**Appendix for Online Publication**

The Appendix provides additional information on the context of the Bogotá metro project, the survey format, summary statistics and balance tests, additional results, and the sensitivity analysis.

1. **Context**

One way to situate the Bogotá case is to look at subways globally.[[1]](#footnote-1) I built an original database of all subways. I define a ‘subway’ as an electric powered urban rail that is completely isolated from interactions with automobile traffic and pedestrians, following Gonzalez-Navarro and Turner (2016). This excludes most streetcars and BRT networks because they interact with vehicle traffic at stoplights and crossings. Some combine both elements making them difficult to code. If they contain some segregated element, they were included. To focus on intra-urban transportation systems, I also exclude heavy rail commuter lines. I do not distinguish between surface, underground, or aboveground subway lines because the exclusive right of way condition is satisfied. These data come from http://www.urbanrail.net/ and are supplemented with information on each subway from news sources where necessary. For the purposes of this article, I simply code whether a subway exists in each city, regardless of length, opening date, etc.

Identifying the relevant universe of cities also was a challenge. The UN World Cities Report provides data on urban agglomerations with 750,000 inhabitants or more.[[2]](#footnote-2) This data set includes city populations in 2015. The drawback of the UN database is that it is limited to quite large cities. Especially within the OECD, many smaller cities also have metro systems. For instance, 53 of 245 cities with metros (22%) do not have information in the UN Cities Database. I therefore supplement the UN data with OECD statistics on metropolitan areas.[[3]](#footnote-3) The OECD defines a metropolitan area as containing 500,000 people or more, and cities often have around 100,000 inhabitants. Country wealth data comes from the World Bank Development Indicators.

Figure A.1 plots the relationship between city size (on the horizontal axis) and country wealth (on the vertical axis). What stands out is that wealthy countries are far more likely to have subways even in quite small cities, as seen in the upper-left quadrant. While 90 percent of cities above five million people in developed countries have subways, only 32 percent of similarly sized cities in the developing world have subways.[[4]](#footnote-4) As cities grow, they are more likely to have subways, as in the upper-right quadrant. Not until cities reach populations above 20 million are they almost guaranteed to have a metro in the developing world.



Figure A.1 Subways by City Population Size and GDP per Capita (2015)

Sources: Author’s Database compiled from Urban Rail Net and UN World Cities Report (2015).

Turning to the Bogotá subway project, I asked several questions on my survey to probe general expectations about the subway in Bogotá summarized in Table A.1. Most respondents had positive impressions that the subway would reduce their commute length and traffic in the city, as well as raising property values. However, as in many cities, Bogotá residents expect substantial corruption in the metro project, doubted when construction would begin, and thought that they would pay for the project through higher taxes. There also was uncertainty about the project: about half of respondents believed that the metro would be under construction in five years.

**Table A.1. Expectations for the Bogotá Subway Project**



Notes: Responses are coded as ‘1’ if respondent thought it was ‘likely’ or ‘highly likely’ that a condition would hold and ‘0’ otherwise. N=907.

1. **Survey Design and Sampling Procedures**

All survey protocols and questionnaires were approved by Harvard University’s IRB (#15-3150) and reviewed by local academics and the partner survey firm, Centro Nacional de Consultoría, in Colombia. The survey was administered to 900 respondents living in Bogotá, Colombia in August to October 2016.

I used a stratified sample to guarantee an adequate sample size living within each geographic area. Following standard practice in urban planning, I divided the city into three zones shown in Figure 1 of the article. The survey firm then selected city blocks within each transportation zone using government-defined class stratifications. Colombia divides households into six ‘strata,’ which often are grouped together with strata 1 and 2 as low-income households (‘lower’), strata 3 as lower-middle (‘middle’), and strata 4 and above as upper-middle class (‘upper’). Table A2 shows the division of class groups in the sample and their relationship to Bogotá’s population. This method of respondent recruitment was preferable to an online convenience sample because it covers the entire city and different class groups. I conducted a pilot using a Facebook sample but recruited a much younger, wealthier, and more centrally located sample than possible through direct recruitment in different neighborhoods.

**Table A.2: Sampling Procedure**

|  |  |  |  |
| --- | --- | --- | --- |
| Geographic Zone | **Lower Class**  (Strata 1 and 2) | **Middle Class**  (Strata 3 and 4) | **Upper Class**  (Strata 5 and 6) |
| **Unaffected zone** | 131 (121) | 130 (112) | 40 (30) |
| **Secondary influence zone** | 132 (123) | 135 (127) | 39 (35) |
| **Direct influence zone** | 130 (124) | 130 (121) | 40 (35) |
| *Total in Sample (907)* | *393 (368)* | *395 (360)* | *119 (100)* |
| *As Percent of Bogotá’s population* | 48.8% | 35.7% | 13.8% |

Notes: Software glitches meant that some QV responses were not recorded. The numbers indicate the completed surveys on demographic questions. The numbers in parentheses indicate the completed survey on both demographic and QV questions.

The QV instrument must be administered online, which created some logistical challenges in a city with spotty internet access and high rates of crime. Security concerns also meant that enumerators were unwilling to carry portable electronic devices in all areas of the city. But Internet cafes are widespread. Rather than conduct a household survey, enumerators approached respondents in public spaces and at starting places on the selected blocks. They first verified that the respondent lived in the defined area around the metro project and their class stratification. They then recruited them to take the survey in an Internet café set up to administer the survey. To encourage respondents to take the survey, they were offered a small gift for completing the survey. The preliminary screening interview ensured that respondents met the distance and demographic criteria even in more dangerous neighborhoods. Response rates were 18 percent, which is somewhat lower than face-to-face surveys conducted in Bogotá by the same survey company but higher than most telephone surveys. The survey firm verified that observable demographic features are like the features of the underlying city population.

Even using the Internet cafes to increase the chances of a reliable Internet connection, software problems emerged in recording the QV responses on the first module of the survey. The numbers in parentheses in Table 3 indicate the number of surveys with complete responses on the QV/Likert conditions and experiment. Luckily, there was no systematic missingness to induce bias into the survey responses. But the lost data reduce the power to detect significant effects, especially in the subgroup analyses.

The experimental vignettes were assigned through simple randomization. Block randomization is even more powerful but wasn’t possible given that the QV software required first administering the specialized QV modules and then directing respondents to demographic and political questions. Power calculations were complicated given the novelty of the survey technique and subject matter, which gave me little sense of a reasonable effect size and standard deviation. I simulated the different sample sizes required using <https://egap.shinyapps.io/power-app/> in which I looked for a significance level of 0.5 (alpha), assumed a standard deviation of 2, a power target of 0.8. Given four experimental groups, 225 respondents received each control or treatment. I have statistical power to distinguish only relatively large heterogeneous treatment effects (i.e., different reactions to the experimental treatments between renters and owners, affected and unaffected areas). The high costs of geographic sampling and recruitment limited the ability to increase the sample size.

Survey enumerators were trained to assist the respondents with any questions regarding the instructions and implementation of the survey. An alternative design would be to use a linear budget to assign priority across issues. Lalley and Weyl (2018) show formally that a quadratic price results in a more accurate reporting of preference intensity relative to a linear price. However, linear budgets are easier to understand and have been used successfully in survey research. For instance, Altamirano et al. (2020) ask respondents to assign 100 points across pensions, public security, health, and education to understand the priorities assigned to different spending issues in high-crime settings in any way that they like. The main risk is that respondents allocate all their points to one issue and zero to the rest. No direct comparisons exist of quadratic and linear budget survey questions, though Casella and Sanchez (2019) compare QV to linear vote systems in referenda settings and find advantages to both. Both systems create better incentives for voters to express the intensity of their preferences. In a referenda context, QV achieves higher average welfare, but also leads to inefficient minority victories.

Initial focus groups highlighted the need to change the terminology around QV. QV was set up to involve the ‘purchase of votes,’ but I instructed respondents only to ‘allocate credits’ across questions, given the largely condemned practice of vote buying in Colombia.

1. **Descriptive Statistics and Balance Tests**

Table A.3 presents descriptive statistics for the main dependent and independent variables.

Table A.3. Descriptive Statistics

****

Table A.4 presents the balance tests on the key covariates across the two experiments based on rush hour assignment and information treatments. The results are based on two-sided t-tests with unequal variances, as well as an F-test with all covariates. Age is statistically different in the logrolling and value capture treatment assignment, but age has no association with the outcome and the results are unchanged controlling for all observable variables.

Table A.4. Balance Tests for Rush Hour and Informational Experiments

****

1. **Alternative Explanations**

Table A.5 considers additional explanations based on property value expectations, transit use, and government attitudes. First, it’s possible that respondents simply were unsure about how the metro would affect property values. Model 1 shows that property owners generally behave as expected: owners in the closest zone are more likely to expect that their property value will increase. Models 2 and 3, however, demonstrate that this expectation doesn’t shape attitudes towards the metro more broadly. Model 2 indicates that nearby property owners are no more supportive of the metro, and Model 3 shows that owners that expect their property values to rise are no more supportive of the metro. These results provide further evidence that geography and property ownership doesn’t seem to be shaping preferences as expected. An important limitation is that the sample covers individuals who live at their residence (homeowners), rather than landlords or businessowners who may be less worried about construction nuisances or have different expectations for property value rises.

Table A5. Expectations of Property Value Change and Metro Support



Notes: OLS regressions, standard errors are included in parentheses, \* indicates significance at the 95 percent level.

Model 4 turns to theories based on existing transit use. Redistributive theories can be based on the uneven distribution of costs of projects or the uneven benefits they provide. One way to think about the benefits is by asking individuals about their current commute. Car owners might be less supportive of the metro, given their access to alternative means of transportation. Model 4 includes an indicator variable for whether an individual owns a car (*Car*). I also look at whether reliance on the existing bus system (*Public Transit)* or informal transport providers (*Informal Transit*) is associated with less support for the metro, given possible crowding out of investments in these alternative transit networks. Car ownership doesn’t predict support for the metro, which is consistent with an externalities explanation in which wealthy car owners don’t intend to take the metro. However, public transit use is associated with greater support for the metro, suggesting that lower-class transit users are more likely to support the metro.

An alternative theory looks at how trust in government conditions support for public works projects. Model 5 examines whether attitudes toward government or the mayor in power drive support. To capture anti-government views, I look at a standard question that asks the respondent to choose between “the less government, the better” or “there are more things that the government should be doing” (*Anti-Government*). Wanting a smaller government is somewhat different than trust in government. I also probe whether individuals think that the government will lose money to corruption in the metro project *(Corruption*) and whether respondents voted for the mayor in office (*Voted Mayor).* Individuals who support a larger government role in general, perceive less corruption, or supported the current mayor should be enthusiastic about the metro.

Attitudes towards government do condition support toward public works, as the literature would predict. Anti-government attitudes predict less support for the metro. But anti-government attitudes also are rare: in Bogotá, just 20 percent of the sample agrees that less government would be better. By comparison, this question draws support from 59 percent of American voters.[[5]](#footnote-5) This discrepancy suggests that theories rooted in American conservatism may have less leverage in understanding public works in the developing world. In contrast, concerns about corruption in project implementation are widespread, with two-thirds of respondents suspecting a loss of funds. But corruption concerns are not associated with support for the metro. Voting for the elected mayor is associated with greater support for the metro. This is somewhat surprising, given that the mayor at the time of the survey (Peñalosa) turned against the underground subway project.

1. **Sensitivity to the Choice Set**

An important concern with QV is that the question responses are sensitive to the choice set provided. For instance, it might be the case that low-income individuals place all their votes on the peace process, but once this issue is off the political agenda in Colombia, the response patterns would look quite difference and the poor would support infrastructure or social policy or any other issue. A robustness check therefore involves removing certain issues from the choice set and rescaling the responses to understand what preferences look like on a subset of the issues analyzed. As long as the questions were included in the choice set, I can take a non-linear transformation of the data to understand how much individuals care about a given issue relative to a subset of those explained. This method assumes the independence of irrelevant alternatives, meaning that choices depend only on the issue at stake and not on the characteristics of other items in the choice set. This assumption could be violated in practice. As often is argued in behavioral research, perceptions of one item are often linked to perceptions of others, leading responses to depend on other, correlated attributes. Future research may examine whether this assumption is violated in practice in the spirit of Bansak et al.'s (2018) work on changing the number of attributes in a conjoint experiment. For the moment, I concentrate on the formal analysis and assume the independence of irrelevant alternatives.

In general terms, consider the set of propositions initially included in the QV module with a budget of *B*. For the full question set, individual allocates their votes on subjects such as denoted by to the constraint that

Now suppose we are interested in the subset of issues , where . In the survey, each individual *j* will have allocated a total number of credits to the issues in equal to

By the assumption that, regardless of budget, individuals vote across issues in proportion to their utilities, under any rescaling of the budget individuals would simply have scaled up or down their votes.[[6]](#footnote-6) Using this assumption, I can derive what individuals would have done had they been presented only with a subset of the issues . In particular, I suppose that everyone had a budget of to vote on the issues . Then, there exists a scaling factor such that *j*’s votes would have been solving:

So, all I need to do to obtain the counterfactual votes in this setting is scale up each individual’s votes on the issues in by . Another way to think of this is that there is a Euclidean length to all the issues that an individual cares about, and this scaling factor extends the Euclidean length as appropriate for each individual by the amount that is removed by taking some elements out of the choice set so that each individual in the survey has the same budget for the subset of issues under consideration.

Again, rescaling cannot account for omitted issues or framing effects. In other words, this formula works for any hypothesis that we have about what public opinion looks like when removing elements from the choice set; it does not allow us to test hypotheses about scenarios in which different issues are added to the choice set. It also assumes that there are no framing effects caused by seeing a proposition in the choice set. It is possible that the mere presence of an issue in the choice set makes respondents think differently about other issues. For instance, including a question about taxation may lead the rich to voice more opposition to other forms of government spending than would be the case if no questions were presented on taxation.

To give an example, I may be interested in focusing on the trade-off between the metro project and social expenditures relative only to each other. If so, imagine that an individual had placed 5 votes on the metro and 3 votes on redistribution. They spent a total of 34 credits () on those two issues. We then calculate their individual following the formula derived above as =1.7, which when scaled up implies that they placed 8.5 votes on the metro and 5 votes on redistribution. We would repeat the same procedure for each individual in the sample and then be able to make statements about the relative choice between transit and redistribution.

There will be different views on which comparisons represent the best model of politics. Surveys using Likert long assumed that there were no trade-offs on the table. Much theorizing remains to be done on how to best conceptualize the meaningful relative comparisons under QV. But this scaling approach suggests a tractable way to examine the way that different issues are linked across the presented choice set.

Applied to this paper, I removed the redistributive issues from the choice set. These included three questions on poverty relief, taxes on the rich, and health and education expenditures. Table A.5 shows the results from removing each of these issues on support for the metro. Model 1 replicates the core model used in the paper prior to removing any issues. Models 2 through 4 remove each issue individually, and Model 5 shows the results to removing the trio of redistributive issues. The most interesting fact is that class differences on the metro actually are strengthened by the removal of redistributive issues. It therefore does not seem like class cleavages around the metro are simply the byproduct of low-income voters prioritizing redistributive issues.

Table A.6 considers whether class cleavages on other issues drive the metro results. Model 1 through 3 disaggregate the redistributive issues bundled together in the paper into their component parts. Class doesn’t predict the votes places on any of the redistributive issues, so can’t induce the divisions around the metro project under QV. Models 4 and 5 show the two issues on which class differences emerge on the QV survey, street vending and corruption. In both cases, upper-class groups put more votes on the issue than lower-class respondents. Models 6 and 7 use the sensitivity analysis to remove street vending and corruption from the choice set and reweight the votes placed on the metro. The dependent variable then is the reweighted votes on the metro. Removing these issues from the choice set and reweighting the data strengthens the class differences on the metro. It therefore doesn’t seem like class differences on other issues included on the choice set are driving the metro results.

Table A6. Class Differences on Other Issues in the Choice Set



Notes: OLS regressions, standard errors are included in parentheses, \* indicates significance at the 95 percent level.

**References**

Altamirano M, Berens S, and Ley S (2020) Between Security and Social Policy Spending. Working Paper.

Bansak K, Hainmueller J, Hopkins DJ, and Yamamoto T (201) The Number of Choice Tasks and Survey Satisficing in Conjoint Experiments. Political Analysis 26(1), 112–19.

Cavaillé C, Chen DL, and Van der Straeten K (2019) A Decision Theoretic Approach to Understanding Survey Response: Likert vs. Quadratic Voting for Attitudinal Research. The University of Chicago Law Review 87, 22–43.

Gonzalez-Navarro M and Turner MA (2016) Subways and Urban Growth: Evidence from Earth. Working Paper.

Lalley SP and Weyl EG (2018) Quadratic Voting: How Mechanism Design Can Radicalize Democracy. American Economic Association Papers and Proceedings 108, 33–37.

1. The subway database that I constructed is available through the article’s Dataverse, GlobalSubways.csv. [↑](#footnote-ref-1)
2. Available at https://unstats.un.org/unsd/demographic/products/dyb/dyb2015.htm [↑](#footnote-ref-2)
3. Available at https://stats.oecd.org/Index.aspx?DataSetCode=CITIES [↑](#footnote-ref-3)
4. ‘Developed’ is defined by the World Bank’s high-income threshold of a GDP per capita of $12,475 in 2015. [↑](#footnote-ref-4)
5. This estimate comes from ANES, VCF9131, 2012. [↑](#footnote-ref-5)
6. For a more extensive model of this property, see Cavaillé et al. (2019). [↑](#footnote-ref-6)