, a popular transportation and navigation service, I assigned each city an index measure for their average level of traffic for the year 2018. *Traffic Index* is measured on a 0 to 100 scale and calculated— through the use of GPS data—by measuring the average additional travel time traffic congestion adds to a standard 30-minute trip within a given city. For ease of calculation and interpretability, I employ a hierarchical linear regression that includes both state and city level random effects and display my results in the table below.

Table 3: Hierarchical Linear Regression of Local Roads Evaluation on Partisan Identity: Full Results

	Service Evaluation:
	Roads
Republican	0.030
	(0.026)
Traffic Index	−0.003
	(0.003)
Age	-0.005^{*}
	(0.001)
Gender (M)	0.020
	(0.022)
Black	-0.175^{*}
	(0.033)
Income	0.016^{*}
	(0.004)
Constant	−0.003
	(0.077)
Observations	6,928
State Random Effects	✓
Municipality Random Effects	✓
Additional Controls	✓
AIC	18,191.980

Note: The above table presents the results of a hierarchical linear regression with random effects. All respondents are nested within municipalities within states. Full regression results can be found within the appendix. Standard errors are presented in parentheses. $^{*}p<0.05$

The results from Table 3 above suggest that one's partisan identity does not appear to play a significant role in their evaluations of local roads. However, it also appears as though the average traffic conditions don't have a significant effect on an individual's evaluations either. This finding could be because individuals use alternative cues to measure road quality such as pavement conditions. Alternatively, individuals may not have a conception for the annual performance of roads and instead base their evaluations on road conditions immediately surrounding when the survey was fielded. If this were the case annual measures of road conditions may be poor objective performance measures for local road quality. Despite this finding, the results surrounding partisanship do appear to support my theory of polarized local services. Additionally, these results fall in line with previous work examining road conditions such as Arceneaux (2005) or de Benedictic-Kessner (2018) which find an inconsistent correlation between road and traffic conditions and perceived mayoral performance.

4.4 Party Position Taking or Elite Valence?

While the results thus far have generally supported my claim that for polarized services individuals partisan identities bias their evaluations, it remains unclear as to the exact mechanism behind this bias. As mentioned previously in this paper, this bias could be individuals adopting their national party's stance towards a service (i.e. favorability in the case of policing); however, it could also be caused by some form of partisan penalty/reward depending on which party holds political control. In order to examine this possibility, I record the partisanship of each city's mayor as of 2018 and each state's governor. I opt to control for both state and local partisan control due to the potentially blurred lines of responsibility for some local goods.[12] In the case that the mayoral election was non-partisan, I rely on news articles, campaign websites, and future political offices to assign party identification. I employ a similar hierarchical regression and present the results in the table below.

[12]While work such as Arceneaux (2006) suggests that individuals are capable of making meaningful distinctions between levels of government, we know that this is dependent on their perceived responsibilities of local, state and federal governments (Arceneaux 2005). The growing trend of nationalization and politicization may lead individuals to view state and federal governments as increasingly responsible for some local policy areas.

Table 4: Hierarchical Linear Regression of State and Local Co-partisanship on Local Service Evaluations

	Service Evaluation:		
	Police	Roads	Schools
Co-partisan Mayor	−0.038	0.013	0.042
	(0.034)	(0.032)	(0.030)
Co-partisan Governor	0.015	0.045	0.052
	(0.031)	(0.025)	(0.028)
Republican	0.246*	0.023	-0.065
	(0.038)	(0.033)	(0.033)
Constant	−0.057	−0.021	−0.630*
	(0.122)	(0.081)	(0.085)
Observations	4,750	6,846	6,908
AIC	12,635.960	18,002.590	18,953.690
State Random Effects	✓	✓	✓
Municipality Random Effects	✓	✓	✓
Objective Measures	✓	✓	✓
Additional Controls	✓	✓	✓

Note: The above table presents the results of a hierarchical regression for my large cities sample. Respondents are nested within cities and then states. Co-partisanship was determined by whether the respondent shared the party of the acting Governor or Mayor in 2018. Standard errors are presented in the parentheses. *$p<0.05$

In the case of all three local services, co-partisanship with either the city's mayor or governor does not appear to have a significant effect on one's evaluations of local services. Even in the case of police, an individual's evaluations appear to be based far more on their partisan identity rather than their shared partisanship with local officials. While these null results don't necessarily mean shared partisanship has no effect on local evaluations, it appears for polarized services the party's average position matters far more. These null results may be at least partially explained by the lack of party competition often found within municipal governments (Bucchianeri 2020).

5 Conclusion

At the beginning of this article, I argued that individuals' evaluations of municipal services are not necessarily based on the quality of the service or their access to it. Instead, some local services have become polarized, leading to partisan distortions of evaluations in the direction of party preferences. Through an analysis of local police and schools, I provided strong correlational evidence that the polarized discourse surrounding police has led Republicans to evaluate the service more positively than other out partisans. Additionally, I showed how a similar process is not present in non-polarized cases of local schools and local roads.

My work adds to our understanding of retrospective evaluations within local politics in several ways. First, it highlights how increased nationalization and partisan sorting may inhibit retrospective voting within a local context. Second, the data used in my analysis departs from the traditionally geographically narrow approaches used to examine local politics and expands my analysis to a national sample. Lastly, it adds to our broader understanding of how polarization shapes electoral accountability. As individuals become increasingly sorted along party lines, my work suggests that retrospective evaluations may become less and less likely at all levels of government.

These findings suggest that partisanship may play a far more prominent role in the realm of local politics than previously conceived. For policymakers, this conclusion is potentially problematic as individuals may begin holding local services accountable for national opinions rather than their objective performance. This analysis of local evaluations also raises several vital questions surrounding local accountability and responsiveness. Notably, for two of the three examined services, objective measures of the said service failed to have a significant impact on an individual's evaluations. Future work should examine how individuals construct these perceptions of local services and which measures impact which group's evaluations. Additionally, while I've shown that partisanship may shape local assessments, it is unclear whether this translates into any policy or electoral change? Examining how local politics' nationalization and potential polarization affect municipal policy-making and spending should be an important focus of future work.

6 Appendix

6.1 Full Regression Results

Table 5: OLS Regression of Local School and Police Evaluations on Partisan Identity: Full Results

	Service Evaluation:	
	Police	School
Republican	0.264*	-0.035
	(0.018)	(0.028)
Property Crime Rate	0.00000	
	(0.00003)	
Violent Crime Rate	-0.0004*	
	(0.0001)	
School Rating		0.340*
		(0.014)
Age	0.007*	0.0002
	(0.0005)	(0.001)
Gender (M)	0.004	0.034*
	(0.010)	(0.011)
Parent	0.018	0.087*
	(0.012)	(0.015)
Education	0.013*	0.006
	(0.005)	(0.004)
Rural Res.	-0.175*	-0.028
	(0.018)	(0.017)
Homeowner	0.053*	0.032*
	(0.012)	(0.015)
Black	-0.282*	-0.016
	(0.033)	(0.035)
Hispanic	-0.038*	0.040
	(0.015)	(0.027)
Asian	0.060	0.118*
	(0.046)	(0.048)
Income	0.025*	0.016*
	(0.002)	(0.002)
State Fixed Effects	✓	✓
Observations	29,020	34,812
R^2	0.078	0.103
Adjusted R^2	0.076	0.101
F Statistic	187.880 (df = 13; 28958)	332.247 (df = 12; 34749)

Note: The above table presents the results of a OLS regression with state fixed effects and clustered standard errors. Crime rates are standardized to incident per 100,000 citizens. Standard errors are presented in parentheses.$^*p<0.05$

Table 6: Hierarchical Linear Regression of Local Roads Evaluation on Partisan Identity: Full Results

	Service Evaluation:
	Roads
Republican	0.030
	(0.026)
Traffic Index	−0.003
	(0.003)
Age	−0.005*
	(0.001)
Gender (M)	0.020
	(0.022)
Parent	0.049*
	(0.024)
Education	0.016
	(0.008)
Rural	−0.102
	(0.095)
Homeowner	0.075*
	(0.025)
Black	−0.175*
	(0.033)
Hispanic	−0.019
	(0.036)
Asian	−0.036
	(0.056)
Income	0.016*
	(0.004)
Constant	−0.003
	(0.077)
Observations	6,928
State Random Effects	✓
Municipality Random Effects	✓
AIC	18,191.980

Note: The above table presents the results of a hierarchical linear regression with random effects. All respondents are nested within municipalities within states. Standard errors are presented in parentheses.*$p<0.05$

6.2 Primary Results Including Independents

Table 7: OLS Regression of Local School and Police Evaluations on Partisan Identity: Full Results and Independents

	Service Evaluation:	
	Police	School
Republican	0.261*	-0.037
	(0.018)	(0.028)
Independent	-0.048*	-0.185*
	(0.018)	(0.021)
Property Crime Rate	-0.00000	
	(0.00003)	
Violent Crime Rate	-0.0004*	
	(0.0001)	
School Rating		0.341*
		(0.013)
Age	0.007*	0.0003
	(0.0004)	(0.001)
Gender (M)	-0.001	0.029*
	(0.009)	(0.011)
Parent	0.021	0.089*
	(0.011)	(0.014)
Education	0.012*	0.004
	(0.005)	(0.004)
Rural Res.	-0.165*	-0.020
	(0.017)	(0.016)
Homeowner	0.055*	0.034*
	(0.013)	(0.014)
Black	-0.289*	-0.022
	(0.033)	(0.032)
Hispanic	-0.046*	0.048
	(0.014)	(0.026)
Asian	0.073	0.115*
	(0.040)	(0.044)
Income	0.026*	0.017*
	(0.002)	(0.002)
Observations	32,994	39,548
R^2	0.078	0.106
Adjusted R^2	0.077	0.104

Note: The above table presents the results of a OLS regression with state fixed effects and clustered standard errors. Crime rates are standardized to incident per 100,000 citizens. Standard errors are presented in parentheses.$^*p<0.05$

6.3 Alternative Model Specifications

Table 8: Hierarchical Linear Regression of 7-Point Party ID on Local School and Police Evaluations

	Service Evaluation:	
	Police	School
Party ID	0.050*	−0.010*
	(0.002)	(0.002)
Crime Rate	−0.0001*	
	(0.00001)	
School Rating		0.321*
		(0.007)
Age	0.007*	0.0005
	(0.0003)	(0.0003)
Gender (M)	−0.006	0.023*
	(0.010)	(0.010)
Parent	0.013	0.086*
	(0.011)	(0.011)
Education	0.013*	0.004
	(0.004)	(0.003)
Income	0.025*	0.017*
	(0.002)	(0.002)
Rural	−0.179*	−0.045*
	(0.013)	(0.013)
Homeowner	0.051*	0.030*
	(0.012)	(0.011)
Black	−0.251*	0.002
	(0.020)	(0.018)
Hispanic	−0.032	0.057*
	(0.020)	(0.019)
Asian	0.069*	0.118*
	(0.032)	(0.030)
Constant	−0.254*	−0.682*
	(0.031)	(0.044)
State Fixed Effects	✓	✓
Municipality Fixed Effects	✓	✓
Observations	32,994	39,548
AIC	86,091.060	106,191.300

Note: The above table presents the results of a hierarchical linear regression with random effects. All respondents are nested within municipalities within states. Standard errors are presented in parentheses.*$p<0.05$

Table 9: Hierarchical Regression Results With Alternative Measures of Police Performance

	(1)	(2)
2017 Crime Rate	0.00001	
	(0.00003)	
2018 Crime Rate	-0.0001^{*}	
	(0.00004)	
Crime Rate Difference		0.00000
		(0.00003)
Constant	-0.138^{*}	-0.186^{*}
	(0.031)	(0.030)
State Fixed Effects	✓	✓
Municipality Fixed Effects	✓	✓
Additional Controls	✓	✓
Observations	32,115	32,115
AIC	83,722.270	83,730.040

Note: The above table presents the results of a hierarchical linear regression with random effects. Model (1) includes both current and lagged crime rates. Model (2) includes only the year change in crime rates. All respondents are nested within municipalities within states. Standard errors are presented in parentheses.$^{*}p<0.05$

6.4 Variable Distributions

Table 10: Descriptive Statistics of Independent and Dependent Variables: Full CCES

	Mean	Median	SD	Min	Max
Individual Covariates:					
Age	48.680	49	17.608	18	93
Gender (M)	0.445	0	0.497	0	1
Parent	0.658	1	0.475	0	1
Republican	0.377	0	0.485	0	1
Independent	0.139	0	0.346	0	1
Black	0.079	0	0.270	0	1
Hispanic	0.087	0	0.282	0	1
Asian	0.029	0	0.169	0	1
Homeowner	0.622	1	0.485	0	1
Income	6.412	6	3.306	1	16
Rural	0.150	0	0.357	0	1
Objective Measures:					
School Rating	2.651	2.605	0.826	0	5
Crime Rate (2018)	482.325	300.771	534.689	0	$3,063.617$
Crime Rate (2017)	576.387	355.356	658.426	0	$13,124.290$
Crime Rate Difference	-97.506	-27.720	259.851	$-2,013.473$	$1,195.273$
Service Evaluations:					
School	0.335	0	0.998	-2	2
Police	0.458	0	0.931	-2	2

Table 11: Difference in Covariate Means Between CCES and Large Cities Subset

	CCES	Large City Subset	Δ	P-Value
Demographics:				
Age	48.667	46.537	2.130	0
Gender (M)	0.451	0.459	-0.008	0.156
Parent	0.654	0.572	0.082	0
Black	0.091	0.145	-0.054	0
Hispanic	0.080	0.124	-0.044	0
Asian	0.028	0.041	-0.013	0
Homeowner	0.630	0.536	0.094	0
Rural Res.	0.191	0.030	0.160	0
Income	6.520	6.530	-0.010	0.786
Partisanship:				
Republican	0.385	0.279	0.106	0
Independent	0.122	0.111	0.011	0.001
Objective Measures:				
School Rating	2.628	2.271	0.357	0
Violent Crime Rate	82.940	69.789	13.151	0
Property Crime Rate	462.579	395.580	66.999	0

Note: The above table presents results of a two sample t-test between the 2018 CCES and the new merged data set. Covariates such as Republican, Independent, and Black are coded as a binary choice.Income (1-10) is measured on its own categorical scale.Crime rates are standardized to incidents per 100,000 citizens.