Public Information is an Incentive for Politicians: Experimental Evidence from Delhi Elections

Abhijit Banerjee, Nils Enevoldsen, Rohini Pande and Michael Walton*

March 30, 2021

Abstract

Two years prior to elections, roughly two-thirds of Delhi municipal councilors were informed that they had been randomly chosen for a pre-election newspaper report card. Relative to their control counterparts, treated councilors in high-slum-density areas increased pro-poor spending. Treated incumbents ineligible to rerun in their current ward because of randomly assigned gender quotas were substantially likelier to run elsewhere, but only if their report card showed a strong pro-poor spending record. Parties also benefited electorally from their councilor's high pro-poor spending. In contrast, in a cross-cut experiment, councilors did not react to actionable information that was not publicly disclosed.

1 Introduction

Can transparency help democracy work better for poor citizens? In this paper, we provide two pieces of evidence on this question based on an at-scale field experiment. First, anticipation of report cards about their performance being published in a newspaper made politicians more responsive to needs of lower-income voters. Second, their response was consistent with

*Authors affiliation (in alphabetic order): MIT (Banerjee), Yale (Pande), Harvard (Walton), and unaffiliated (Enevoldsen). Emails: banerjee@mit.edu, nils@wlonk.com, rohini.pande@yale.edu, Michael_Walton@hks.harvard.edu. For fieldwork and research assistance we thank Aditya Balasubramanian, Gaurav Chiplunkar, Sarika Gupta, Sandra Levy, Patrick Mayne, Cory Smith, Yashas Vaidya, Jeff Weaver, and Ariel Zucker. We thank Dominic Leggett, Fred Finan and seminar participants for comments, and USAID DIV (AID-OAA-G-12-00006), National Science Foundation (1063693), 3ie (OW2.099) and International Growth Center for financial support. AEA RCT Registry is AEARCTR-0004975.

how political parties and voters subsequently reacted to the disclosures. Both incumbents with a more pro-poor performance in public disclosures and their parties were electorally rewarded.

Our experiment focuses on elected city councilors in one of the world's most populous cities: Delhi. Each councilor represents a single ward, with several wards home to large slum populations (Jha et al. 2007). Councilors receive discretionary spending funds for slum infrastructure improvements. They also represent constituent interests in the legislature and in standing committees that oversee welfare schemes. In collaboration with one of India's largest media houses, we experimentally study the chain of responses when councilors are informed that a leading newspaper will report on their performance two years hence, just before the next city election.

We observe an incentive effect: the treatment better aligns councilor spending in high-slum-density wards with slum-dweller preferences.[1] Such an incentive effect makes sense only if disclosed performance is expected to influence electoral outcomes. We observe such an effect in the subsequent election: parties preferentially allocated tickets to rerun for election to incumbents with report cards demonstrating a good pro-poor spending record.

These results, in conjunction with evidence that absent newspaper disclosures slum dwellers lack information about councilor performance, support a central prediction of political agency models: transparency, by enhancing retrospective electoral accountability, incentivizes politicians to appear to perform well (Besley 2007). At baseline, councilors' spending decisions systematically diverged from the spending preferences of the poor. Only 16% of councilor funds were spent on sanitation, something 70% of respondents were concerned about. The majority of funds went to improve roads, which was not a priority for slum dwellers. At the same time, only a third of slum dwellers claimed to know how councilors spent their discretionary funds. The perception of councilor performance among slum dwellers was more negative than warranted. 68% believed that their councilor spent less money than other councilors and a majority underestimated their councilor's investment in sanitation, perhaps because these investments tend to be less visible.

Our experiment was carried out at scale, covering 240 of Delhi's 272 wards. In 2010, a random sample of 168 councilors were informed that a major newspaper would publish report cards on their performance using data from administrative records immediately before the 2012 election. At the same time, to establish credibility, the newspaper published midterm

[1] A ward has high slum-density if its fraction of slum area exceeds the median.

report cards for a random subset of treated councilors. These midterm report cards could have directly influenced councilors' choices, even absent an incentive response. For example, they might have encouraged councilors to focus more on indicators the report cards make salient. To directly test whether councilors react to information in a report card when it is not tied to public disclosure, we implemented a cross-cutting experiment in a subset of high-slum-density wards. In the year prior to the election, treated councilors received two State of Sanitation Information (SSI) report cards that were not made public. These were based on toilet and garbage point audits in slums, and provided information about problems that councilors could act upon prior to the election.

Our electoral analysis utilizes two additional sources of exogenous variation. First, the fraction of wards reserved for female candidates was unexpectedly increased just before the election. Random assignment of ward reservation status created exogenous variation in the set of (mostly male) incumbents ineligible to contest the election in their current wards (henceforth "ineligible incumbents"). An ineligible incumbent's party had to decide whether to allocate him a "ticket" as the party candidate for a different (unreserved) ward. Party ticket allocation occurred after most report cards had been published, and parties could have reacted to them. Second, we randomized the sets of four councilors featured in a single newspaper edition, creating, for each councilor, variation in the strength of the other three report cards appearing in the same edition.

Turning to the results, anticipation of public disclosures led to 0.65 standard deviations more pro-poor spending allocations by treated councilors in high-slum-density wards. (We construct pro-poor weights for different spending categories using slum-dweller preferences from our baseline surveys.) Paralleling this, pro-poor spending *fell* by a noisily estimated 0.32 standard deviations in low-slum-density treatment wards. Councilors' legislative attendance also improved by 0.29 standard deviations in treated high-slum-density wards, and fell by 0.17 standard deviations in treated low-slum-density wards, though neither with statistical significance. Estimated effects were similar across councilors only promised end-of-term disclosure in 2012 and those who, in addition, received a midterm report card in 2010. Finally, the private SSI treatment did not improve sanitation measures, indicating the importance of public disclosures.

The aforementioned expansion of gender quotas rendered a third of incumbents ineligible for re-election in their current wards. An ineligible treated incumbent was 16 percentage points more likely to receive a party ticket for another ward relative to his ineligible control counterpart, for whom the mean is zero. The effect on eligible incumbents running in any

ward is also positive, but not significant. The choice of who gets to rerun in treatment is linked to performance: relative to his control counterpart, a one standard deviation improvement in an ineligible treated incumbent's pro-poor spending record increased his probability of getting a party ticket by 13 percentage points. High-performing ineligible councilors who received a party ticket were more likely to come from, and be moved to, high-slum-density wards. A treated incumbent's vote share (setting vote share of those who did not run to zero) increased by roughly the same magnitude as the probability of running in a different ward, suggesting that displaced treated incumbents received electoral support. The party also benefited: it received higher vote shares in treatment wards where high-performing incumbents were rendered ineligible. Given that the number of ineligible councilors who received a ticket was small, we confirm that our results are robust to a randomization inference procedure. Finally, consistent with yardstick competition, the party of a councilor with a better pro-poor performance *rank* among the four councilors featured in the same newspaper edition saw electoral gains in that ward, controlling for the councilor's *absolute* performance.

A report card prepared by a reputable NGO and published in a leading newspaper provided councilors an opportunity to publicize performance in a credible way to poorly informed voters and, in doing so, increased councilors' incentives to do what these voters want. Councilors may have also anticipated that report cards would influence party ticket allocation. Our finding that parties favor incumbents with report cards showing good performance is consistent with parties believing that voters will reward reported good performance. Anecdotal evidence suggests that parties use report cards to advertise both high-performing candidates and a general commitment to pro-poor policies.[2] This advertisement might be especially important when the candidate is new to the ward. Furthermore, since pro-poor spending matters most to poor voters, it makes sense that the party redirects ineligible high-performers to high-slum-density wards.

On yardstick competition, relative rank in the newspaper contains additional information whenever incumbents face a common shock, e.g. the relative ease of using councilor funds to improve sanitation versus roads (Lazear and Rosen 1981; Green and Stokey 1983).[3] It can also explain incumbents' response to anticipated disclosures. Since the report cards didn't mention that councilors had been previously informed, it is plausible that voters did not

[2] A video of a party using a report card in a pre-election rally, from a different part of India, is available from the authors.

[3] It could be that permits for construction of different kinds in the city may be easier or harder to get.

know that incumbents were reacting to the report card. Hence, incumbents get credit on election day for a better report card.

Our assumption that voters favor high-performing incumbents with report cards over comparable incumbents without report cards is likely valid for voters outside the incumbent's current ward, who lack information about the incumbent beyond his report card. However, as we show using a simple theoretical example, this assumption may not hold for an incumbent rerunning in the current ward. For instance, voters with very high expectations may be disappointed even when the incumbent performed relatively well. The party may, in this case, be more likely to give the ticket to an outside incumbent with no report card. The evidence on yardstick competition offers a novel way of addressing this concern. We compare the party's electoral outcomes across incumbents with similar spending performance who featured in different newspaper editions but, as a result of random variation in who else featured in the same edition, were ranked differently. Since we control for the incumbent spending record, it is reasonable to assume that voters have similar priors about them. Moreover it is implausible that they had prior information about the rank beyond that contained in the incumbent's spending record. Indeed, we show that voters favor the party with a higher ranked incumbent after controlling for her spending record.[4]

In sum, multiple pieces of evidence point in the same direction: that anticipation of performance disclosures incentivizes councilors to respond to the needs of poor voters in high-slum-density areas, and that they are justified in doing so, since parties and voters favor better performers.

Multiple papers demonstrate that credible media reports on incumbent performance and corruption impact voters' choices (Banerjee et al. 2011; Ferraz and Finan 2008; Larreguy et al. 2020). The evidence on pre-election voter information campaigns (PEVACs) via flyers or video screenings is more mixed (Chong et al. 2015; Adida et al. 2020; Boas et al. 2018; de Figueiredo et al. 2012; Dunning et al. 2019), with some evidence that treatment effects of PEVACs vary with voters' prior beliefs.[5] Relevant for us, Kendall et al. (2015); Arias et al. (2018) argue that voters with sufficiently negative priors may be positively surprised by information of less corruption than expected leading to an incumbency advantage.[6]

[4]Our use of the evidence on yardstick competition may be novel, but the idea that voters compare own incumbent with others is not. See, for example, Besley and Case (2003), for the case of US governors.

[5]Chong et al. (2015) show that disclosures can discourage voter turnout. That said, we have evidence that well-targeted non-media sources can matter: Kendall et al. (2015) show informative phone calls from incumbent's campaign influenced vote shares in an Italian mayoral election, while Buntaine et al. (2018) show that text messages about incumbent's corruption reduces his vote share.

[6]The impacts of soft information – for instance, candidate debates – are even more sensitive to what

Our results also relate to the political economy literature on the media, where observational studies report a positive link between media presence and quality of policies, typically at a macro level (Besley and Burgess 2002; Strömberg 2004; Snyder Jr and Strömberg 2010; Enikolopov et al. 2011). More closely related are papers studying the impact of a specific information revelation mechanisms, and in particular, three papers on the long term effects of anticipated disclosures.[7] Bobonis et al. (2016) in Puerto Rico and Avis et al. (2018) in Brazil show that anticipated pre-election publication of corruption audits of mayors reduces corruption at least temporarily. However, corruption differs from the kind of performance information that is our focus. First, accusations of corruption are only credible when there is an expensive and well-institutionalized process behind them (Avis et al. 2018). Second, corruption tends to be big news when revealed. Third, investigations of corruption often open the possibility of prosecution, which may partially account for the observed effect on corruption. Voters may worry about their mayor ending up in court or prison, paralyzing subsequent performance. In contrast, the spending information we consider comes from routinely collected administrative data, obtained under India's Right to Information Act. There are no judicial penalties associated with councilors' spending choices. Closely related is Grossman and Michelitch (2018), who evaluate the impact of annual performance score cards for politicians prepared by an Ugandan CSO and disseminated through public meetings. A key difference, other than the use of mass media to enable an at-scale intervention, is that the CSO provided *summary politician evaluations*. This is more directive (and politically sensitive) than our *modus operandi* of reporting the spending categories.

Our other innovation is to connect the effect of report cards on politician performance to their impact on candidate selection by parties and voters. As better report cards are electorally rewarded, politicians are justified in trying to improve them.[8]

To the best of our knowledge this paper provides the first at-scale experimental evidence that newspaper report cards move both the performance and the selection of rerunning incumbents in a pro-poor direction.[9] The report cards are based, in principle, on publicly available administrative data. Similar interventions are possible elsewhere since discretionary constituency funds, like the one we study, are increasingly common in developing countries (International Budget Partnership 2010) and the spread of freedom of information acts

voters infer from it (Bidwell et al. 2020; Wantchekon 2003; Fujiwara and Wantchekon 2013).

[7] See Bidwell et al. (2020), for instance, on short-term incentive impacts of anticipated disclosures.

[8] The idea that gender quotas accentuate competition between male candidates for party tickets and leads to the selection of high male performers echoes Besley et al. (2017) results for Sweden.

[9] Both statements are relative to the control group, where impacts may be in the opposite direction.

makes public disclosures of spending information increasingly feasible (Mendel 2014).

The rest of the paper proceeds as follows. Section 2 describes context and experimental design. Section 3 presents empirical strategy and data. Sections 4 and 5 discuss treatment impacts on politician behavior and electoral outcomes, respectively. Section 6 concludes.

2 Setting and experimental design

2.1 Context

With a population of over 18 million, Delhi ranks high among the world's mega-cities and, as in most mega-cities, inequality is built into the human geography: between a quarter and a half of Delhi's population lives in slums spread across the city (Delhi Human Development Report 2006; Jha et al. 2007).[10] Our focus is on Delhi's municipal government; every five years, each of the 272 city wards elects a councilor via first-past-the-post elections.

Ethnographic and political science literatures highlight the importance of the slum-dweller vote in lower-income countries (Post 2018). Political participation is high among slum dwellers and they often vote en bloc. Furthermore, the uncertain legal status of the slums – which are often built on encroached land and for that reason often denied recognition – makes them particularly reliant on politicians (Holland 2016). In our baseline surveys, 86% of slum dwellers reported being registered to vote and, of those registered, 83% reported having voted in the last city election. 30% reported participating in a political rally or march prior to that election. They also reported significant dependence on public facilities: for instance, 76% of those who reported that a public toilet is available to them reported using one, while complaining about their cleanliness. Finally, they often don't know much about how the system works: for instance, only 30% of slum dwellers reported knowledge of the discretionary funds controlled by the councilors.

City elections are contested by some of the major national parties, with candidates selected by party leadership. (There are no party primaries.) Our experiments were carried out in the context of 2012 election. Four months prior to these elections, the fraction of wards reserved for women candidates was unexpectedly increased from 33% to 50%. The assignment of reservation status was random.[11]

[10] The difference in estimates reflect whether slum definition only includes legally notified slums or also non-notified slums, typically illegal encroachment on land (DUSIB 2010).

[11] There is also non-random reservation for scheduled castes (SC). In 2012, wards were listed by serial number with every odd seat reserved for a woman (separately for SC reserved and SC unreserved categories).

2.2 Policy levers and citizen preferences

Elected councilors have access to multiple policy levers. First, they vote on policies in the city legislature; in the five years between 2007 and 2012, the city legislature met, on average, 24 times per year. They are also assigned to standing committees which are intended to meet regularly and respond to citizen grievances.

Second, councilors receive an annual discretionary fund which averaged 200,000 USD per year during the 2007–2012 term. Councilors allocate these funds for local infrastructure. Public goods particularly relevant for slums include public toilets, garbage removal, sewers, and drains, all of which fall into the broad category of sanitation. Toilets and garbage removal are managed on a contractual basis by private organizations.[12] Councilors help select contractors, and can pressure or sanction them for non-delivery.

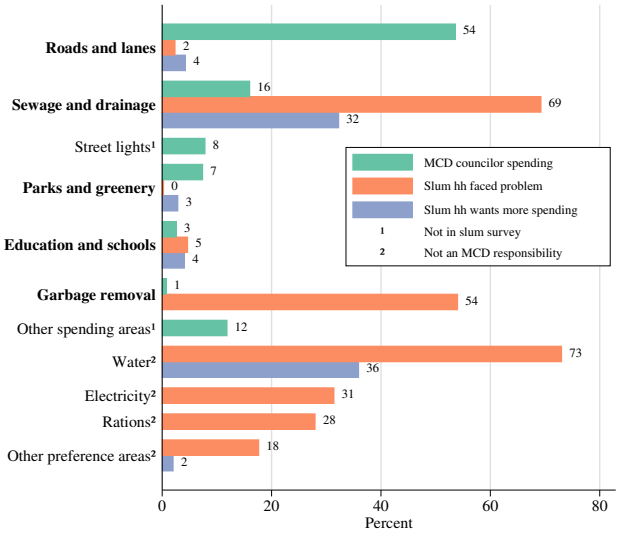
Reflecting our focus on councilor accountability to Delhi's poor, our baseline data collection focused on 107 high-slum-density wards defined using aerial data on slum incidence.[13] In April–May 2010 we surveyed 5,481 slum households in these wards; the average respondent was a long-term migrant who had lived in Delhi for 17 years. The survey collected data on slum dwellers' access to, usage of, and difficulties with various public goods and social services, as well as knowledge of local government, interactions with public officials, political preferences, and participation in politics.

Figure 1 reports on the slum dwellers' concerns with infrastructure and compares them with how councilors spend their funds. Councilors spend a majority of their funds on roads, which are a visible form of investment but also one which has significant scope for corruption in contracting (Lehne et al. 2018). Slum dwellers neither report problems with road infrastructure nor favor more road spending. They focus on sanitation – lack of clean water, sewage issues, and irregularities in garbage collection – as the most problematic set of issues. While water in Delhi is privately provided and is beyond the remit of the councilors, drains and toilets are not. Yet spending on sanitation is only 16% compared to 54% on roads. Other large spending categories that do not align with slum dwellers' priorities include provision and repair of lights (8% of spending) and improvement of parks and greenery (7% of spending). In particular, slum dwellers express no interest in spending

[12] The typical toilet contract is usually 20–30 years and sets maximum user price, states necessary facilities and requires regular cleaning. The typical garbage contract is for nine years and stipulates that operators provide garbage bins, and segregate and collect waste daily (IL & FS Ecosmart Limited 2006).

[13] On average, high-slum-density wards had 3.1 slums while low-slum-density wards had 0.7 slums; see Appendix A.1.

Figure 1: Councilor spending and constituent preferences



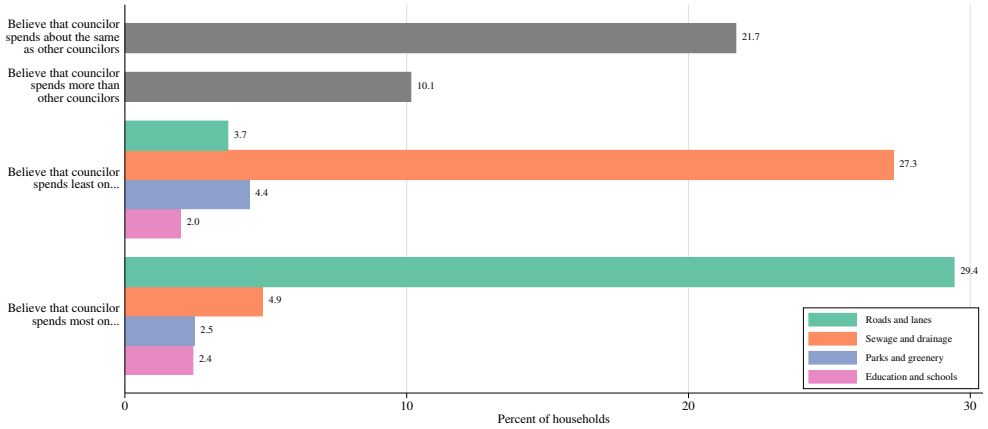
Reported ward means for: "Councilor spending" is annual fraction of councilor spending (2007–2012). "Slum household faced problem" is fraction of slum households who specify category in response to, "In which of following areas have you personally faced problems in the last year?" ("Other preference areas" comprises health, law and order, and pensions.) "Slum household wants more spending" is fraction who specify category in response to, "What type of project would you like your councilor to spend more on?" ("Other preference areas" comprises health.) Bold face indicates areas included in preference-weighted spending measures.

on parks and greenery, which are typically located outside slums.

Figure 2 examines slum-dweller knowledge about spending patterns in their ward. Over a quarter believe councilors spend the most on roads, and over a quarter believe councilors spend the least on sanitation. In reality, councilors spend twice as much on sanitation as on parks and greenery. Constituents are pessimistic about their councilors: of the 81% of slum dwellers who answered a question on their councilor's relative performance, 68% thought he spent less money than other councilors (11% thought more, and 22% thought about the same), even though nearly all councilors spend nearly all their funds. Slum-dweller beliefs on councilor spending by category – both in absolute terms and relative to other councilors – is uncorrelated with actual spending.

While we did not conduct household surveys outside of slums, in 94 survey wards we surveyed Resident Welfare Association (RWA) office-holders. RWAs represent residents of legally recognized housing colonies in the ward, who are typically richer than slum dwellers.

Figure 2: Constituent beliefs and preferences about councilor spending



Each colored bar represents ward-mean of slum households who specify each category in response to, "What type of project do you think your councilor spends most (least) on?" We show only the categories included in preference-weighted spending measures.

Relative to slum dwellers, RWA officeholders are ten times more likely to name roads as an important problem, put less (but still substantial) emphasis on sewage, drainage, and garbage disposal, and are keen on parks and street lights, two items that slum dwellers never name. Thus, policy preferences differ across these two groups, though arguably there is too much road spending even relative to RWA preferences.

2.3 Experimental Design

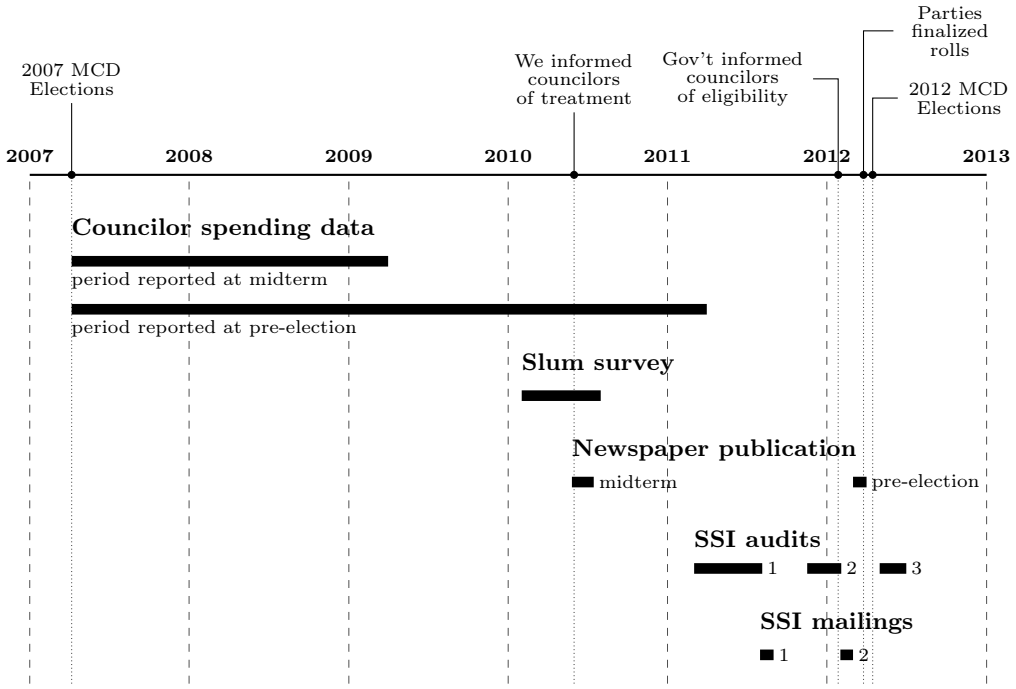
Our study uses two cross-cutting experiments: one that provided public information in the form of newspaper report cards, and one that provided private information to councilors in the form of sanitation facility audits. Figure 3 summarizes the different aspects of our study and their timelines.

A. Newspaper report card experiment

Our sample consists of 240 wards, with treatment assignment stratified by incumbent party and geographic zone.[14] We randomly assigned 72 wards to be control, 58 to only receive

[14] We excluded 32 wards: 5 where our partner NGO was doing community work, 10 where by-elections occurred less than two years ago, and two zones (comprising 17 wards) with few or no slums.

Figure 3: Experiment timeline



pre-election report cards (T1) and 110 to receive both midterm and pre-election report cards (T2). Appendix Figure A.1 shows a map of wards by treatment status.

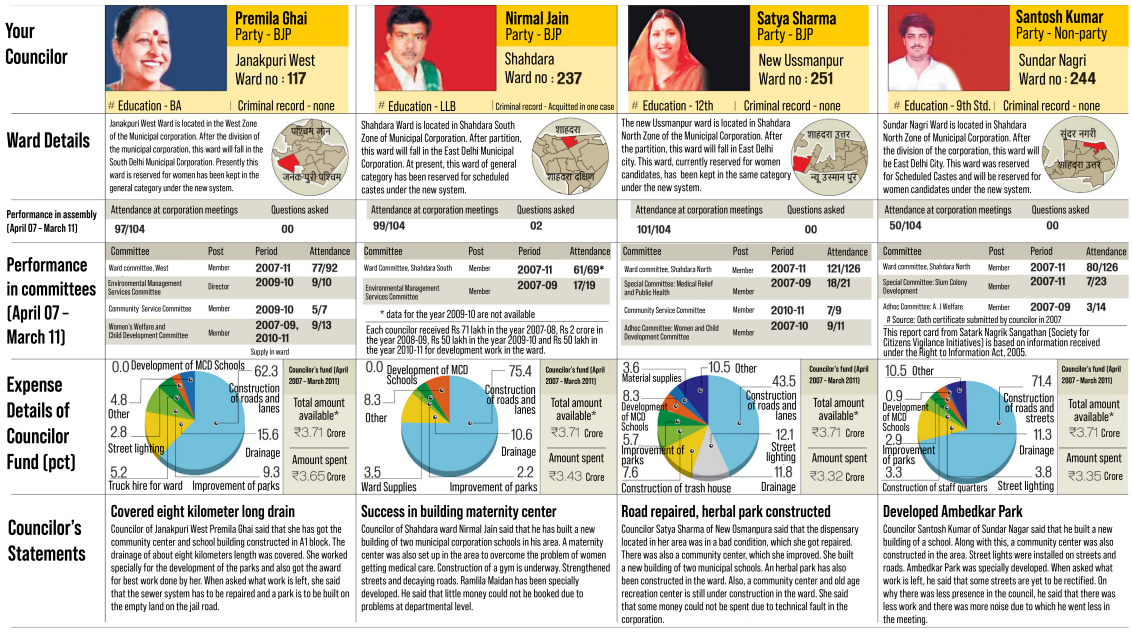
Using the 2005 Right to Information Act (RTI), our NGO partner, Satark Nagrik Sangathan (SNS) obtained data on councilor spending, meeting attendance, and committee membership.[15] Figure 4 depicts one set of four report cards published in 2012 by our media partner: the Hindi daily newspaper *Dainik Hindustan*, which, at the time of our experiment, was Delhi's second largest newspaper by market share. For each councilor, the report card includes her photo, committee and assembly attendance, and category-wise discretionary fund spending.[16]

Midterm report cards were published over the course of a month starting May 27, 2010, with performance data spanning April 2007 to March 2009. Pre-election report cards were published over the course of a month starting March 1, 2012, with performance data spanning

[15] SNS (English name: Society for Citizens Vigilance Initiatives) has a history of creating and disseminating, in partnership with media outlets, report cards on incumbent performance and candidate qualifications for state and national legislatures. Banerjee et al. (2011) found their 2010 Delhi state election campaign increased the vote share of better performing candidates.

[16] It also reported incumbent qualifications and an incumbent statement on an infrastructure success.

Figure 4: Example councilor report cards in *Dainik Hindustan*, 2012 (translated)



April 2007 to March 2011.[17]

Informing councilors In June 2010, councilors in both treatment arms received a letter from SNS, with an appended newspaper page featuring the first midterm report card. Each letter described SNS's mandate for using the RTI Act to create report cards on elected representative performance and to disseminate them via the media, listing previous report card campaigns and media partners. The letter then stated,

> SNS has prepared midterm report cards of MCD councilors to inform people of the development work being done by councilors for the welfare of ward residents. Due to limited resources, we have not been able to prepare report cards for all 272 councilors in Delhi. We have randomly selected 110 wards for which report cards are being prepared. The party-wise break up of the sample is the same as in the MCD. As you might be aware, these report cards are being published by the Hindi newspaper *Hindustan*. (Please find attached the report cards that appeared on May 27, 2010 in *Hindustan*.)

The T1 letter then went on to state,

[17] As data is released by fiscal year, we could not report spending data for FY 2011–12.

> Your ward is not in the list of 110 wards for which report cards are being prepared this time. However, in 2012, we will again be preparing report cards for these and more wards in Delhi. In 2012, we intend to include your ward.

The T2 letter instead stated,

> Your ward is one of the 110 wards for which the report card has been prepared. In 2012, we will again be preparing report cards for these 110 wards and for more wards in Delhi. We hope that dissemination of these report cards based on objective information will help people understand the development efforts being made by councilors for the welfare and betterment of their wards.

The councilors in control wards received a letter that they were not selected for report cards and would not be reported on until at least 2014, two years after the election. All treatment councilors received a reminder letter in February 2011.

Potential for yard-stick competition We began publishing pre-election report cards six weeks prior to the election. Each daily edition featured four councilors, chosen through a two-step process.[18] First, wards were categorized into comparable groups (by zone, incumbent eligibility, and slum-density) and within each group wards were randomly paired. Two pairs were randomly chosen to feature on the same page to evaluate yardstick competition, i.e. whether, controlling for absolute performance, parties and voters compare a councilor's performance relative to his peers featured in the same edition of the newspaper.

B. Private State of Sanitation Information (SSI) experiment

We conducted a cross-cutting experiment to study the impact of providing a councilor with actionable private information on sanitation problems in slums. A JPAL South Asia team audited sanitation facilities in, on average, three slums in 108 high-slum-density wards (for a total of 310 slums).[19] We randomized wards into treatment or control status.[20] Treated councilors received *private* report cards on the state of public toilets, formal garbage dumps, and informal garbage piles in slums in their wards. This information was not released to the public.

[18] Appendix A.2 has details.
[19] Surveying errors led to audit and survey samples differing in the selection of four wards.
[20] Appendix A.3 has details.

Informing councilors Three SSI audit rounds were conducted, with reports from the first two rounds mailed to treatment councilors. The first round summarized SSI audits between April and June 2011 and was sent in August 2011. The second round, summarized audits between November 2011 and January 2012 and was sent in February 2012, two months prior to elections. The final SSI audits occurred immediately after elections, from April through June 2012.

Using data from our baseline surveys, the cover letter highlighted what fraction of households reported public toilets, sanitation, and garbage to be a major problem. It then explained that the attached audit report card provided a summary of the status of audited toilet and garbage facilities, and then their details. It also referenced the enclosed map that highlighted all facilities that needed attention.

Information on the state of drains was collected, but not shared with politicians, during the second and third rounds of SSI audits. This was to pick up potential spillover effects, which could be positive if, for example, the audits made the councilors more aware of the problems in those slums, or negative if there was a diversion of effort away from drains.

3 Data and Empirical Strategy

3.1 Data

Our experiments and subsequent analysis draw upon multiple sources of data.

SNS collated councilor performance data for all wards for the period 2007–12; only treatment ward data was subsequently released as newspaper report cards. 2007–10 spending allocations occurred prior to our intervention. Hence, to examine incentive effects on councilor spending we use 2010–12 data to construct the relevant outcome measure. The 2012 report cards were based on spending data for 2007–11 (as the 2011–12 fiscal year hadn't yet ended) and so our electoral analysis uses performance measures based on 2007–11 data. Throughout, we use a five category spending classification: Roads and lanes, Sewage and drainage, Parks and greenery, Education and schools, and Garbage removal.[21]

Our *pro-poor spending* index combines these categorized councilor spending data with the survey data on slum dwellers' infrastructure preferences, described in Section 2.2. We create three separate indices, each of which weights spending on the different categories

[21] The omitted category contained MCD office/staff-quarter construction and repairs, construction of tenements, community centers, health centers, boundary walls, monument gates, and street name boards.

with a different preference measure: fraction of slum households in the city reporting that the issue is (1) the biggest problem in their area, (2) a problem for them, or (3) a problem for the community. For each weighting criteria, we calculate the log of the sum of the weighted spending amounts. (Wherever the value was zero it was bottom-coded to the first percentile of the spending distribution prior to taking logs.) The mean z-score of the three log weighted spending measures serves as our pro-poor spending index.

As a robustness check, we construct a *directly-elicited spending* index which includes the same spending categories – except garbage collection – weighted by fraction of slum households reporting that their councilor should spend more on it minus the fraction reporting that they should spend less on it.[22] The z-score of the inverse hyperbolic sine of the sum is our directly-elicited spending index.[23]

The report card also reported the councilor's attendance in legislature and committees, which averaged 81% and 66%, respectively. Our attendance index is the mean z-score of attendance in: (1) the legislature and (2) the committees of which the councilor is a member. The pro-poor spending index and attendance index are positively correlated with $\rho = 0.10$.

Finally, based on report cards published in the same edition we allocate each report card a rank between 0–3 depending on the incumbent's spending performance relative to the three others featured in the same newspaper edition. This rank is based on the pro-poor spending index variable which was not explicitly featured in the report card.

All three sanitation facility audits covered public toilets and garbage points, and the latter two also covered drains. During a facility audit the surveyors noted its quality and interviewed two respondents for data on frequency of cleaning and prices. For toilet usage, the surveyor counted how many people used the toilet in a randomly chosen 15 minute interval between 3–5 PM. We construct a slum-level *toilet availability* index as the mean z-score of three elements: log total public toilets, log open public toilets, and log adult public toilet users. We separately examine toilet price. For garbage, we counted the total number of formal and informal piles, and construct a *garbage collection* index which averages the fraction of regularly cleaned piles across formal and informal points.

To measure electoral outcomes, we use ward-level 2012 electoral data: turnout, candidate list, and candidate-wise vote share. We include a dummy for each incumbent councilor that

[22] Respondents were asked on which single project their councilor should spend less, and on which more, and prompted with categories: roads, water, sanitation, health, education, parks, and community centers. We identify sanitation with sewage and drainage.

[23] We use inverse hyperbolic sine instead of log since, unlike pro-poor spending index, the operand can be negative.

equals one if the quota rendered him ineligible to recontest in his current ward in 2012. One third (80 of 240) of our incumbent councilors were ineligible. The fraction of incumbents who did not rerun exceeds the fraction affected by quotas: Only two fifths (95 of 240) of incumbent councilors contested an election from any ward, including just over half (87 of 160) of those who were eligible to rerun in the current ward.

At the councilor level, we consider two related electoral outcomes: whether a councilor gets a party ticket (for the current ward or for a different ward) and her vote share. Given that the choice to rerun is endogenous, we set the vote share of non-running incumbents to zero, rather than exclude them. (In this we follow Ferraz and Finan (2008).) At the party level, we consider the vote share of the incumbent councilor's party.

3.2 Analysis plan

Our experiment and data collection exercises sought to test multiple related hypotheses. First, if councilors believe newspaper report cards inform voters, or make known information salient, then treatment should incentivize better councilor performance. In high-slum-density wards, we hypothesize that pro-poor spending by councilors will increase. In contrast, private disclosures to councilors should improve performance only if councilors lack information about public good quality and are either intrinsically motivated or believe that their party or a sufficient number of citizens will learn about improvements they make.

Second, public disclosures should incentivize councilors only if voters and, therefore, parties, favor those incumbents the report cards designate as better performers.

Third, a well-performing incumbent's party should also benefit from her performance.

Finally, based on the predictions of political agency models, we examine heterogeneous treatment effects by councilor performance.

One concern is that performance in the final year of our published data was influenced by treatment. This could potentially generate spurious evidence of an interaction effect between treatment and performance if the main effect of performance on electoral outcomes is non-linear yet we estimate a linear model.[24] In the appendix we estimate an alternative specification where we drop the final year of our performance data. The drawback is that this is no longer the measure of performance that the voters see in the report card, though

[24] If performance is better in treatment wards and, say, performance only impacts electoral outcomes at high levels of performance, then the slope of electoral outcomes with respect to performance will be higher in treatment. If we nevertheless estimate a model where performance enters linearly, we would find a positive interaction between treatment and performance, even if there is no such interaction.

the two measures have a correlation of $\rho = 0.92$. Reassuringly, the two specifications deliver very similar results.

We also estimate an alternative specification where we only compare across wards whose incumbents featured in published report cards, exploiting the variation coming from the random choice of report cards to published together in the newspaper. We ask how being ranked higher among those published in the same edition of the newspaper affects party fortunes, *after controlling for our measure of performance.*

Our experiment was launched in 2010, prior to creation of the AEA RCT registry. The experimental design and baseline data collection capture our interest in whether future disclosures strengthen retrospective electoral accountability towards poor citizens, as described in our subsequent trial registration. We report three types of heterogeneity, which were not explicitly accounted for as stratification variables at the time of our randomization.

First, our restriction of baseline data collection on spending preferences to slum dwellers drives the way we estimate the politician response to treatment. We hypothesize that, in the presence of spending preference differences across slum dwellers and richer households, treatment impacts on pro-spoor spending will be larger in high-slum-density wards.

Second, an exogenous source of variation – not anticipated at the time of experimental design – relates to councilor eligibility status. As the party ticket allocation decision is qualitatively different when choosing whether to move an ineligible councilor to a new ward, we consider the electoral impacts separately for eligible and ineligible councilors.[25]

Finally, as discussed above, we examine whether electoral gains, if any, associated with treatment are concentrated among better performing incumbents.

3.3 Implementation and Balance Checks

In late May and June 2010, midterm report cards were published. Actual treatment was close to ITT: every one of the 110 councilors were reported on, excepting one councilor who had died. In June 2010, letters were sent out to councilors informing them of treatment status.

In March 2012, pre-election report cards were published. Newspaper space constraints resulted in actual treatment being significantly lower than ITT. Across ITT councilors, we

[25]The vote share effect for councilors combines their probability of rerunning with the actual votes they get if they rerun. Eligible incumbents are much more likely to rerun than ineligibles, and moreover information is likely to play a very different role where the councilor is already an incumbent and therefore known, than where he is not.

randomized report card publication order and prioritized report cards for eligible councilors. Of 168 report cards for ITT councilors, 124 were published: 45 of 58 T1 wards and 79 of 110 T2 wards. This covered all eligible ITT councilors and half of ineligible ITT councilors.[26] Throughout, we estimate ITT effects.

Appendix Table A.I contains balance checks. Panels A and B consider pre-treatment councilor spending outcomes (the three fiscal years between 2007–10) for all wards and for slum-survey wards, respectively. Panel C considers slum households' self-reports on problem areas over the last year. We observe no differences by ward treatment status. Appendix Table A.II considers 2007 electoral outcomes as the dependent variable. We observe no treatment differences for number of registered voters, turnout, seat reservation, or number of candidates, but a slightly higher 2007 vote share among eligible ITT councilors (significant at 10%).

4 Does information influence councilor behavior?

In this section we focus on whether the promise of public report cards just before elections influences the performance of incumbent councilors. We compare this with the effect of report cards that are privately given to the same councilors.

4.1 Anticipated public disclosures: Newspaper report cards

A. Performance outcomes and empirical specification

Based on the report card, we examine three councilor performance metrics. First, total discretionary fund spending which, arguably, all voters want maximized. Second, a meeting attendance index (averaged over legislature and committee attendance), also an unambiguous good, unless citizens favor councilor actions unobserved to us and there is a time budget constraint. Third, composition of councilor infrastructure spending. As we expect the preferred composition to vary across neighborhoods and our preference data is restricted to slum dwellers, we evaluate a narrower hypothesis: in high-slum-density wards, treated councilors should better align spending decisions with the public goods preferences of low-income voters. Newspaper report cards provide voters with information about overall public goods spending patterns, whereas otherwise they might only know what is happening in certain neighborhoods, and therefore incentivize the councilor to deliver what voters want.

[26] Appendix A.2 has details.

We measure the extent of alignment by the pro-poor spending index in a ward, and check results are robust to an alternative set of direct-elicited preference weights.

For all performance metrics, we report ITT estimates using ANCOVA specifications, which allow smaller changes in ward w with high baseline investments in that outcome:

$$y_{w1} = \alpha y_{w0} + \beta Treat_w + X_w + \epsilon_w, \tag{1}$$

$$y_{w1} = \alpha y_{w0} + \beta Treat_w + \gamma High_w + \lambda Treat_w \times High_w + X_w + \epsilon_w, \tag{2}$$

$$\begin{aligned} y_{w1} = \alpha y_{w0} + \beta Treat_w + \gamma High_w + \delta Inelig_w + \lambda Treat_w \times High_w + \mu Treat_w \times Inelig_w \\ + \nu High_w \times Inelig_w + \xi Treat_w \times High_w \times Inelig_w + X_w + \epsilon_w, \end{aligned} \tag{3}$$

$Treat_w$ is ward treatment status (pooling across T1 and T2 treatment arms), $High_w$ is a high-slum-density ward indicator, $Inelig_w$ is whether the councilor is ineligible for reelection in that ward, and X_w is a stratum fixed effect (zone–party combination). The councilor performance outcomes y_{w1} are measured for post-treatment fiscal years 2010–11 and 11–12. The lagged dependent variable, y_{w0} represents the average of the same variables for pre-treatment fiscal years 2007–08, 08–09, and 09–10. The specification in equation (1) reports average treatment effects, while equation (2) allows for heterogeneous treatment effects across high- and low-slum-density wards. Equation (3) controls, in addition, for councilor ineligibility.

B. Results

Table I presents our main specifications. Councilors typically spent their entire budget. Reflecting this, columns (1)–(3) show that anticipation of report cards has no effect on log total spending, either on average or in high-slum-density wards. Councilor ineligibility to stand for reelection from current ward in 2012 elections also doesn't affect overall spending.

Columns (4)–(6) focus on our primary incentive outcome: the pro-poor spending index. Column (4) shows no overall impact on spending. However, column (5) shows a large *negative* (but noisy) impact on pro-poor spending in low-slum-density wards, and relative to these treatment wards, a large positive impact on high-slum-density treatment wards (higher by 0.66 standard deviations, $p = 0.03$). Column (6) shows councilor eligibility has no impact on nature of spending, in high- or low-slum-density wards. In columns (7)–(9) we

observe similar patterns when we use the alternative directly-elicited spending measure – where weights reflect directly-elicited citizen preferences on councilor spending.

Appendix Table A.V examines individual spending categories as spending outcomes. Treated councilors in high-slum-density wards spent more on drainage and sewage – column (2) – and less on schools – column (4).

Columns (10)–(12) report a similar, but noisier, pattern of impacts of report cards on attendance – the treatment effect on attendance in high-slum-density wards exceeds that in low-slum-density wards. Future ineligibility doesn't influence councilor spending in high-slum-density wards. Attendance of ineligible councilors from low-slum-density wards is negatively impacted by treatment, and that of ineligible councilors from high-slum-density wards is positively impacted.[27]

As a robustness check, Appendix Table A.VI shows that results are similar when we consider an alternative definition of high-slum-density wards: whether the ward was in our household survey sample. Recall that a selection criterion for survey wards was (verified) slum density.

Appendix Table A.VII separates T1 and T2 treatment arms. In both arms, councilors received a letter in 2010 informing them that they would be reported on prior to the election but in T2 a midterm report card was also published when the letter was sent. We expect incentive effects in both arms, but T2 might be associated with stronger incentive effects. First, the midterm report card may make the process more credible. Second, by raising salience of some issues or providing new information midterm report cards may create a councilor information effect. Finally, they might have spurred citizen activism.

We cannot reject identical impacts on spending and attendance indices across the two arms, but T2 effects are more precisely estimated due to a larger sample. Of course, this could be because midterm reports in T2 were also read by T1 councilors and citizens.[28] To further explore the *councilor information* channel, we consider the SSI experiment which disclosed service quality in slums *just to the councilor*.

[27] Appendix Table A.IV columns (3), (4), and (5) shows findings are robust across the three components of the pro-poor spending index. Columns (9) and (10) show similar patterns across attendance index components. Differences in attendance effects across low- and high-slum-density neighborhoods potentially reflects greater middle-class cynicism about what politicians do, and in particular the importance of the social support programs that the councilors are supposed to monitor through their committees.

[28] Report cards had specific information only about T2 wards, but that difference may be less relevant.

4.2 Private disclosures: State of Sanitation Information

A. Audit outcomes and empirical specification

Our audit performance metrics are based on the two public goods included in the private SSI reports: sanitation and garbage. Both lend themselves to simple improvements by a councilor, but are also goods where improvements are either less visible (sewer repairs) or require ongoing maintenance (toilets and garbage). On sanitation, our main outcomes are a toilet availability index (based on total number of toilets, open toilets, and adult users) and the price of using a toilet. For garbage, we separately measure the incidence of formal and informal garbage piles, but also a composite collection rate index.

For comparison with Table I, we first report regressions of the same form as equation (2) where the treatment indicator is whether the councilor received an SSI report and the outcome is pro-poor spending for our audit sample of 103 wards. Next, we use the three audit rounds to estimate slum-level difference-in-differences ITT regressions:

$$Service_{sw} = \beta Post_t \times Treat_{sw} + \gamma Post_t + X_w + \epsilon_{sw}, \tag{4}$$

where $Treat_{sw}$ is the slum's ward treatment status, $Post_t$ is a midline or endline audit indicator, X_w is a ward fixed effect, and $Service_{sw}$ are public services outcome variables. We cluster standard errors by ward.

B. Results

Table II presents results. For (ward-level) pro-poor spending we observe no significant impact of the SSI (column 1). In contrast, Appendix Table A.VI columns (3) and (4) show that, consistent with Table I results, newspaper report cards differentially raised pro-poor spending in treated wards in the audit sample as well. In columns (2)–(6), we consider all the audit performance metrics and observe no treatment impacts. Relative to control wards, slums in treatment wards show an insignificant decline in the toilet availability index (column 2).[29] Toilet prices are unaffected (column 3). We observe no impact of SSI treatment on incidence of formal or informal garbage collection piles (columns 4–5). Garbage collection

[29] In Appendix Table A.VIII we consider report on the individual elements of the toilet availability index. The decline in the toilet availability index is driven by a significant decrease in the number of open toilets, despite no overall change in the number of toilets. One possibility is that the councilor asked contractors to improve the worst toilets and they chose to close them. A less charitable possibility is that the councilor favored closing dirty toilets in order to improve appearances at the cost of usage.

rates are also unaffected.

Given these results, it is unsurprising that Appendix Table A.VIII shows no spillover impacts on the quality of the unreported public good, drains.

The absence of impacts of the SSI treatment suggests that actionable information by itself may not suffice to get councilors to act.[30] In particular, the SSI audits highlighted localized and less-visible public goods. Councilors may have felt that, absent public disclosures, it was unlikely that a substantial number of voters and/or the party would observe improvements in the quality of local public goods *and* attribute these improvements to the councilor.

5 Does information influence party and voter behavior?

The anticipation of newspaper disclosures impacted councilor performance.[31] Did the disclosures themselves also influence electoral outcomes? The answer is both of independent interest, and helps identify whether voter and party reactions justify councilors' performance response.

A. What impacts should we expect?

87 of the 160 eligible incumbents reran for election. Of the 80 incumbents ineligible to run from their current ward, eight ran in a different ward. Foreshadowing the results later in this section, all eight came from treatment wards. The facts that relatively few incumbents reran, and that most ineligible incumbents were dropped, suggests an important role for the party in allocating tickets to rerun. The party may replace an eligible incumbent who had opted not to run by an ineligible incumbent, or conversely, choose to allocate a party ticket to an ineligible councilor rather than allow the eligible incumbent to rerun as party candidate. It could also bring in a new candidate altogether.

In Appendix B, we examine the party's decision problem when it allocates one party ticket to one of two incumbents with identical performance measures – one based in a treatment ward and has a public report card, while the other is in a control ward and does not. The party assumes that voters use knowledge about incumbent's past performance to predict future performance and vote for the incumbent with expected better performance.

[30] This offers one explanation of similar T1 and T2 impacts, despite T2 councilors receiving advanced warning about own performance. However other explanations are plausible – for example, T1 councilors could have seen T2 reports cards and privately constructed their own based on that template.

[31] The SSI treatment did not impact electoral outcomes. Analysis available from authors.

In the world of this very simple model, if, absent a report card, voters know everything about the incumbent, then treatment status is moot. The party can choose either of two equally good incumbents. When voters lack information about the incumbent in control, the party decision rule is to select the treated incumbent if and only if his performance is good enough. Finally, if voters have noisy information about control incumbent's performance, then the rule varies with nature of information. In particular, we show by example that the party may favor high performing control incumbents over comparable incumbents in treatment and this effect may be strongest among best performing incumbents.

Applied to our context, this suggests the interaction of treatment and report card performance should yield a positive effect on the probability of the incumbent being selected to rerun (and hence their vote share) in low-information settings. This fits best the situation when an ineligible incumbent from a different ward is running in a new ward where people do not know him. Given that the set of ineligible incumbents selected to rerun is small, those selected will come mostly from the right tail of performance distribution in treatment. In this case, treatment will have an average positive effect on being selected for rerunning. Further, as our measure of performance captures pro-poor spending, we anticipate that the best report cards will be the ones most attractive to low income voters. Hence high-performing treated ineligible incumbents should be targeted to high-slum-density wards.

The situation is less clear cut when the party has an eligible incumbent candidate. Our prediction of the party choosing high performers in treatment over comparable politicians in control, may not hold in such cases, if voters have a significant amount of information about candidates even without a report card. Moreover, the party may account for information that voters are privy to but that is not captured by our performance measure. (The person could be famously efficient or scrupulously honest.) Indeed voters may punish the party for replacing an incumbent who they know and like, even if by our measures of performance he or she is not stellar. All these factors can weaken the link between our measure of performance, treatment and the probability of rerunning for eligible incumbents.

At the party level, the logic is clear-cut where an incumbent's ineligibility requires the party to field a new candidate. In that case, the new candidate will benefit if the previous incumbent performed well and enhanced the party reputation. This is less obvious where the incumbent is eligible. When he reruns, past performance should indeed have a positive impact, though the effect may be muted because voters have other information about the incumbent. However, when an eligible incumbents does not rerun (recall only about half of them rerun) and in particular a well-performing eligible incumbent gets replaced (say,

because of internal party politics), voters may punish the party. In other words, for eligible candidates who do not rerun, the party may lose vote share precisely when they are high performers and perhaps, especially in the case where there is a report card showing that the performed well.

Finally, high pro-poor spending in high-slum-density wards potentially sends a clearer positive signal to low-income voters than a similar spending pattern in low-slum-density wards. Voters in high-slum-density wards may think that the reported high levels of spending on sanitation in low-slum-density wards was not in slums. (However, one could also make the reverse argument – that being very pro-poor in a low-slum-density ward shows real commitment to the poor.)

The final set of hypotheses concern yardstick competition. In the presence of common shocks the incumbent's rank relative to the set of incumbents with report cards in the same newspaper edition contains information beyond our performance measure. Hence, we should expect a voter response to the rank, even after we control for our performance measure.

Unlike the performance measure, where voters might have independent information about their incumbent, voters probably had no way of knowing relative rank. For this reason, there is a stronger case to expect an effect of rank even among eligible incumbents. Apart from the independent interest of this potential result, this is useful because, relative to ineligible incumbents, the pool of eligible incumbents is larger and they are more likely to rerun. Hence, small sample concerns are lessened.

B. Specification

We report ITT estimates of electoral impacts pooled across the two treatment arms. For each outcome, we consider average effect of report cards followed by whether impacts differ by councilor eligibility status. Specifically, we estimate:

$$Outcome_c = \beta Treat_c + X_c + \epsilon_c, \tag{5}$$

$$Outcome_c = \beta_1 Treat_c + \beta_2 Treat_c \times Inelig_c + \gamma Inelig_c + X_c + \epsilon_c, \tag{6}$$

where $Treat_c$ is councilor treatment status, $Inelig_c$ is whether the councilor is ineligible for reelection in the current ward, and X_c is a fixed effect for randomization stratum (zone–party combination). To account for small number of ineligibles running in other wards, we report, in addition to heteroscedasticity-robust standard errors, p-values for each coefficient using

randomization inference.

Finally, we ask whether treatment effects differ based on councilor performance:

$$Outcome_c = \beta_i Treat_c \times Y_c + \beta_j Treat_c \times Inelig_c \times Y_c + \gamma_i Inelig_c \times Y_c + \gamma_j Y_c + X_c + \epsilon_c, \quad (7)$$

where Y_c is a vector of the councilor's pro-poor spending index and attendance index. Both performance measures are based on four years of data (the fiscal years 2007–08, 08–09, 09–10 and 10–11). This was the data that was presented for treated councilors in March 2012 report cards, and is the metric that parties and voters arguably based their decisions on. As discussed earlier, treatment impacts on councilor performance in fiscal year 2010–11 could lead to potential misinterpretations. To account for this endogeneity, Appendix Table A.X show versions of the results which only include performance data for fiscal years 2007–08, 08–09, and 09–10, i.e. prior to the councilor letters being released.

We present analysis at both councilor- and party-level. First, we consider the councilor as the relevant unit of analysis: treatment, ineligibility, performance, and electoral outcomes are characteristics that "follow" a councilor if he runs in a different ward. Next, we consider incumbent party vote share as a ward-level outcome, $Outcome_w$, together with ward-level treatment, ineligibility, and performance variables, e.g. $Treat_w$, that are associated with the incumbent councilor rather than the candidate who ultimately runs in that ward. If report cards influence party ticket allocation, then even if voter support for the incumbent is unaffected by treatment, a party selection effect can generate a positive *incumbent* treatment effect. In contrast, a *ward* treatment effect is only possible if the report card alters voter preferences.[32]

Finally, we undertake an analysis of yardstick competition within the treated sample. We define $Rank_c$ as councilor c's pro-poor spending index rank, calculated across the four councilors randomly assigned to the same newspaper edition. All regressions control for absolute councilor spending performance. Time constraints prevented us from featuring our full sample of treated councilors, so the sample comprises 124 councilors for whom report cards were published. Furthermore, we prioritized publishing report cards for eligible councilors. Given this, the sample of ineligible incumbents who featured in a report card and subsequently received a party ticket is arguably too small to estimate treatment effects. These sample size concerns do not impact the party-level analysis where we ask whether the

[32]Appendix Table A.XI shows that the treatment didn't affect voter registration or voter turnout on average, though wards with eligible incumbents with high reported attendance did see increased turnout.

party's electoral performance was improved by the incumbent's performance. Hence, we restrict our yardstick analysis to a party-level analysis and estimate:

$$Outcome_w = \beta_1 Rank_c + \beta_2 Inelig_w \times Rank_c + \gamma_1 Inelig_w + \gamma_2 D_w + \gamma_3 D_w \times Rank_c + Y_w + X_w + \epsilon_w, \quad (8)$$

D_w, the newspaper edition publication date, controls for our prioritization of publishing report cards for eligible incumbents. Y_w is a vector of the councilor's pro-poor spending index, this index separately interacted with ineligibility, and the publication date, all of which is included to separate absolute performance from relative performance. X_w denotes the randomization stratum.

C. Results

Councilor electoral performance

Table III presents results for pooled treatment, and Appendix Table A.IX separates T1 and T2 treatment arms. In column (1) we consider the set of eligible incumbents' and examine their probability of running for the 2012 election from any ward. 48% of eligible incumbents in control wards ran for re-election, all but two of them in their current wards. Treatment has a positive but insignificant impact on the propensity of eligible incumbents to rerun. This propensity is not impacted by the incumbent's performance record, measured by either of two metrics: pro-poor spending index and attendance index (column 2).

Why did treatment not impact eligible incumbents propensity to rerun, especially for better performers? Given that eligible incumbents almost always run in their current wards if they run at all, it could be that voters, both in treatment and control, know enough about them to make the report cards redundant. But it could also be that performance as measured in the report card does not matter for eligible incumbents either in treatment or in control. There is some weak evidence against the latter hypothesis: specifically, the (non-experimental) correlation between performance and electoral outcomes for eligible incumbents in the control group suggests that performance does matter. A one standard deviation increase in pro-poor spending increases the probability that an eligible incumbent in the control group reruns by seven percentage points ($p = 0.46$), omitted from the table but available from the authors. For the subsample of high-slum-density wards, this rises to 15 percentage points ($p = 0.46$).

In columns (3)–(5) we examine whether incumbents ran outside their current ward. Here,

we consider the full sample of 240 councilors since, irrespective of eligibility, incumbents are on *de jure* equal footing. We observe a four percentage point increase ($p = 0.07$) in the likelihood that a treated councilor runs in a different ward.[33] This rises to a 16 percentage point increase ($p < 0.01$) for the subset of ineligible councilors (column 4). As noted earlier, none of the 26 ineligible councilors from control wards ran for re-election. In contrast, ineligible incumbent councilors in treatment were significantly and substantially more likely to get a ticket to run in a different ward, with eight of the potential 54 doing so. Moreover, the increased likelihood of receiving a party ticket was concentrated among ineligible councilors with a more pro-poor spending record (column 5); for each standard deviation increase in pro-poor spending record, ineligible treated incumbents were 13 percentage points more likely ($p = 0.03$) to get a ticket to run in another ward.

In Table IV we explore party placement of incumbents who receive a ticket for a new ward. The set of high-performing treated ineligible incumbents who get to run in a new ward come mainly from high-slum-density wards (column 2). Performance along both dimensions – spending and attendance – matters, though the latter effect is more noisily estimated. At the margin, ineligible treated incumbents from high-slum-density wards are 31 percentage points more likely ($p = 0.03$) to be given a ticket to run elsewhere, for each standard deviation increase in their pro-poor spending record. Those with a similarly good spending record from low-slum-density wards do not get the same advantage, though small sample sizes restrict our ability to say anything definitive about this comparison.

High-performing ineligible incumbents from treatment wards tend to displace incumbents from the same party in other high-slum-density wards. The latter is consistent with the idea, suggested above, that incumbents with report cards showing high pro-poor spending are most likely to be popular in high-slum-density wards. At the margin, ineligible treated incumbents are 22 percentage points more likely ($p = 0.04$) to be given a ticket to run in another high-slum-density ward controlled by their party for each standard deviation increase in their pro-poor spending record. We see some evidence that a good attendance record matters, though as before this effect is more noisily estimated.

Parties appear to base selection among ineligible incumbents on report card metrics. Eligible incumbents are almost never moved to other wards – if they are popular they can remain in their own ward and if they are unpopular, they are unlikely to be worth moving, especially since they compete with the entire pool of ineligible incumbents, which, being

[33] The p-value associated with the randomization inference test is noisier at 0.19.

randomly selected, contains some of the best performing incumbents.

Returning to Table III, columns (6)–(8) show that treatment-induced increases in propensity to run in a different ward translated into higher vote shares. This is partly, but not entirely, mechanical, since voters could reject the new candidate. The gains are concentrated among ineligible incumbents with better pro-poor spending records (column 8). At the margin, ineligible treated incumbents earn a nine percentage points higher ($p = 0.05$) vote share for each standard deviation increase in their pro-poor spending record. There is weak evidence that the same is true among control incumbents. Among eligible incumbents in the control group, the coefficient on pro-poor spending is a positive four percentage points ($p = 0.30$), omitted from the table but available from the authors. In the subsample from high-slum-density wards, this coefficient rises to positive ten percentage points ($p = 0.11$). Once again, these non-experimental correlations are consistent with view that performance matters for eligible incumbents, but having a report card has no additional effect.

As a robustness check, Appendix Table A.X shows very similar patterns when we re-estimate the performance heterogeneity regressions, using pre-treatment performance measures (spanning the fiscal years 2007–08, 08–09, and 09–10).

Our findings support the idea that retrospective electoral accountability underlies the responsiveness of councilor performance to anticipated public disclosures. An important channel is ticket allocation for ineligible incumbents: parties field ineligible incumbents with good performance records in serving their own *low-income* wards in other slum-dense wards. These benefits do not accrue to untreated ineligible incumbents, even if they have high pro-poor spending, suggesting either that parties access performance information via media, or that parties value the credibility associated with newspaper disclosures.

Party electoral performance

Table V presents ward-level results for electoral rewards that may accrue to a party whether or not their incumbent councilor reruns. This evidence sheds light on whether treatment altered voter preferences across candidates (rather than just affecting the party candidate selection). Columns (1) and (2) show that there is no residual vote share benefit for the party from simply having had a report card published. In contrast, column (3) shows that in wards where the incumbent can't rerun, the party benefits from its incumbent featuring in a report card showing high pro-poor spending. At the margin, a party earns 10 percentage points higher ($p < 0.01$) vote share for each standard deviation increase in the pro-poor

spending record of a treated, ineligible incumbent. The estimated impacts for attendance index are noisier but point in the same direction: better performance improves incumbent party vote share. Once again, the results suggests that voters have a different relation with eligible incumbents; perhaps, as suggested above, they already know what they like (and dislike) about them (which may or may not be reflected in the report card).

In column (4) we consider yardstick competition *within the sample of wards with treated councilors with published report cards.* These results are important for two reasons. First, one reason our treatment may not have an effect on *eligible* incumbents is that they typically run in their current wards and therefore voters may have enough information about their performance even without a report card. However, absent the report card voters were unlikely to have the rank information and, therefore, this is one place where we may expect an impact of performance for treated eligible incumbents (as well as ineligible ones). Second, our focus so far has been on the interaction between treatment and performance, but the latter is not randomly assigned, and may be correlated with other candidate or ward characteristics. The variation identifying the yardstick regressions comes from the random choice of report cards featuring on the same page of the newspaper. This provides some confidence that the voters are actually reacting to performance information.

We find robust evidence that the incumbent's rank impacts party vote share in the incumbent's current ward (whether or not the incumbent runs). A worsening of one rank reduces the party vote share by 7 percentage points ($p = 0.02$). Consistent with the argument above, there is no difference between eligible and ineligible candidates.

6 Conclusion

Anticipation of public disclosures motivates councilors representing high-slum-density wards to better align spending priorities with their constituents. On the other hand, disclosing information privately to politicians has no discernible positive effects on their spending priorities or provision of public infrastructure in slums. We posit that, absent a dissemination mechanism like newspapers, councilors do not anticipate electoral benefits from investments in valued but less visible local public goods.

Our second finding is that councilors are right to respond to the promise of forthcoming report cards. Parties are more likely to give tickets to treated councilors from high-slum-density wards, who, according to their report cards, have served slum dwellers well. Among high-performing incumbents, those rendered ineligible to run from their current ward because

of gender quotas benefit, and get to run from other wards. These effects are large; ineligible treated incumbents from high-slum-density wards are 29 percentage points more likely to receive a ticket to run elsewhere, for each standard deviation increase in their pro-poor spending record. Finally, voters in high-slum-density wards also reward the party associated with high-performing incumbents who get a report card. Being ranked higher among those reported in the same edition also benefits the incumbent's party.

Quota-induced ineligibility mimics, in many ways, term limits for politicians. Furthermore, it is common for term-limited politicians to seek other political offices. Our findings suggest that public disclosures could be valuable for this group, especially since they do not face electoral accountability pressures in their current seat.

Unfortunately, in a time of increasing financial pressure, credible print or online media may lack financial incentives to publish these disclosures. To support a better functioning democracy, there is a need to explore how independent institutions that perform this role can be formed and funded.

References

Adida, C., J. Gottlieb, E. Kramon, and G. McClendon (2020). When does information influence voters? the joint importance of salience and coordination. *Comparative Political Studies 53*(6), 851–891.

Arias, E., H. Larreguy, J. Marshall, and P. Querubín (2018, August). Priors rule: When do malfeasance revelations help or hurt incumbent parties? Working Paper 24888, National Bureau of Economic Research.

Avis, E., C. Ferraz, and F. Finan (2018). Do government audits reduce corruption? Estimating the impacts of exposing corrupt politicians. *Journal of Political Economy 126*(5), 1912–1964.

Banerjee, A., S. Kumar, R. Pande, and F. Su (2011). Do informed voters make better choices? Experimental evidence from urban India. Unpublished.

Besley, T. (2007). *Principled Agents?: The Political Economy of Good Government*. Oxford University Press.

Besley, T. and R. Burgess (2002). The political economy of government responsiveness: Theory and evidence from India. *The Quarterly Journal of Economics 117*(4), 1415–1451.

Besley, T. and A. Case (2003, March). Political institutions and policy choices: Evidence from the united states. *Journal of Economic Literature 41*(1), 7–73.

Besley, T., O. Folke, T. Persson, and J. Rickne (2017). Gender quotas and the crisis of the mediocre man: Theory and evidence from Sweden. *American Economic Review 107*(8), 2204–42.

Bidwell, K., K. Casey, and R. Glennerster (2020). Debates: Voting and expenditure responses to political communication. *Journal of Political Economy 128*(8), 2880–2924.

Boas, T. C., F. D. Hidalgo, and M. A. Melo (2018). Norms versus action: Why voters fail to sanction malfeasance in Brazil. *American Journal of Political Science 63*(2), 385–400.

Bobonis, G. J., L. R. Cámara Fuertes, and R. Schwabe (2016, August). Monitoring corruptible politicians. *American Economic Review 106*(8), 2371–2405.

Buntaine, M. T., R. Jablonski, D. L. Nielson, and P. M. Pickering (2018). Sms texts on corruption help Ugandan voters hold elected councillors accountable at the polls. *Proceedings of the National Academy of Sciences 115*(26), 6668–6673.

Chong, A., A. L. De La O, D. Karlan, and L. Wantchekon (2015). Does corruption information inspire the fight or quash the hope? A field experiment in Mexico on voter turnout, choice, and party identification. *The Journal of Politics 77*(1), 55–71.

de Figueiredo, M., F. Hidalgo, and Y. Kasahara (2012). When do voters punish corrupt politicians? Experimental evidence from Brazil. Presented at 7th Annual Conference on Empirical Legal Studies.

Delhi Human Development Report (2006). Oxford University Press, New Delhi.

Dunning, T., G. Grossman, M. Humphreys, S. D. Hyde, C. McIntosh, G. Nellis, C. L. Adida, E. Arias, C. Bicalho, T. C. Boas, M. T. Buntaine, S. Chauchard, A. Chowdhury, J. Gottlieb, F. D. Hidalgo, M. Holmlund, R. S. Jablonski, E. Kramon, H. Larreguy, M. Lierl, J. Marshall, G. McClendon, M. A. Melo, D. L. Nielson, P. M. Pickering, M. Platas, P. Querubin, and N. Sircar (2019). Voter information campaigns and political accountability: Cumulative findings from a preregistered meta-analysis of coordinated trials. *Science Advances 5*(7), 1–10.

DUSIB (2010). The Delhi Urban Shelter Improvement Board Act. http://delhishelterboard.in/main/wp-content/uploads/2012/02/DUSIB-ACT-2010.pdf.

Enikolopov, R., M. Petrova, and E. Zhuravskaya (2011, December). Media and political persuasion: Evidence from russia. *American Economic Review 101*(7), 3253–85.

Ferraz, C. and F. Finan (2008). Exposing corrupt politicians: The effects of Brazil's publicly released audits on electoral outcomes. *The Quarterly Journal of Economics 123*(2), 703–745.

Fujiwara, T. and L. Wantchekon (2013). Can informed public deliberation overcome clientelism? Experimental evidence from Benin. *American Economic Journal: Applied Economics 5*(4), 241–55.

Green, J. and N. Stokey (1983). A comparison of tournaments and contracts. *Journal of Political Economy 91*(3), 349–64.

Grossman, G. and K. Michelitch (2018). Information dissemination, competitive pressure, and politician performance between elections: A field experiment in Uganda. *American Political Science Review 112*(2), 280–301.

Holland, A. (2016). Forbearance. *American Political Science Review 110*(2), 232–246.

IL & FS Ecosmart Limited (2006). City development plan: Delhi. Ministry of Urban Development, Government of National Capital Territory of Delhi. Submitted under JNNURM.

International Budget Partnership (2010, Jan). Constituency development funds: Scoping paper. Technical report, International Budget Partnership.

Jha, S., V. Rao, and M. Woolcock (2007). Governance in the gullies: Democratic responsiveness and leadership in delhi's slums. *World Development 35*(2), 230–246.

Kendall, C., T. Nannicini, and F. Trebbi (2015, January). How do voters respond to information? Evidence from a randomized campaign. *American Economic Review 105*(1), 322–53.

Larreguy, H., J. Marshall, and J. Snyder, James M (2020, 05). Publicising malfeasance: When the local media structure facilitates electoral accountability in Mexico. *The Economic Journal 130*(631), 2291–2327.

Lazear, E. P. and S. Rosen (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy 89*(5), 841–864.

Lehne, J., J. N. Shapiro, and O. Vanden Eynde (2018). Building connections: Political corruption and road construction in india. *Journal of Development Economics 131*, 62–78.

Mendel, T. D. (2014). Right to information: Recent spread of RTI legislation. Technical report, World Bank Group.

Post, A. (2018). Cities and politics in the developing world. *Annual Review of Political Science 21*, 115–133.

Snyder Jr, J. M. and D. Strömberg (2010). Press coverage and political accountability. *Journal of Political Economy 118*(2), 355–408.

Strömberg, D. (2004). Mass media competition, political competition, and public policy. *The Review of Economic Studies 71*(1), 265–284.

Wantchekon, L. (2003). Clientelism and voting behavior: Evidence from a field experiment in Benin. *World Politics 55*(3), 399–422.

7 Tables

Table I: Impact of report card on councilor performance by slum density

	Log total spending (2010–12)			Pro-poor spending index (2010–2012)			Directly elicited spending index (2010–2012)			Attendance index (2010–12)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treatment × High slum		−0.059	−0.111		0.656	0.818		0.498	0.638		0.285	−0.023
		(0.074)	(0.098)		(0.291)	(0.373)		(0.205)	(0.261)		(0.197)	(0.234)
Treatment	0.005	0.055	0.058	−0.010	−0.325	−0.414	−0.020	−0.271	−0.317	0.030	−0.173	0.101
	(0.040)	(0.056)	(0.077)	(0.133)	(0.198)	(0.269)	(0.096)	(0.146)	(0.194)	(0.100)	(0.146)	(0.174)
Ineligible × Treatment × High slum			0.130			−0.541			−0.473			0.848
			(0.160)			(0.614)			(0.446)			(0.503)
Ineligible × Treatment			0.028			0.286			0.151			−0.769
			(0.117)			(0.418)			(0.316)			(0.381)
Remaining interactions	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Pre-treat outcome control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Strata (zone–party) FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-treat control mean	5.709	5.721	5.721	−0.000	0.019	0.019	0.000	0.010	0.010	0.000	−0.004	−0.004
Pre-treat control s.d.	0.157	0.109	0.109	1.000	0.923	0.923	1.000	0.952	0.952	0.893	0.912	0.912
p-value: $T \times High + T + High = 0$		0.959			0.797			0.868			0.924	
p-value: $T \times High + High = 0$		0.240			0.014			0.010			0.165	
p-value: $T \times High + T = 0$		0.947			0.105			0.101			0.409	
Observations	240	227	227	240	227	227	240	227	227	240	227	227

Heteroskedasticity-robust standard errors in parentheses.

Ward-level OLS regression. The pro-poor spending index is the mean z-score of log spending on issues, with each issue weighted by the fraction of slum households in the city reporting that (1) it is the most problematic in the area, (2) it is a problem for them, and (3) it is a problem for the community. The directly elicited spending index is the z-score of the inverse hyperbolic sine of spending on issues, with each issue weighted by the fraction of slum households in the city reporting that their councilor should spend more on it minus the fraction reporting that they should spend less on it. The attendance index is the mean z-score of councilor attendance at (1) the general assembly and (2) councilor committee meetings. "Treatment" indicates observations of a ward in which a report card on the performance of the MCD councilor was published in a newspaper during the 2012 pre-election period (T1 or T2, ITT). Spending is categorized by lexical heuristic.

Table II: Impact of private information on public services

	(1) Pro-poor spending index	(2) Toilet availability index	(3) Toilet price	(4) Total formal piles	(5) Total informal piles	(6) Garbage collection index
SSI treatment	−0.049					
	(0.190)					
Post × SSI treatment		−0.048	0.065	−0.087	0.607	−0.055
		(0.031)	(0.108)	(0.058)	(2.224)	(0.042)
Post		0.052	0.047	0.196	−2.520	0.332
		(0.023)	(0.061)	(0.049)	(1.423)	(0.028)
Ward FE	No	Yes	Yes	Yes	Yes	Yes
Strata (zone–party) FE	Yes	No	No	No	No	No
Pre-treat outcome control	Yes	No	No	No	No	No
Baseline control mean	0.102	0.935	0.973	0.436	17.85	0.256
Slums in sample		312	156	312	312	309
Observations	103	932	430	932	932	894

Standard errors clustered by ward in parentheses.

Column (1) is a ward-level OLS regression; the remaining columns are slum-level OLS regressions. "SSI treatment" indicates observations in a slum of which the MCD councilor received State of Sanitation Information (ITT). "Post" indicates observations that took place in the second or third round of audits. Of 932 slum-round observations, 529 have a public toilet. "Toilet availability index" is the mean z-score of log total toilets, log open toilets, and log adult toilet users, with first percentile values imputed for zeroes. "Toilet price" is the average price of toilets in the slum. The sample is restricted to slums with one or more toilets. "Total formal piles" is the number of formal garbage piles (dhalaos) in the slum. "Total informal piles" is the number of informal garbage piles in the slum. "Garbage collection index" is the mean of the fraction of formal piles in the slum regularly collected, relative to the number of formal piles at the baseline, and the fraction of informal garbage piles in the slum collected in the past week. The sample is restricted to slums with one or more garbage piles.

Table III: Impact of report card on incumbent electoral outcomes

	Eligible incumbent runs in any ward		Incumbent runs in other ward			Incumbent's vote share (0 if didn't run)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.078	0.391	0.039	−0.012	0.040	0.047	0.032	0.078
	(0.089)	(0.320)	(0.022)	(0.019)	(0.071)	(0.027)	(0.035)	(0.127)
	[0.443]	[0.263]	[0.186]	[0.637]	[0.493]	[0.094]	[0.373]	[0.617]
Treatment × Ineligible				0.156	0.134		0.024	−0.197
				(0.056)	(0.226)		(0.047)	(0.171)
				[0.048]	[0.604]		[0.612]	[0.334]
Treatment × ...		−0.034			−0.048			−0.041
... Pro-poor Spending Index		(0.104)			(0.034)			(0.038)
		[0.709]			[0.069]			[0.250]
Treatment × ...		−0.460			−0.086			−0.065
... Attendance Index		(0.468)			(0.108)			(0.178)
		[0.359]			[0.373]			[0.770]
Treatment × Ineligible × ...					0.126			0.094
... Pro-poor Spending Index					(0.059)			(0.047)
					[0.093]			[0.049]
Treatment × Ineligible × ...					0.043			0.347
... Attendance Index					(0.338)			(0.240)
					[0.914]			[0.234]
Additional Interactions	No	Yes	No	Yes	Yes	No	Yes	Yes
Strata (zone–party) FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ineligible control mean			0	0	0	0	0	0
Eligible control mean	0.478	0.478	0.0217	0.0217	0.0217	0.171	0.171	0.171
Observations	160	160	240	240	240	240	240	240

Heteroskedasticity-robust standard errors in parentheses.

Councilor-level cross section estimated with OLS and with randomization inference (1,000 repetitions). Robust standard errors in parentheses. Randomization inference p-values in brackets. Randomization strata are fixed. "Attendance Index" is overall councilor attendance at MCD committees of which they are a member (2007–11). "Pro-poor Spending Index" is the mean z-score of three log preference-weighted spending amounts (2007–11), analogous to the dependent variable in Table I, column 3, for the pre-publication period. "Additional interactions" are "Eligible", "Pro-Poor Spending Index", and "Eligible × Pro-Poor Spending Index", "Attendance Index", and "Eligible × Attendance Index".

Table IV: Impact of report card on incumbent electoral outcomes by slum density

	Incumbent runs in other ward			Incumbent runs in other ward controlled by party		
	(1)	(2) *from* high-slum ward	(3) *from* low-slum ward	(4)	(5) *to* high-slum ward	(6) *to* low-slum ward
Treatment	0.040	0.381	−0.013	0.031	0.198	0.063
	(0.071)	(0.221)	(0.098)	(0.054)	(0.136)	(0.118)
	[0.493]	[0.045]	[0.748]	[0.540]	[0.174]	[0.452]
Treatment × Ineligible	0.134	−0.490	0.175	0.062	−0.412	0.526
	(0.226)	(0.308)	(0.334)	(0.201)	(0.251)	(0.401)
	[0.604]	[0.238]	[0.457]	[0.793]	[0.225]	[0.317]
Treatment × ...	−0.048	−0.156	−0.006	−0.006	−0.040	0.031
... Pro-poor Spending Index	(0.034)	(0.097)	(0.027)	(0.014)	(0.041)	(0.042)
	[0.069]	[0.014]	[0.597]	[0.693]	[0.412]	[0.154]
Treatment × ...	−0.086	−0.474	0.000	−0.043	−0.226	−0.127
... Attendance Index	(0.108)	(0.291)	(0.175)	(0.074)	(0.166)	(0.171)
	[0.373]	[0.060]	[0.997]	[0.491]	[0.186]	[0.353]
Treatment × Ineligible × ...	0.126	0.310	0.079	0.076	0.222	−0.008
... Pro-poor Spending Index	(0.059)	(0.136)	(0.066)	(0.052)	(0.107)	(0.088)
	[0.093]	[0.078]	[0.465]	[0.300]	[0.163]	[0.943]
Treatment × Ineligible × ...	0.043	0.885	−0.098	0.092	0.748	−0.540
... Attendance Index	(0.338)	(0.476)	(0.499)	(0.308)	(0.423)	(0.561)
	[0.914]	[0.190]	[0.860]	[0.817]	[0.203]	[0.428]
Additional Interactions	Yes	Yes	Yes	Yes	Yes	Yes
Strata (zone–party) FE	Yes	Yes	Yes	Yes	Yes	Yes
Ineligible control mean	0	0	0	0	0	0
Eligible control mean	0.0217	0.0400	0	0	0	0
Observations	240	114	114	240	113	116

Heteroskedasticity-robust standard errors in parentheses.

Councilor-level cross section estimated with OLS and with randomization inference (1,000 repetitions). Robust standard errors in parentheses. Randomization inference p-values in brackets. Randomization strata are fixed. "Attendance Index" is overall councilor attendance at MCD committees of which they are a member (2007–11). "Pro-poor Spending Index" is the mean z-score of three log preference-weighted spending amounts (2007–11), analogous to the dependent variable in Table I, column 3, for the pre-publication period. "Additional interactions" are "Eligible", "Pro-Poor Spending Index", "Eligible × Pro-Poor Spending Index", "Attendance Index", and "Eligible × Attendance Index".

Table V: Impact of report card on incumbent party vote share

	Incumbent's party's vote share in same ward			
	(1)	(2)	(3)	(4)
Treatment	−0.002	−0.001	−0.034	
	(0.018)	(0.022)	(0.081)	
	[0.905]	[0.983]	[0.772]	
Treatment × Ineligible		−0.000	−0.133	
		(0.041)	(0.150)	
		[0.995]	[0.476]	
Treatment × ...			−0.017	
... Pro-poor Spending Index			(0.025)	
			[0.429]	
Treatment × ...			0.048	
... Attendance Index			(0.116)	
			[0.766]	
Treatment × Ineligible × ...			0.097	
... Pro-poor Spending Index			(0.036)	
			[0.025]	
Treatment × Ineligible × ...			0.230	
... Attendance Index			(0.222)	
			[0.401]	
Rank				−0.073
				(0.031)
Ineligible × Rank				0.006
				(0.055)
Additional Interactions	No	Yes	Yes	Yes
Strata (zone–party) FE	Yes	Yes	Yes	Yes
Ineligible control mean	0.408	0.408	0.408	
Eligible control mean	0.347	0.347	0.347	
Ineligible treated mean	0.390	0.390	0.390	0.370
Eligible treated mean	0.352	0.352	0.352	0.349
Observations	240	240	240	124

Heteroskedasticity-robust standard errors in parentheses.

Ward-level cross section estimated with OLS and with randomization inference (1,000 repetitions). Robust standard errors in parentheses. Randomization inference p-values in brackets. Randomization strata are fixed. "Attendance Index" is overall attendance at MCD committees of which they are a member (2007–11). "Pro-poor Spending Index" is the mean z-score of three log preference-weighted spending amounts (2007–11), analogous to the dependent variable in Table I, column 3, for the pre-publication period. "Rank" is the rank of the incumbent councilor's mean z-score of three log preference-weighted spending amounts among the four councilors whose report cards appeared in the same newspaper. (Highest = 0; lowest = 3.) "Publication Date" is scaled between 0 (first issue published) and 1 (last issue published). "Additional interactions" are "Eligible", "Pro-Poor Spending Index", and "Eligible × Pro-Poor Spending Index", as well as "Attendance Index" and "Eligible × Attendance Index" for the "Treatment" regressions and "Publication Date", "Publication Date × Rank", and "Publication Date × Pro-Poor Spending Index" for the "Rank" regressions.

A Data Appendix

A.1 Slum Identification

We identify slums following a methodology based on the UN-HABITAT and Indian census definition of slums.[34] A list of nine common criteria closely correlated to the census definition of slums was drawn up and included high density of housing, poor quality housing structure and material, lack of internal household infrastructure, poor road infrastructure, access to water and water infrastructure, uncovered and unimproved drains, low coverage of private toilet facilities, high incidence of trash piles and frequent cohabitation with animals.[35]

We used a two-stage process: first, we compiled a list of potential areas from inspection of the visual appearance from aerial photographs of Delhi using satellite imagery, based on housing density and appearance, complemented by Delhi government listings. This was then verified by field visits; locations that prominently featured at least five of these nine characteristics were marked as more slummy and others as less slummy.

Between 9 and 126 households were surveyed in each ward[36], with the exact number in a slum dependent on the number of potential slums identified by satellite image in each ward and the physical size of the slum. To the extent that population density is similar across different slums, this approximates a Probability Proportional to Size (PPS) sampling procedure. To select households within slums we also used a spatial method: an overall map

[34] The 2011 Indian census defines a slum as a "compact housing cluster or settlement of at least 20 households with a collection of poorly built tenements which are, mostly temporary in nature with inadequate sanitary, drinking water facilities and unhygienic conditions will be termed as slums."; UN-HABITAT defines a slum household as "a group of individuals living under the same roof that lacks any one of meet the following conditions: insecure residential status, inadequate access to safe water, inadequate access to sanitation and other infrastructure, poor structural quality of housing and overcrowding." The main difference between the two is UN-HABITAT's inclusion of insecure residential status; this is an issue that will be explored within the survey work, but since this is the case to some degree in most Delhi slums, we safely omit it.

[35] Housing: Whether the space separating households was sufficiently wide for vehicles larger than motorcycles; housing materials: Whether the majority of houses are made of unimproved brick or lower quality material, including metal and plastic sheeting; internal household infrastructure: Whether household chores (e.g. washing, cooking) were frequently done outside of the house as a proxy for the quality of households' internal infrastructure, since households who conduct these activities outside tend to lack household water supply/drainage or ventilation for cooking smoke; road infrastructure: Whether the majority of roads in the area were unpaved, badly maintained, and of poor quality; water: Whether households receive water from hand pumps, tanker trucks, or lower-grade options; animal cohabitation: Whether non-domestic animals (buffalo, goats, pigs, donkeys) resided in the same tenements as people.

[36] In ten wards, it was found that surveys had been conducted in the wrong areas. In these cases, surveyors were sent back out, and the surveys were redone in the proper areas. In some cases, the incorrect surveys were still conducted in slum areas, so have been included in the data; thus ten wards have sixty or more surveys. In other cases, the wrongly done surveys were dropped.

of each slum was created, and then surveyors were stationed at randomly selected points within the slums. Surveyors then followed the "right hand rule," where each surveyor moves from their start point along the right hand side of the wall, interviewing every X households (where X is determined by the population of the slum).

Overall, we had just over 3,400 households in high-slum-density areas and 2,000 households in 8 low-slum-density neighborhoods (fewer than five slum characteristics).[37]

A.2 2012 Newspaper Report Card Publication Sequence Randomization

All 2010 report cards were published as intended, but subject to space constraints in each newspaper and time constraints, we were unable to publish all 2012 report cards before the 2012 election. We did not publish any report cards after election day. Of 168 report cards for ITT councilors, 124 were published. Six wards were dropped because the councilor was suspended for corruption or died; seven were dropped because they were never sampled (in two of these cases another ward was sampled instead); one was dropped because there were no slums in the ward; and the last thirty were dropped because they could not be published by election day. Out of the 58 T1 wards, 45 were treated, and of the 110 T2 wards, 79 were treated.

First, report cards were categorized according to zone (of which there were ten), to whether the councilor was eligible for reelection, and to whether the ward had above- or below-median slum fraction by area. Then, within these forty categories, each of which had about four report cards, we randomly assigned report cards to pairs. We then assigned a stratum to each pair according to treatment status (T1/T1, T1/T2, or T2/T2) and political party affiliation (no BJP councilor, or at least one BJP councilor). We randomly assigned publication sequence to each pair of report cards for eligible councilors, distributing the above six strata evenly across the publication sequence. Then we repeated the process for half (i.e. as many as possible) of the report cards for ineligible councilors, and these were placed after the eligible councilors in the publication sequence. Report cards for ineligible councilors were published after the deadline for parties to assign candidate tickets. Two pairs of report cards were published in each daily issue of the newspaper.

[37] The survey was typically carried out with the household head (in 51% of the cases) or, in the case where the household head was unavailable or away on two consecutive visits made to the household, with his or her spouse (49% of the cases) or other household member. If a household proved unwilling or unavailable after multiple visits, another was selected using the same method.

A.3 State of Sanitation Information Randomization

Our audits covered the sample of 108 high-slum-density wards that entered our baseline survey. These wards, in turn, were situated in 55 state assembly constituencies (ACs). All ACs were randomized into treatment and control, followed by a balanced randomization of the wards within an AC. In the event that a ward was split across two ACs, it was put in the AC with an unbalanced number of wards. We then separately randomized report card distribution across the two levels of government: 51 wards were randomly assigned to have the MCD councilor receive a ward report card and, out of the 55 ACs, 27 were randomly assigned to receive a AC report cards.[38] In each ward we audited, on average, three slums giving us a sample of 310 slums across 108 wards.

B Appendix on an illustrative example

We are interested in the decision problem of party P_A that selects a candidate to run against party P_B's candidate. Party P_A chooses between two candidates who are identical in all respects except one: the amount of public information available about them. This is to mirror the situation where the party chooses between similar incumbents in treatment and control wards to run in a single seat. We start with voters' decisions for a fixed set of candidates, where Party candidate $i \in \{A, B\}$ will implement policy $\tilde{\theta}_i$, if elected.

Voter choice Voters make their choice based on the distributions of θ_i , a random variable that represents voters' (shared) perception of $\tilde{\theta}_i$. They form the distribution θ_i based on their priors and on signals they receive which, in turn, are influenced by $\tilde{\theta}_i$.

We assume that every voter receives an identical candidate-specific signal. Signal s stands for a policy previously chosen by the candidates (if any). We assume that A is an incumbent with a track record, s_A. If A is 'treated' then s_A is revealed to the voters. For convenience, we assume $s_A = \tilde{\theta}_A$, i.e. the incumbent continues his previous policy. If A is in control, $s_A = \phi$, an uninformative signal. On the other hand, B is new with no track record and is never treated: $s_B = \phi$. Voters might get other informative signals about θ as well. Assume r_A is the signal about θ_A but there is no signal about θ_B; $r_B = \phi$. This is to

[38] Because wards and ACs are not perfectly aligned, this made for a total of 118 Ward–AC combinations: 30 control, 30 where only the MLA received a report card, 32 where only the MCD councilor received a report card, and 26 were both the MLA and MCD councilor received report cards.

simplify the exposition and limit the number of cases. Finally, voter's prior belief about θ_i is represented by a density function $f_i(\theta_i)$ with mean $\bar{\theta}_i$.

In their voting decision voters compare $\mathbb{E}[\theta_A|s_A, r_A] + \varepsilon_A$ to $\mathbb{E}[\theta_B|s_B, r_B] + \varepsilon_B$ where ε_i are candidate specific shocks drawn independently for each candidate and each voter. We assume a large population of voters, such that A's expected vote share is increasing in

$$\mathbb{E}[\theta_A|s_A, r_A] - \mathbb{E}[\theta_B|s_B, r_B].$$

Party choice Assuming ε_i are realized after candidates are chosen, party A chooses a candidate to maximize this expected difference in θs. Based on our assumptions $\mathbb{E}[\theta_B|s_B, r_B] = \mathbb{E}[\theta_B|\phi, \phi] = \bar{\theta}_B$, the voters' (shared) prior. So the party chooses its candidate to maximize $\mathbb{E}[\theta_A|s_A, r_A]$.

Assume that party A has two potential candidates, A_t and A_c, both incumbents, one in treatment and one in control. It has to choose only one to run. We know that $\mathbb{E}[\theta_{A_t}|s_{A_t}, r_{A_t}] = s_{A_t}$ because A_t is in treatment, and $\mathbb{E}[\theta_{A_c}|s_{A_c}, r_{A_c}] = \mathbb{E}[\theta_{A_c}|\phi, r_{A_c}]$ because A_c is in control. $\mathbb{E}[\theta_{A_c}|\phi, \phi] = \bar{\theta}_A$, the voters' shared prior.

Finally let actual θ of both incumbents be the same, which is denoted, as before by $\tilde{\theta}_A$. For the comparative statics, we need variation in $\tilde{\theta}_A$. We assume the actual distribution of $\tilde{\theta}_A$ is given by the density function $g(\tilde{\theta}_A)$, which is the same in treatment and control. Let the support of $g(\tilde{\theta}_A)$ be the same as the support of $f(\theta_A)$ and that the mean of $g(\tilde{\theta}_A)$ is $\bar{\theta}_A$.

The party's choice depends on how voters will react to the treatment information. Whether it makes the party more or less favorable to the incumbent from treatment or control clearly depends on the sign of

$$s_{A_t} - \mathbb{E}[\theta_{A_c}|\phi, r_{A_c}]$$

This, in turn, depends on r_{A_c}, which depends on the realized value, $\tilde{\theta}_A$. Assume that $h(r_A|\tilde{\theta}_A)$ is the density of r_A, conditional on $\tilde{\theta}_A$, which represents the process that generates the signals. Finally to interpret the signal voters need to have a belief about the distribution of r_A conditional on θ_A. Assume this is also given by $h(r_A|\theta_A)$, i.e. voters have the correct model of the signal generating function.

Next, assume the party observes s_{A_t} and $\tilde{\theta}_A$ but not r_{A_c} while choosing the candidate,

and knows voters' priors about θ and $h(r_A|\theta_A)$. Party candidate choice will depend on

$$\mathbb{E}[s_{A_t} - \mathbb{E}[\theta_{A_c}|\phi, r_{A_c}]|\tilde{\theta}_A]$$

The no-information case Suppose r_{A_c} is uninformative (or almost uninformative). This may occur if the incumbent is ineligible to run in their current ward and is otherwise not well-known. Voters only get information beyond their priors from the report card. Then

$$\mathbb{E}[s_{A_t} - \mathbb{E}[\theta_A|\phi, r_{A_c}]|\tilde{\theta}_A] \simeq \mathbb{E}[s_{A_t}|\tilde{\theta}_A] - \bar{\theta}_A = \tilde{\theta}_A - \bar{\theta}_A$$

which is increasing in $\tilde{\theta}_A$. The party should favor high-performing incumbents in treatment over high performing incumbents in control and the reverse for low-performing incumbents. Since it wants the candidates to win, it will want to choose those for whom $s_A > \bar{\theta}_B$. Therefore, if the number of incumbents to be selected to rerun is small, then it always makes sense to choose high-performing incumbents in treatment wards. In this case the average treatment effect across all types of incumbents will be positive, since only (high-performing) treated incumbents will be chosen to rerun (and no one from control).

The full-information case At the other extreme, consider where $r_{A_c} = \tilde{\theta}_A$. This may be the case in an incumbent's current ward, where he is well-known. Then

$$\mathbb{E}[s_{A_t} - \mathbb{E}[\theta_A|\phi, r_{A_c}]|\tilde{\theta}_A] \simeq \mathbb{E}[s_{A_t}|\tilde{\theta}_A] - \tilde{\theta}_A = 0.$$

There is no reason to choose treated incumbents over control incumbents, but there is reason to choose high-performing incumbents to rerun and drop low-performing incumbents.

The partial information case The most interesting case is when voters have limited information about the incumbent in control from the signal r_{A_c}. A voter with the incorrect prior that most candidates are either very good or very bad, for example, might be disappointed to discover that the incumbent is only moderately good but still hold onto the hope that the incumbent in control will be better. Moreover as the quality of the incumbents gets better, the signal about them can improve in such a way the incumbent in control looks better and better to the voters compared to the one in treatment, at least over a range.

This is the argument we now formalize. Consider the following somewhat special case: *ex ante* voters believe candidates are almost always either very good, $\theta_A = G$ or very bad

$\theta_A = -G$ with probability equal to $\frac{1-\eta}{2}$ for η positive but close to zero. The rest of the probability is distributed uniformly between $-G$ and G. On the other hand, $\tilde{\theta}_A$ is uniformly distributed between $-G$ and G. Also assume that r_{A_c} always takes one of two values, r_g and r_b (good and bad) and that

$$h(r_g|\theta_A = \theta) = h(r_g|\tilde{\theta}_A = \theta) = p(\theta)$$

with $p(G) = 1$, $p(0) = \frac{1}{2}$, $p(-G) = 0$ and $p'(\cdot) \geq 0$.

These assumptions immediately give us that $\mathbb{E}[\theta_A|\phi, r_g] \approx G$, $\mathbb{E}[\theta_A|\phi, r_b] \approx -G$, and $\mathbb{E}[[\theta_A|\phi, r_{A_c}]|\tilde{\theta}_A] = p(\tilde{\theta}_A)G + (1 - p(\tilde{\theta}_A))(-G)$. It follows that

$$\mathbb{E}[s_{A_t} - \mathbb{E}[\theta_A|\phi, r_{A_c}]|\tilde{\theta}_A] \simeq \tilde{\theta}_A - [2p(\tilde{\theta}_A) - 1]G.$$

The key observation is that this expression may be negative for some values of $\tilde{\theta}_A \in [G-x, G]$ for $G > x > 0$, such that the party no longer favors high-performing treated incumbents over equally high-performing control incumbents. Moreover for G large enough, it is possible that there exists $g > 0$ such that for $-g < \tilde{\theta}_A < g$,

$$\frac{d(\tilde{\theta}_A - [2p(\tilde{\theta}_A) - 1]G)}{d\tilde{\theta}_A} = 1 - 2p'(\tilde{\theta}_A)G < 0.$$

In other words, improvements in performance may help treatment incumbents less than control incumbents in the relevant range. For example, take the case where

$$p(\tilde{\theta}_A) = \frac{1}{1 + \exp(-\tilde{\theta}_A)}$$

and G is very large. Clearly for $\tilde{\theta}_A = 0$, $p(\tilde{\theta}_A) = \frac{1}{2}$ and

$$\tilde{\theta}_A - [2p(\tilde{\theta}_A) - 1]G = \tilde{\theta}_A = 0.$$

Also

$$p'(0) = \frac{1}{4},$$

so as long as $G > 2$, $1 - 2p'(0)G < 0$. In other words, for $\tilde{\theta}_A$ just below zero, $\tilde{\theta}_A - [2p(\tilde{\theta}_A) - 1]G > 0$ and for $\tilde{\theta}_A$ just above zero, $\tilde{\theta}_A - [2p(\tilde{\theta}_A) - 1]G < 0$. Clearly for g small enough $\tilde{\theta}_A - [2p(\tilde{\theta}_A) - 1]G$ can be positive everywhere on $[-g, 0)$ and negative everywhere on $(0, g]$.

In this example, among high performing incumbents, the party prefers those in control unlike in the no information case.

Moreover if we add to the model some amount of randomness so that the choice of the incumbent is monotonically increasing in $\tilde{\theta}_A - [2p(\tilde{\theta}_A) - 1]G$ (rather than always zero-one), we see that because $1 - 2p'(0)G < 0$ in the neighborhood of $\tilde{\theta}_A = 0$ as long as $G > 2$, it is possible that the probability that the incumbent in control gets chosen is increasing in $\tilde{\theta}_A$ over some range $[-g, g]$.

Summary If voters have little information about candidates other than via report cards, and there are few wards to fill, then parties favor high-performing incumbents from treatment wards over comparable candidates from control wards. When voters have full information about incumbents even without the treatment, then parties are indifferent between treatment and control incumbents with similar records. Finally, when voters have partial information about the incumbent in control, it is possible that the party prefers an incumbent in control over one in treatment with a similar record and that this bias in favor of control incumbents increases with incumbent performance.

C Appendix Tables and Figures

Table A.I: Balance table of councilor performance and constituent preferences

	(1) Roads and lanes	(2) Sewage and drainage	(3) Parks and greenery	(4) Education and schools	(5) Garbage removal	(6) Other areas	(7) Total
Panel: Spending (all wards)							
Treatment	−0.004	−0.005	0.007	0.004	0.001	−0.003	0.000
	(0.022)	(0.016)	(0.008)	(0.007)	(0.002)	(0.013)	−−
Strata (zone–party) FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	0.552	0.178	0.060	0.024	0.008	0.178	1.000
Control s.d.	0.183	0.130	0.067	0.034	0.013	0.114	0.000
Observations	240	240	240	240	240	240	240
Panel: Spending (slum survey wards)							
Treatment	−0.017	0.018	−0.003	0.001	−0.001	0.002	0.000
	(0.037)	(0.028)	(0.015)	(0.005)	(0.003)	(0.019)	−−
Strata (zone–party) FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	0.542	0.187	0.073	0.019	0.008	0.171	1.000
Control s.d.	0.194	0.142	0.070	0.023	0.012	0.094	0.000
Observations	106	106	106	106	106	106	106
Panel: Slum HH preference							
Treatment	0.005	−0.014	0.001	−0.006	0.026	−0.079	−0.067
	(0.011)	(0.040)	(0.003)	(0.016)	(0.045)	(0.073)	(0.099)
Strata (zone–party) FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	0.020	0.688	0.002	0.046	0.513	1.581	2.851
Control s.d.	0.043	0.180	0.007	0.063	0.195	0.389	0.411
Observations	106	106	106	106	106	106	106

Heteroskedasticity-robust standard errors in parentheses.

Ward-level OLS regression. "Spending (all wards)" is the fraction of total MCD councilor spending (calculated over pre-treatment period) booked for each area. "Spending (slum survey wards)" is an equivalent measure restricted to wards in which we surveyed slum households. Spending is categorized by lexical heuristic. "Slum HH preference" is the ward-mean of households in slum areas who specify each area in response to the question, "In which of the following areas have you personally faced problems in the last year?" (The total is the mean number of areas named by households.) Household responses are weighted within wards to correct for differential coverage of surveys between slums.

Table A.II: Balance table of incumbent electoral outcomes

	2007 Election					2012 Election
	(1) Log registered voters	(2) Log turnout	(3) Seat reserved for minority	(4) Number of candidates	(5) Winner's vote share	(6) Eligible for reelection
Treatment × Ineligible (2012)	−0.051	−0.009	0.072	−0.214	−0.032	
	(0.047)	(0.057)	(0.128)	(1.162)	(0.041)	
Treatment	−0.004	−0.022	−0.060	0.113	0.056	0.040
	(0.032)	(0.038)	(0.086)	(0.767)	(0.034)	(0.067)
Ineligible (2012)	0.002	−0.017	−0.438	0.405	0.013	
	(0.038)	(0.044)	(0.106)	(0.948)	(0.022)	
Control mean	10.500	9.643	0.472	9.472	0.395	0.639
Control s.d.	0.168	0.198	0.503	4.121	0.097	0.484
Observations	240	240	240	240	240	240

Heteroskedasticity-robust standard errors in parentheses.

Ward-level cross section estimated with OLS. "Treatment" indicates observations of a ward in which a report card on the performance of the MCD councilor was published in a newspaper during the 2012 pre-election period (T1 or T2, ITT). "Ineligible (2012)" indicates observations of a ward in which the incumbent councilor became ineligible to run in the same ward in the 2012 elections due to a gender (commonly) or caste (rarely) quota; this is also the dependent variable in column (6). Columns (1)–(5) are dependent variables pertaining to the immediately previous elections in 2007. "Seat reserved for minority" indicates that the seat was reserved for a woman, scheduled caste person, or both, for the 2007–12 term.

Table A.III: Electoral variables of wards by slum density

Variable	(1) Low slum Mean/SE	(2) High slum Mean/SE	T-test P-value (1)-(2)
Incumbent eligible	.66	.68	.78
	(.045)	(.044)	
Incumbent's pro-poor spending index	-.21	-.012	.19
	(.12)	(.09)	
Incumbent's attendance index	.63	.67	.19
	(.018)	(.018)	
Incumbent runs in any ward	.37	.42	.42
	(.045)	(.046)	
Incumbent runs in other ward	.026	.053	.31
	(.015)	(.021)	
Incumbent wins in same ward	.21	.18	.62
	(.038)	(.036)	
Incumbent's vote share in same ward	.14	.13	.71
	(.019)	(.018)	
Incumbent's party wins in same ward	.51	.54	.69
	(.047)	(.047)	
Incumbent's party's vote share in same ward	.37	.37	.71
	(.016)	(.012)	
Voter turnout	.54	.54	.58
	(.0061)	(.0049)	
Voter registration	41715	41193	.7
	(1043)	(840)	
N	114	114	

Heteroskedasticity-robust standard errors in parentheses.

Ward-level cross section. Wards are classified as "low slum" if they have below-median slum fraction by area, and "high slum" otherwise. "Pro-poor Spending Index" is the mean z-score of three log preference-weighted spending amounts (2007–11), analogous to the dependent variable in Table I, column 3, for the pre-publication period. "Attendance Index" is overall councilor attendance at MCD committees of which they are a member (2007–11). All other variables pertain to 2012 MCD elections. "Eligible" indicates observations of a ward in which the incumbent councilor was not rendered ineligible to run in the same ward in the 2012 elections due to a gender or caste quota. "Vote share in same ward" is set to zero if the incumbent does not rerun in that ward.

Table A.IV: Impact of report card on councilor performance index components by slum density

	Pro-poor spending index (2010–2012)	Spending index components			Directly elicited spending index (2010–2012)	Spending index components		Attendance index (2010–12)	Attendance index components	
	(1)	(2) Biggest problem	(3) Problem for individual	(4) Problem for community	(5)	(6) Should spend more	(7) Should spend less	(8)	(9) Assembly	(10) Committees
Treatment × High slum	0.656	1.519	1.338	1.368	0.498	0.548	0.302	0.285	0.052	0.037
	(0.291)	(0.638)	(0.615)	(0.625)	(0.205)	(0.254)	(0.198)	(0.197)	(0.034)	(0.039)
Treatment	−0.325	−0.724	−0.679	−0.689	−0.271	−0.308	−0.218	−0.173	−0.035	−0.014
	(0.198)	(0.432)	(0.420)	(0.427)	(0.146)	(0.183)	(0.153)	(0.146)	(0.024)	(0.030)
High slum	−0.275	−0.640	−0.563	−0.572	−0.200	−0.287	−0.158	−0.124	−0.025	−0.007
	(0.252)	(0.551)	(0.532)	(0.540)	(0.175)	(0.225)	(0.171)	(0.164)	(0.028)	(0.033)
Pre-treat outcome control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Strata (zone–party) FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-treat control mean	0.019	3.196	6.480	6.403	0.010	0.010	−0.030	−0.004	0.816	0.651
Pre-treat control s.d.	0.923	2.034	1.941	1.973	0.952	0.928	0.979	0.912	0.127	0.190
p-value: $T \times High + T + High = 0$	0.797	0.740	0.834	0.818	0.868	0.816	0.626	0.924	0.702	0.587
p-value: $T \times High + High = 0$	0.014	0.010	0.017	0.017	0.010	0.060	0.183	0.165	0.172	0.236
p-value: $T \times High + T = 0$	0.105	0.076	0.125	0.120	0.101	0.165	0.493	0.409	0.480	0.393
Observations	227	227	227	227	227	227	227	227	224	227

Heteroskedasticity-robust standard errors in parentheses.

Ward-level OLS regression. The pro-poor spending index is the mean z-score of log spending on issues, with each issue weighted by the fraction of slum households in the city reporting that (1) it is the most problematic in the area, (2) it is a problem for them, and (3) it is a problem for the community. The directly elicited spending index is the z-score of the inverse hyperbolic sine of spending on issues, with each issue weighted by the fraction of slum households in the city reporting that their councilor should spend more on it minus the fraction reporting that they should spend less on it. The directly elicited spending index components are reported as the z-scores of the inverse hyperbolic sines of spending on issues, weighted with "should spend more" and "should spend less" (the latter with flipped sign). The attendance index is the mean z-score of councilor attendance at (1) the general assembly and (2) councilor committee meetings. "Treatment" indicates observations of a ward in which a report card on the performance of the MCD councilor was published in a newspaper during the 2012 pre-election period (T1 or T2, ITT).

Table A.V: Impact of report card on incumbent's discretionary fund allocation by slum density

	Log spending...					
	(1) ...total	(2) ...on drains	(3) ...on garbage/malba	(4) ...on schools	(5) ...on roads	(6) ...on parks
Treatment × High slum	−0.059	0.700	0.044	−0.549	−0.074	−0.028
	(0.074)	(0.400)	(0.121)	(0.317)	(0.146)	(0.474)
Treatment	0.055	−0.358	−0.101	0.150	0.013	0.474
	(0.056)	(0.279)	(0.085)	(0.221)	(0.104)	(0.352)
High slum	0.007	−0.183	0.021	0.384	0.002	0.089
	(0.065)	(0.345)	(0.102)	(0.268)	(0.125)	(0.411)
Pre-treat spending control	Yes	Yes	Yes	Yes	Yes	Yes
Strata (zone–party) FE	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	4.555	1.943	0.140	0.009	3.865	0.651
Control s.d.	0.248	1.392	0.453	1.122	0.500	1.713
p-value: $T \times High + T + High = 0$	0.959	0.596	0.656	0.943	0.586	0.135
p-value: $T \times High + High = 0$	0.240	0.019	0.289	0.372	0.407	0.813
p-value: $T \times High + T = 0$	0.947	0.217	0.477	0.074	0.543	0.164
Observations	227	227	227	227	227	227

Heteroskedasticity-robust standard errors in parentheses.

Ward-level OLS regression. "Treatment" indicates observations of a ward in which a report card on the performance of the MCD councilor was published in a newspaper during the 2012 pre-election period (T1 or T2, ITT). Amount of spending is in lakh rupees. Spending is categorized by lexical heuristic.

Table A.VI: Impact of report card on councilor performance by ward survey status

	Log total spending (2010–12)		Pro-poor spending index (2010–2012)		Directly elicited spending index (2010–2012)		Attendance index (2010–12)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment × Survey ward		−0.006		0.529		0.381		0.025
		(0.080)		(0.281)		(0.204)		(0.211)
Treatment	0.005	0.006	−0.010	−0.234	−0.020	−0.179	0.030	0.025
	(0.040)	(0.063)	(0.133)	(0.194)	(0.096)	(0.147)	(0.100)	(0.144)
Survey ward		−0.008		−0.084		−0.045		0.039
		(0.068)		(0.240)		(0.170)		(0.170)
Pre-treat outcome control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Strata (zone–party) FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Post-treat control mean	4.552	4.552	−1.854	−1.854	−1.253	−1.253	−0.394	−0.394
Post-treat control s.d.	0.243	0.243	1.053	1.053	0.774	0.774	1.162	1.162
p-value: $T \times Svy + T + Svy = 0$		0.904		0.311		0.315		0.557
p-value: $T \times Svy + Svy = 0$		0.742		0.003		0.003		0.587
p-value: $T \times Svy + T = 0$		0.999		0.130		0.134		0.734
Observations	240	240	240	240	240	240	240	240

Heteroskedasticity-robust standard errors in parentheses.

Ward-level OLS regression. The pro-poor spending index is the mean z-score of log spending on issues, with each issue weighted by the fraction of slum households in the city reporting that (1) it is the most problematic in the area, (2) it is a problem for them, and (3) it is a problem for the community. The directly elicited spending index is the z-score of the inverse hyperbolic sine of spending on issues, with each issue weighted by the fraction of slum households in the city reporting that their councilor should spend more on it minus the fraction reporting that they should spend less on it. The attendance index is the mean z-score of councilor attendance at (1) the general assembly and (2) councilor committee meetings. "Treatment" indicates observations of a ward in which a report card on the performance of the MCD councilor was published in a newspaper during the 2012 pre-election period (T1 or T2, ITT). Spending is categorized by lexical heuristic.

Table A.VII: Impact of report card treatment arm on councilor performance by slum density

	Log total spending (2010–12)		Pro-poor spending index (2010–2012)		Directly elicited spending index (2010–2012)		Attendance index (2010–12)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1: 2012 report (ITT) × High slum		−0.094		0.528		0.378		0.048
		(0.103)		(0.379)		(0.268)		(0.236)
T2: 2010/12 reports (ITT) × High slum		−0.032		0.724		0.557		0.406
		(0.081)		(0.301)		(0.216)		(0.223)
T1: 2012 report (ITT)	0.072	0.135	0.018	−0.238	−0.003	−0.208	0.037	−0.063
	(0.054)	(0.075)	(0.162)	(0.242)	(0.118)	(0.172)	(0.115)	(0.156)
T2: 2010/12 reports (ITT)	−0.031	0.009	−0.024	−0.375	−0.029	−0.306	0.026	−0.237
	(0.042)	(0.061)	(0.143)	(0.205)	(0.105)	(0.154)	(0.112)	(0.176)
High slum		0.005		−0.277		−0.200		−0.127
		(0.065)		(0.253)		(0.176)		(0.166)
Pre-treat outcome control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Strata (zone–party) FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
p-value: $T1 \times High + T1 = T2 \times High + T2$		0.359		0.809		0.662		0.223
p-value: $T1 = T2$	0.042	0.093	0.772	0.475	0.815	0.475	0.923	0.316
Observations	240	227	240	227	240	227	240	227

Heteroskedasticity-robust standard errors in parentheses.

Ward-level OLS regression. The pro-poor spending index is the mean z-score of log spending on issues, with each issue weighted by the fraction of slum households in the city reporting that (1) it is the most problematic in the area, (2) it is a problem for them, and (3) it is a problem for the community. The directly elicited spending index is the z-score of the inverse hyperbolic sine of spending on issues, with each issue weighted by the fraction of slum households in the city reporting that their councilor should spend more on it minus the fraction reporting that they should spend less on it. The attendance index is the mean z-score of councilor attendance at (1) the general assembly and (2) councilor committee meetings. Spending is categorized by lexical heuristic.

Table A.VIII: Impact of private information on public services: spillovers and index components

	Drains outcomes			Toilet availability index	Toilet availability index components			Garbage collection index	Garbage collection index components	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Total drains	Drains with proper disposing (%)	Drains clogged (%)		Log total toilets	Log open toilets	Log adult toilet users (#)		Formal piles regularly collected (%)	Informal piles recently collected (%)
SSI treatment	−0.032	−0.087	−0.048							
	(0.069)	(0.052)	(0.133)							
Post × SSI treatment				−0.048	−0.010	−0.043	−0.091	−0.055	−0.004	−0.038
				(0.031)	(0.026)	(0.024)	(0.067)	(0.042)	(0.089)	(0.044)
Post				0.052	0.041	0.044	0.071	0.332	−0.054	0.416
				(0.023)	(0.020)	(0.020)	(0.050)	(0.028)	(0.065)	(0.026)
Ward FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control mean	0.271	0.136	0.500							
Baseline control mean				0.935	0.723	0.517	1.566	0.256	0.304	0.253
Slums in sample	311	66	66	312	312	312	312	309	110	307
Observations	620	132	132	932	932	932	932	894	328	867

Standard errors clustered by ward in parentheses.

Slum-level OLS regression. "SSI treatment" indicates observations in a slum of which the MCD councilor received State of Sanitation Information (ITT). "Post" indicates observations that took place in the second or third round of audits. "Total drains" is the number of drains in the slum. "Drains with proper disposing (%)" is the fraction of drains from which extracted garbage was taken to a formal garbage pile or to a landfill, rather than left by the drain or burned. "Drains clogged (%)" is the fraction of drains which are so clogged with trash at any point that the water is not visible. "Toilet availability index" is the mean z-score of log total toilets, log open toilets, and log adult toilet users, with first percentile values imputed for zeroes. "Log total toilets" is the log of the number of toilets in the slum. "Log open toilets" is the log of the number of open (i.e. not locked shut) toilets in the slum. "Log adult toilet users" is the log of the mean number of adults who used each toilet in a randomly chosen 15 minute interval between 3–5 PM. "Garbage collection index" is the mean of the fraction of formal piles in the slum regularly collected, relative to the number of formal piles at the baseline, and the fraction of informal garbage piles in the slum collected in the past week. The sample is restricted to slums with one or more garbage piles. "Formal piles regularly collected" is the fraction of formal piles in the slum regularly collected, relative to the number of formal piles at the baseline. "Informal piles recently collected" is the fraction of informal piles in the slum recently collected, relative to the number of informal piles at the baseline.

Table A.IX: Impact of report card treatment arm on incumbent electoral outcomes

	Eligible incumbent runs in any ward		Incumbent runs in other ward			Incumbent's vote share (0 if didn't run)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
T1: 2012 report (ITT)	0.018	0.328	0.035	−0.017	0.067	0.042	0.042	0.168
	(0.115)	(0.413)	(0.032)	(0.019)	(0.073)	(0.034)	(0.048)	(0.166)
T2: 2010/12 reports (ITT)	0.104	0.400	0.041	−0.009	0.023	0.050	0.028	0.038
	(0.093)	(0.346)	(0.025)	(0.021)	(0.082)	(0.029)	(0.037)	(0.141)
T1: 2012 report (ITT) × Ineligible				0.144	−0.665		0.008	−0.470
				(0.086)	(0.482)		(0.059)	(0.223)
T2: 2010/12 reports (ITT) × Ineligible				0.166	0.340		0.032	−0.117
				(0.070)	(0.229)		(0.055)	(0.195)
T1: 2012 report (ITT) × ...		−0.056			−0.040			−0.037
... Pro-poor Spending Index		(0.115)			(0.034)			(0.046)
T1: 2012 report (ITT) × ...		−0.459			−0.137			−0.186
... Attendance Index		(0.604)			(0.113)			(0.232)
T2: 2010/12 reports (ITT) × ...		−0.029			−0.052			−0.043
... Pro-poor Spending Index		(0.110)			(0.033)			(0.040)
T2: 2010/12 reports (ITT) × ...		−0.438			−0.054			−0.008
... Attendance Index		(0.506)			(0.120)			(0.200)
T1: 2012 report (ITT) × Ineligible × ...					0.172			0.110
... Pro-poor Spending Index					(0.067)			(0.051)
T1: 2012 report (ITT) × Ineligible × ...					1.189			0.724
... Attendance Index					(0.721)			(0.310)
T2: 2010/12 reports (ITT) × Ineligible × ...					0.108			0.081
... Pro-poor Spending Index					(0.075)			(0.057)
T2: 2010/12 reports (ITT) × Ineligible × ...					−0.286			0.232
... Attendance Index					(0.306)			(0.271)
Additional Interactions	No	Yes	No	Yes	Yes	No	Yes	Yes
Strata (zone–party) FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ineligible control mean			0.000	0.000	0.000	0.000	0.000	0.000
Eligible control mean	0.478	0.478	0.022	0.022	0.022	0.171	0.171	0.171
Observations	160	160	240	240	240	240	240	240

Heteroskedasticity-robust standard errors in parentheses.

Councilor-level cross section estimated with OLS. Robust standard errors in parentheses. "Attendance Index" is overall councilor attendance at MCD committees of which they are a member (2007–11). "Pro-poor Spending Index" is the mean z-score of three log preference-weighted spending amounts (2007–11), analogous to the dependent variable in Table I, column 3, for the pre-publication period. "Additional interactions" are "Eligible", "Pro-Poor Spending Index", "Eligible × Pro-Poor Spending Index", "Attendance Index", and "Eligible × Attendance Index".

Table A.X: Impact of report card on electoral outcomes using strictly pre-treatment performance

	Eligible incumbent runs in any ward	Incumbent runs in other ward			Incumbent runs in other ward controlled by party			Incumbent's vote share (0 if didn't run)	Incumbent's party's vote share in same ward
	(1)	(2)	(3) *from* high-slum ward	(4) *from* low-slum ward	(5)	(6) *to* high-slum ward	(7) *to* low-slum ward	(8)	(9)
Treatment	0.240	0.064	0.352	−0.035	0.034	0.137	0.029	0.019	−0.068
	(0.306)	(0.083)	(0.228)	(0.115)	(0.059)	(0.121)	(0.136)	(0.122)	(0.081)
Treatment × Ineligible		0.136	−0.403	0.218	0.094	−0.277	0.606	−0.161	−0.161
		(0.239)	(0.309)	(0.382)	(0.211)	(0.240)	(0.472)	(0.168)	(0.158)
Treatment × ...	−0.098	−0.030	−0.081	−0.010	0.002	−0.019	0.014	−0.061	−0.032
... Pro-poor Spending Index	(0.101)	(0.026)	(0.078)	(0.028)	(0.013)	(0.026)	(0.047)	(0.037)	(0.021)
Treatment × ...	−0.228	−0.118	−0.439	0.033	−0.046	−0.138	−0.068	0.021	0.096
... Attendance Index	(0.445)	(0.126)	(0.302)	(0.187)	(0.079)	(0.147)	(0.181)	(0.170)	(0.114)
Treatment × Ineligible × ...		0.123	0.190	0.097	0.070	0.128	0.042	0.124	0.098
... Pro-poor Spending Index		(0.050)	(0.112)	(0.071)	(0.043)	(0.093)	(0.082)	(0.047)	(0.037)
Treatment × Ineligible × ...		0.047	0.753	−0.151	0.046	0.528	−0.653	0.297	0.279
... Attendance Index		(0.348)	(0.459)	(0.550)	(0.313)	(0.393)	(0.649)	(0.232)	(0.229)
Additional Interactions	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Strata (zone–party) FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ineligible control mean		0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.408
Eligible control mean	0.478	0.022	0.040	0.000	0.000	0.000	0.000	0.171	0.347
Observations	160	240	114	114	240	113	116	240	240

Heteroskedasticity-robust standard errors in parentheses.

Councilor-level (1–8) and ward-level (9) cross sections estimated with OLS. Robust standard errors in parentheses. "Attendance Index" is overall councilor attendance at MCD committees of which they are a member (2007–10). "Pro-poor Spending Index" is the mean z-score of three log preference-weighted spending amounts (2007–10), analogous to the dependent variable in Table I, column 3, for the pre-treatment period. "Additional interactions" are "Eligible", "Pro-Poor Spending Index", "Eligible × Pro-Poor Spending Index", "Attendance Index", and "Eligible × Attendance Index".

Table A.XI: Impact of report card on electoral participation

	Voter turnout			Voter registration		
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment	−0.008	−0.008	−0.076	−744.579	−702.669	4437.303
	(0.007)	(0.009)	(0.033)	(1420.318)	(1941.576)	(11217.330)
Treatment × Ineligible		−0.001	0.102		−460.830	−6678.261
		(0.015)	(0.061)		(2519.533)	(13468.438)
Treatment × ...			−0.009			1931.597
... Pro-poor Spending Index			(0.013)			(2687.108)
Treatment × ...			0.104			−7624.440
... Attendance Index			(0.053)			(15964.006)
Treatment × Ineligible × ...			−0.004			−1552.976
... Pro-poor Spending Index			(0.017)			(3397.179)
Treatment × Ineligible × ...			−0.160			9553.231
... Attendance Index			(0.094)			(19916.529)
Nontreatment Interactions	No	Yes	Yes	No	Yes	Yes
Strata (zone–party) FE	Yes	Yes	Yes	Yes	Yes	Yes
Ineligible control mean	0.534	0.534	0.534	40226.3	40226.3	40226.3
Eligible control mean	0.554	0.554	0.554	43269.2	43269.2	43269.2
Observations	240	240	240	240	240	240

Heteroskedasticity-robust standard errors in parentheses.

Ward-level cross section estimated with OLS. "Attendance Index" is overall councilor attendance at MCD committees of which they are a member (2007–11). "Pro-poor Spending Index" is the mean z-score of three log preference-weighted spending amounts (2007–11), analogous to the dependent variable in Table I, column 3, for the pre-publication period.

Figure A.1: Map of Delhi wards by treatment arm