Using Cell-phone Mobility Data to Study Voter Turnout*

Masataka Harada[†], Gaku Ito[‡], and Daniel M. Smith[§]

October 7, 2022

Abstract

Scholars of voting behavior are often confronted with poor data availability or unsuitably large units of aggregation for reported turnout. We demonstrate a big-data solution to this challenge, using fine-grained cell-phone mobility data on millions of GPS locations for more than 300,000 eligible voters in Tokyo. Our approach uses the geolocations of polling stations, combined with GPS data points recorded on election day and a reference day, to measure patterns in individual-level (but anonymized) voting behavior. We first test the validity of the measure by comparing it to official aggregated data on turnout. We then demonstrate the measure's substantive utility with two applications exploring the relationships between turnout decisions and (1) the cost of voting (proxied by distance to the polling station); and (2) the extent of neighborhood-level damages sustained during World War II bombings, building on an emerging literature on the long-run effects of political violence. (147 words)

Keywords: voter turnout, GIS, GPS, cell-phone mobility, Japan
Word count: 11,995 words, excluding appendix

*Prepared for the Toronto Political Behaviour Workshop, November 4-5, 2022. We thank Zachary Steinert-Threlkeld, Paul Lendway, Øyvind Skorge, and Kosuke Imai for helpful feedback. This work was supported by the Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research, KAKENHI (Grant Numbers: JP15K16977, JP18KK0039, JP19K21681, JP20H01466, and JP20K01464), with additional financial support (No. 187102) from the Central Research Institute of Fukuoka University.

[†]Associate Professor, Fukuoka University. Email: masatakaharada@gmail.com.

[‡]Associate Professor, Osaka Metropolitan University. Email: gaku@omu.ac.jp.

[§]Gerald L. Curtis Visiting Associate Professor of Modern Japanese Politics and Foreign Policy, Columbia University. Email: dms2323@columbia.edu.

Voter turnout has long been considered one of the most important barometers of the health of democracy (e.g., Franklin, 2004; Lijphart, 1997; Tingsten, 1937). For democratic elections to accurately represent the will of the people, the people need to show up to the polls. In close elections, small changes in turnout can even determine which party or candidate wins, and consequently, which public policies get implemented (e.g., Fowler, 2013; Hansford and Gomez, 2010; Martinez and Gill, 2005). Reflecting this central importance to both democratic theory and the practical goals of politicians seeking office, an immense comparative literature aims to explain not only which citizens are more likely to vote (e.g., Brady et al., 1995; Fowler et al., 2008; Loewen and Rubenson, 2019; Wolfinger and Rosenstone, 1980), but also why there is variation in participation across different contexts (e.g., Anzia, 2014; Blais, 2006; Cancela and Geys, 2016; Cox et al., 2016).

A key challenge confronting research in this area is the fact that fine-grained (especially individual-level) measures of turnout are often limited in availability due to privacy concerns. Only a few governments disclose the administrative records of voters (e.g., Bhatti et al., 2012; Martikainen et al., 2005), and these are often difficult or costly to obtain.[1] Meanwhile, self-reported turnout in surveys tends to be inaccurate, due to social desirability bias and over-reporting (e.g., Ansolabehere and Hersh, 2012; Brockington and Karp, 2005; Holbrook and Krosnick, 2010). In most cases, official turnout data are only disclosed after some amount of aggregation, often at the level of entire municipalities or counties. To obtain data at smaller units of aggregation, such as polling stations, researchers may need to make cumbersome information disclosure requests to local governments.

When turnout data are not optimally (dis-)aggregated to test a given hypothesis, researchers risk exposure to a number of statistical problems, including a reduction in statistical power, the ecological inference problem and aggregation bias (King, 2013; Matsusaka and Palda, 1993), and the modifiable areal unit problem (Fotheringham and Wong, 1991). As a result, the limited availability of micro-scale turnout data may force researchers to give up

[1] Studies of voting behavior in the United States can use voter files compiled by private companies (e.g., Nyhan et al., 2017).

on linking many geographical factors of interest to their outcome variables, or to compromise their research designs with a suboptimal unit of analysis.

Recent advances in big data provide an opportunity to creatively overcome these kinds of challenges. Researchers can now construct treatment and outcome variables based on geographic information contained in previously underutilized data sources, examine previously untestable hypotheses, and provide new insights into enduring questions (e.g., Moore and Reeves, 2020). These advances include (1) the ability to quantify spatial information such as maps, aerial photography, and satellite photography and sensors; and (2) the availability of mobility data collected from Global Positioning System (GPS) satellites, Wi-Fi spots, and cellular phone base-stations.

In this study, we introduce and explain a method for generating an alternative measure of voter turnout from cell-phone mobility data when official administrative data are unavailable at the desired unit of analysis, and then demonstrate its validity and substantive utility with two applications of interest to political scientists. The key idea behind our approach is that cell-phone users can be considered to have voted if they approached the location of a designated polling station during voting hours on an election day. With some limitations we will describe, measuring turnout with this approach allows researchers to create a proxy for voting behavior at smaller units of aggregation, such as neighborhoods, and even for individual voters.

While we make a unique contribution to the literature and methods for studying turnout, our approach builds on a growing number of studies that use cell-phone mobility data to answer substantive questions of importance to political science and other disciplines (e.g., Chen et al., 2019; Rotman and Shalev, 2020; Sobolev et al., 2020). Some existing studies involve the summation of GPS signals representing pedestrian traffic aggregated at points of interests such as major train stations (e.g., Google, 2020). Others record any movement of cell-phone signals and regard this as a measure of users' activity. This approach has been used to analyze citizens' behavior under the stay-at-home mandates of COVID-19 pandemic (e.g., Clinton et al., 2021; Jay et al., 2020). However, people go to train stations or move about town for many reasons—so in many cases it is difficult to attribute any particular meaning to users' mobility patterns.

Another stream of research complements the semantic vacuum of mobility data by using geo-tagged data from social network services (SNS) such as Twitter and Weibo to analyze the reasons behind a user being in a given location. In these analyses, if an SNS user tweets that he or she is engaging in some activity, the location from which the user tweeted is assumed to be where the activity took place (e.g., Hobbs and Lajevardi, 2019; Steinert-Threlkeld, 2017). This approach opens up the possibility of providing mobility data with broader meaning, but still relies on stated information (i.e., tweets) for the content of the activity. Moreover, SNS users are concentrated in younger generations and tend to have opinions and behavior that are distinct from the general voting population.

Our mobility data come from the traffic records of cell-phone users on an election day and a reference day in the 23 special administrative wards comprising the central metropolitan area of Tokyo, Japan. The reference day data allow us to apply a difference-in-differences (DID) design to estimate voting behavior (e.g., Nunn and Qian, 2011). Altogether, the mobility data contain over 40 million entries for about 300 thousand unique user IDs. We first discuss how we remove "noise," or miscoded turnout, from the data. Then, we tune the parameters for data processing through validating our measure using administrative turnout records and census-based population counts aggregated at the level of a voting precinct (roughly the size of several neighborhoods combined).[2] These tuning and validation processes show a stable performance of the measure with changing parameters, and a strong positive correlation between the estimated population counts and turnout from our data and the actual census population counts and recorded turnout at the precinct level.

[2]In Japan, the smallest administrative unit for which aggregated voter turnout is available is the voting precinct (*tōhyō-ku*) which consists of several neighborhoods. Turnout records at the polling station level are only provided upon request in most municipalities, and some neighborhoods are divided into a number of arbitrarily smaller areas, each of which belongs to a different voting precinct. As a result, using these data in combination with other statistics is a challenge.

We next present two applications to demonstrate the substantive utility of our approach for studying important questions of interest at smaller units of analysis. Our first application takes on the rational voter model (Aldrich, 1993; Downs, 1957; Riker and Ordeshook, 1968) to explore the individual-level relationship between the cost of voting (proxied by the distance to a designated polling station) and the decision to vote. Our analysis using estimated turnout based on cell-phone traffic confirms that election-day turnout is indeed lower for individuals whose designated polling stations are further from their homes, corroborating existing evidence from various contexts based on alternative measures (e.g., Bhatti, 2012; Cantoni, 2020; Garnett and Grogan, 2021; Gibson et al., 2013; Haspel and Knotts, 2005; Nishizawa, 1991). The analysis is followed by two types of sensitivity analyses to gauge the impacts of misspecified tuning parameters and miscoded cell-phone users living especially near to a polling station.

Our second application investigates the relationship between wartime destruction and present-day voter participation. Controversy persists over whether exposure to political violence or wartime destruction increases or decreases the short-term and long-term propensity to vote (e.g., Bellows and Miguel, 2009; Blattman, 2009; Gilligan et al., 2014; Kage, 2021; Lupu and Peisakhin, 2017). We contribute additional empirical evidence to this debate by combining our mobility data with detailed data on neighborhood-level damages caused by the firebombing of Tokyo during World War II (Harada et al., 2021). Our mobility-based estimates show significantly lower turnout in the neighborhoods that were most damaged by the firebombing, a long-term negative effect that runs counter to some of the existing evidence of positive effects of war violence on turnout, but which is consistent with evidence that wartime destruction lowers social capital and other socioeconomic indicators of neighborhood well-being (Harada et al., 2021).

Data and Methods

This section explains our data sources and methods for constructing our GPS-based estimate of voter turnout.

Polling Station Data

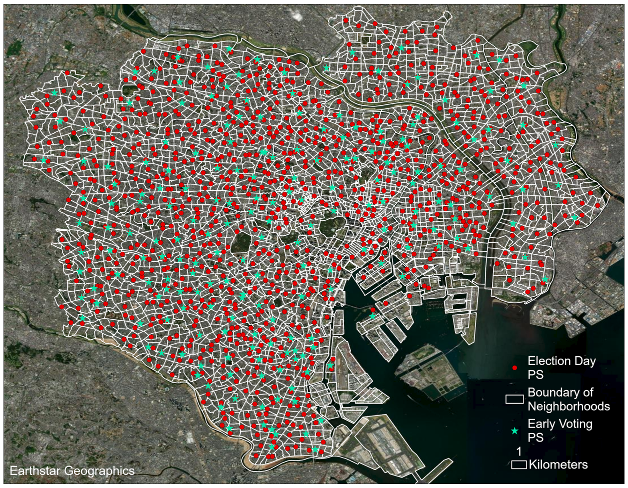
Information on polling stations was collected from the election administration commissions for each of Tokyo's 23 wards or their websites. Specifically, we requested the IDs and names of all polling stations, which households in neighborhoods are assigned to each station, the number of registered voters, and turnout. We then used a correspondence table between neighborhoods and polling stations to refine the prediction. The last two items (registered voters and turnout) are used for validation purposes. The level of detail for the location of polling stations differs across wards.

The locations of the polling stations were identified in the following way. First, we manually searched the location of each polling station from its name (and address whenever available) using Google Maps. Except for several cases, the locations were uniquely identified, and when multiple entries were found for a single name, we collected further information about the polling station, such as its address. When the buildings did not appear in Google Maps due to reconstruction, we referred to aerial photography from several years ago.[3] Through these processes, all of the geographical locations of polling stations were identified.

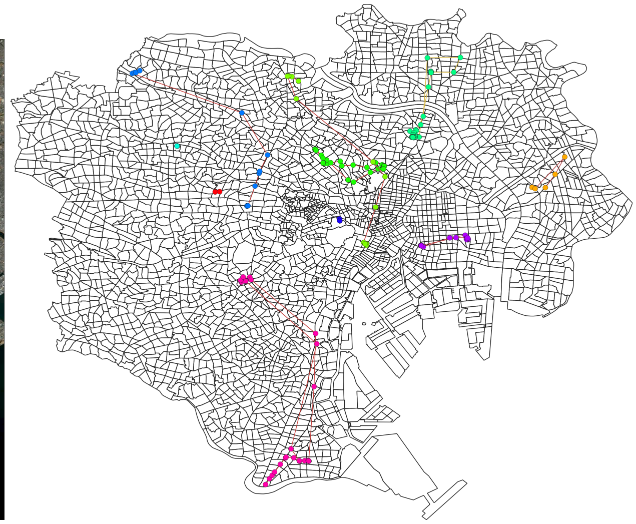
We use the coordinates of the polling stations as the center of a circle encompassing the building objects, and whether a cell-phone user enters within this circle is the basis for our voting measure. The coordinates of the polling stations were therefore manually measured at the midpoint of the longest diameter for each building identified above to minimize the distance to the furthest edge of the building.[4] This approach is also easier to administer

[3]Specifically, we referred to the Geospatial Authority of Japan: https://maps.gsi.go.jp.

[4]Some wards provided only the name of the facility (e.g., ABC elementary school), while other wards provide the exact pinpointed location of the facility (e.g., the gym of XYZ elementary school). In identifying the location (coordinates) of polling stations, we utilized the most detailed information available.



(a) Location of polling stations



(b) 10 sampled tracking records

Figure 1: Illustration of data: polling stations and cell-phone mobility in Tokyo
Note: In the left panel (a), the red and green points represent the locations of election-day polling stations and early-voting polling stations, respectively. The right panel (b) illustrates the cell-phone mobility of 10 sampled users. Both panels were created based on data for the House of Councillors election on July 21, 2019.

than manually calculating the geometric center of a convex hull for each building.[5] The 1,118 election-day polling stations and 205 early-voting polling stations for the 2019 upper house (House of Councillors) election are shown in red and green, respectively in Figure 1(a). We see that the polling stations are located across all 23 wards.

Mobility Data

We obtained cell-phone mobility data (called "fluid population data") from Agoop, a subsidiary company of Softbank, which is the third largest cell-phone carrier in Japan (Agoop Corporation, 2020). This data set consists of the coordinates obtained from the onboard GPS of iOS and Android devices. The coordinates are recorded every five minutes, but only when the device is turned on (for Android devices), or a user explicitly allows the software

[5]Geocoding with Google Maps returns the stored coordinates associated with building names and addresses, but many of them are not close to the coordinates obtained with our approach.

Table 1: Number of daily users and traffic records during the election day and reference day

Date	July 21, 2019	July 28, 2019
Type of Day	Election Day	Reference Day
Number of Daily Users (Original)	158,338	151,725
Number of Traffic Records (Original)	20,811,675	20,151,678
Number of Daily Users (Filtered)	100,906	96,172
Number of Traffic Records (Filtered)	9,336,810	8,962,007

Note: The following criteria were used as filters: Japanese citizens, residents of Tokyo's 23 wards, and signal accuracy below 100 meters.

to use GPS (for iOS devices).[6] To protect privacy, user ID is assigned daily, and GPS entries that help identify users' pinpoint addresses are eliminated by Agoop.[7] Moreover, no user-level demographic information is provided.

To estimate voter turnout, we purchased data on the foot traffic records made within Tokyo's 23 wards on selected days and filtered the data entries of non-residents. Specifically, for our election day, we selected the House of Councillors election held on Sunday, July 21, 2019, prior to the onset of the COVID-19 pandemic. We also selected the following Sunday, one week after the election, as a reference day to control for the regular foot traffic of non-election days. If some cell-phone users regularly pass by a polling station as part of their daily routine, some of their signals will be recorded within a predetermined radius from the station and may therefore be mistakenly counted as votes cast. If their destination is close to the polling station, we expect more miscoding.[8] The basic characteristics of the data are presented in Table 1. Both days have more than 150,000 original daily IDs per day, and the numbers of traffic records recorded every five minutes are over 20 million.[9]

[6] Users agree to provide their GPS information when installing some applications (typically restaurant/healthcare/fitness applications).

[7] Specifically, Agoop drops all traffic records that fall within the 100-meter grid that contains the user's address.

[8] We explain later how we deal with this potential noise in our measure. For applications aimed at estimating cross-sectional differences in turnout across neighborhoods, this possible source of miscoding should be less of a concern.

[9] We initially purchased the data sets of two other elections in 2017. However, the data quality (such as the length of each tracking point) has significantly improved since 2017 because of the introduction of new software to collect GPS information. We therefore use only the data recorded in 2019.

Table 2: Summary of weather on election day and reference day

Date	June 30, 2019	July 21, 2019	July 28, 2019	August 4, 2019
Type of Day	*Neither*	Election	Reference	*Neither*
Weather	Rainy occasionally cloudy	Cloudy intermittent rain	Sunny/cloudy intermittent rain	Sunny
Average Temperature (Celsius)	21.6	25	27.7	29.4
Total Rainfall (mm)	5.5	$0^{(1)}$	$15^{(2)}$	0
Sunlight (hour)	0	0	5.6	11.5
Average Wind (m/s)	1.6	2.1	3.6	2.8
Humidity (%)	99	93	88	76

Source: Japan Meteorological Agency (https://www.data.jma.go.jp/obd/stats/etrn/).
Notes: (1) Tokyo had light rain in the morning, but rainfall less than 1 mm per hour is recorded as 0 mm in Japan; (2) Tokyo had rain until 9:00 AM, but had no rain after that.

Because the purpose of measuring behavior on reference days is to control for regular foot traffic, an ideal reference day should look like an election day except for the fact that no election was held. Elections in Japan are held on Sundays. Therefore, we selected a reference day from among the set of Sundays within a few weeks before or after the election excluding the period of early voting, choosing the date on which weather was most similar. In this potential set of reference days, it turned out that the day one week after the election had the most similar weather, and so was chosen as the reference day. Table 2 presents a summary of weather for the election day, reference day, and two other candidate reference days that were not chosen.

Data Processing

Our primary outcome variables include the GPS-based vote counts and voter turnout rate. Measuring these variables requires several additional coding and cleaning procedures, as the raw Agoop GPS data contain the tracking records of (1) users currently located in the study region, Tokyo's 23 wards, regardless of their residential locations and voting eligibility (e.g., non-Tokyo residents visiting Tokyo), and (2) users residing in the study region regardless of their current locations (e.g., Tokyo residents traveling outside of Tokyo).[10] In addition, as noted earlier, the GPS data set does not provide geographically disaggregated information on the user's residence beyond the city or ward level, which is indispensable to construct the turnout measures.

The coding and cleaning procedures involve the following steps. First, we simply discard any entry that lacks precise geocoordinates by limiting our sample to the records with a GPS accuracy of 100-meters or better.[11] Reflecting the purpose of the analysis, we also drop the entries of users (1) not residing in the 23 wards of Tokyo or (2) with current locations falling outside of the study region. These deletion rules rely on user-specific information provided by the Agoop GPS data.[12] This procedure leaves 197,078 unique users in total, about 100 thousand unique users per day.

Second, to obtain a reliable measure of turnout, it is important to count only the cell-phone users who went to the polling stations designated to them on the basis of their home address. However, users' pinpointed addresses are not disclosed in order to protect their privacy. Therefore, we instead rely on the first signal reception of the day after 6:00 AM as an approximate indicator of users' addresses.[13] The resultant population estimates effectively cover 3,094 (in election day) and 3,093 (in reference day) out of 3,192 neighborhoods in the study region.[14]

Third, to determine whether a user voted during polling hours, we overlay the tracking records between 6:55 AM and 8:05 PM (7:00 AM - 8:00 PM with five-minute margins) within a station-specific radius around the polling stations.[15] As stated previously, we manually

[10] Appendix Table A.1 provides a fictitious sample of track-record data from Agoop to illustrate the structure of the data set.

[11] The 100-meter threshold was determined through parameter tuning, which we will discuss later.

[12] Specifically, we use "accuracy," "citycode," "home_citycode," and "home_countrycode" variables in the original Agoop GPS data.

[13] As with GPS accuracy described earlier, the inclusion rule of 6:00 AM was determined through parameter tuning we discuss later.

[14] The neighborhood boundaries follow the 2015 census, and the total of 3,192 includes neighborhoods without reported residents.

[15] Because GPS signals were delivered only once every five minutes, we set up five-minute margins so that we do not miss the signals from opening/closing time voters.

measured the diameter encompassing each station to identify the center point. We created a station-specific radius by adding a 10-meter margin to one-half of the diameter.[16] A GPS entry is counted as a "vote" if (1) the GPS record falls within the station-specific buffer from a polling station, and (2) the polling station is the designated station for the user.

Validation

If our measurement strategy is ineffective, then any statistical estimates we might obtain from the resulting measure are not credible. Thus, it is important to validate our measure against *known* quantities, such as officially reported turnout at higher levels of aggregation. We estimate two quantities through our approach, population and vote counts, and examine the validity of each against official government statistics in this section.

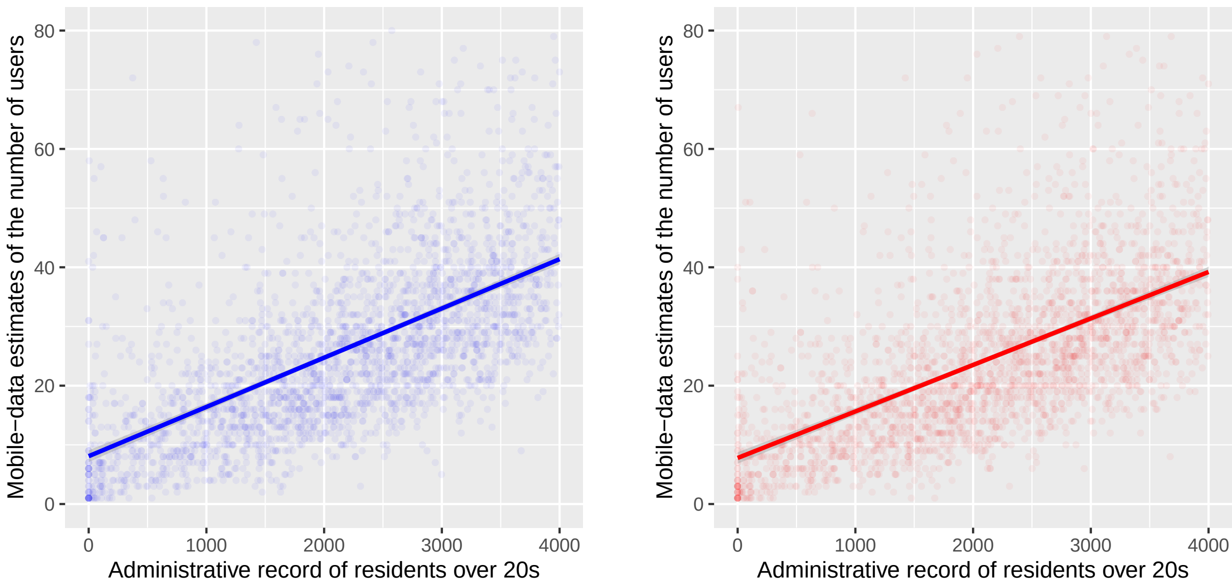
Population Counts

As discussed in the previous section, we regard the first GPS record after 6:00 AM as a user's estimated home address. We then aggregate the numbers of unique users at the neighborhood level and use the neighborhood level unique user count as a GPS-based population count estimate. Official neighborhood-level population statistics are available from *Kokusei Chōsa Shōchiiki Shūkei Kekka* (small-area aggregated census results). This means that we can examine the performance of our strategy for address assignment by calculating the correlation between the census population and the GPS-based population estimate.[17]

Figure 2 shows the result of a performance check using these population counts. The two panels are scatter plots of the GPS-based population estimate in the vertical axes against the census population counts in the horizontal axes on the election day (left panel) and the reference day (right panel). The correlation coefficients are $\rho = 0.69$ and $\rho = 0.67$,

[16]The 10-meter margin was also determined through parameter tuning we discuss later.

[17]We used the population statistics from the 2015 census, and the two-day total for the GPS-based population (election day and reference day).



(a) Election Day ($\rho = .69$)

(b) Reference Day ($\rho = .67$)

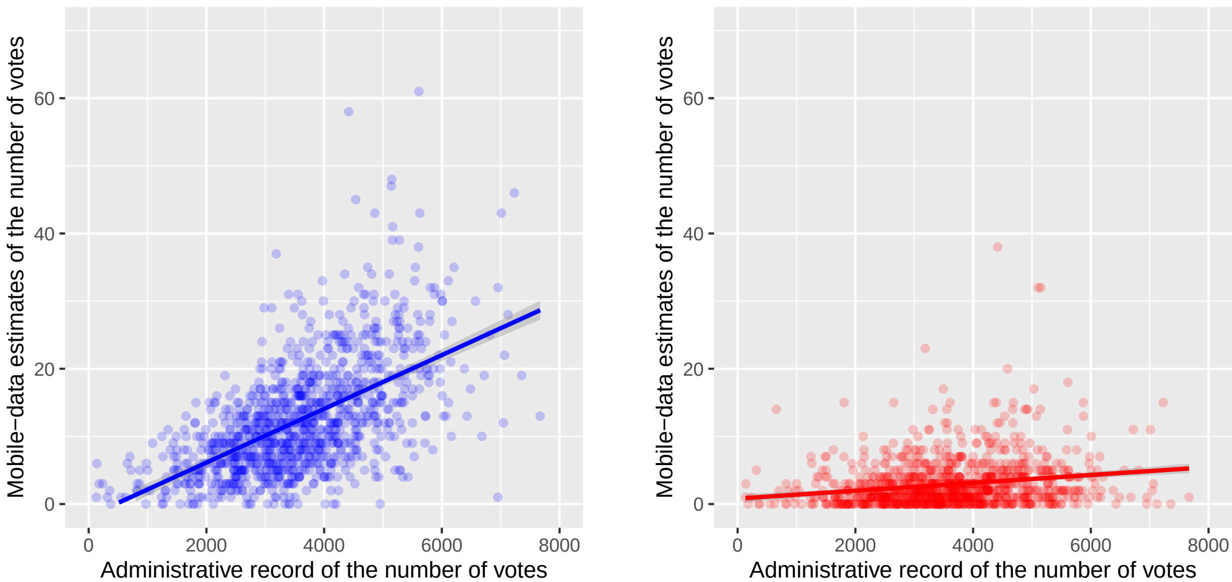
Figure 2: Performance check based on population

which is reasonably high. We do, however, observe dozens of dots that are plotted vertically off the fitted line, indicating overestimation of GPS-based population in these areas. In a later section on sensitivity analyses, we show that our estimation strategy controls for such overestimation.[18]

Vote Counts

In Japan, the smallest unit of observation at which administrative records for voter turnout is available is the polling station. Multiple neighborhoods are usually assigned to each polling station. In addition, neighborhoods are occasionally split into a few smaller blocks when doing so facilitates access to polling stations. As previously mentioned, we obtained the polling station data from Tokyo's 23 wards. One caveat is that although many voters (between 20-30 percent in recent elections) make use of in-person early voting (available since 2003), our data contain only the total number of votes including both early votes and

[18]See Figure A.1 for the census and estimated population counts on the election day projected onto Tokyo's neighborhood polygons.



(a) Election Day ($\rho = .60$) (b) Reference Day ($\rho = .22$)

Figure 3: Performance check based on official voter turnout

election-day votes.[19]

To validate our measurement strategy, we aggregated the GPS-based vote counts to the polling station level using the correspondence table between neighborhoods and polling stations. Figure 3(a) shows a scatter plot of the number of votes estimated from the cell-phone mobility data in the election day (July 21, 2019) against the counterpart for the administrative records. Figure 3(b) depicts the relationship using the reference day (July 28, 2019). The left panel shows an upward pattern with some heteroskedasticity, and the correlation between two data sets is moderate ($\rho = 0.60$), while the counterpart for the reference days in the right panel is small ($\rho = 0.22$).

Several factors might contribute to a reduction in performance of the measure, including: (1) incorrect coordinate inputs; (2) inadequate buffer sizes around polling stations; (3) varying numbers of early voters; (4) demographic/socioeconomic heterogeneities; (5) varying numbers of irrelevant "votes" (or approaches) to polling stations; and (6) general changes in data quality over time. The potential biases due to factors (1) and (2) need to be minimized

[19] About 28.6% of votes within Tokyo's 23 wards were cast through early voting in the House of Councillors election on July 21, 2019 (SEACa, 2019; SEACb, 2019)

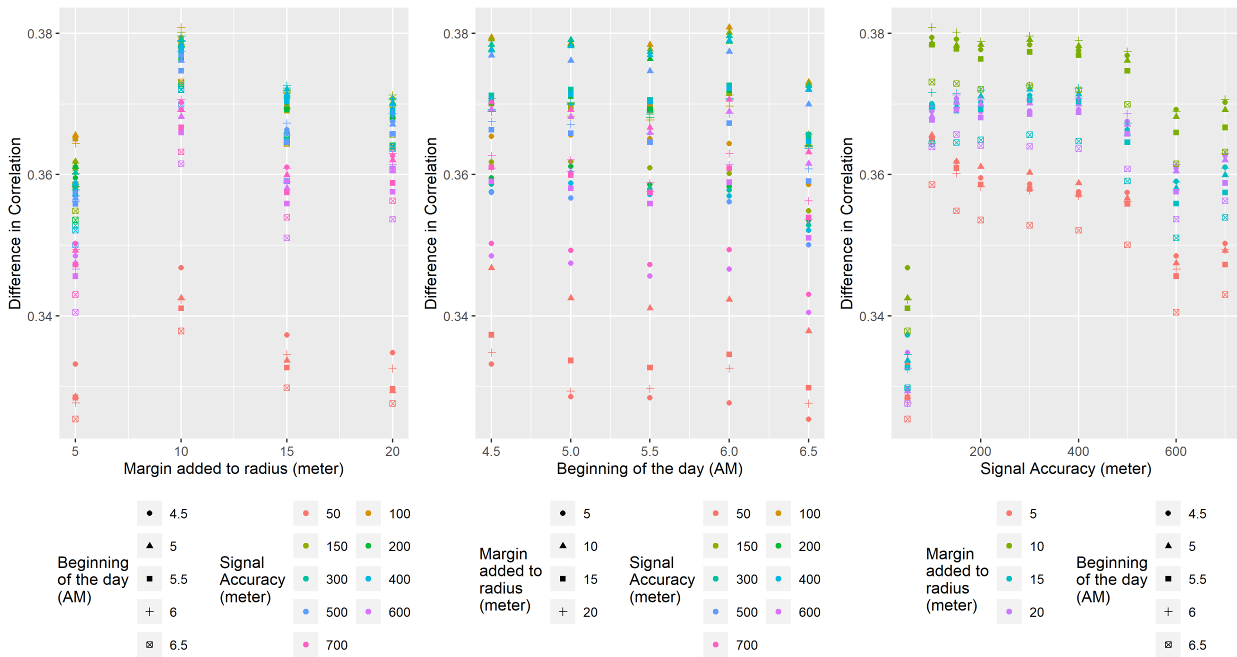
through tuning, and any remaining impacts need to be assessed through sensitivity analyses, which we report later. Factor (3) is controlled for by measuring the distance to the closest early-voting polling station for each user. Factors (4) and (5) are taken into account through our estimation strategy. For now, we make an assumption that the number of early votes are proportional to the number of election-day votes across polling stations. Finally, we initially collected the data for three elections: two in 2017 and one in 2019. We decided to use the latest election data on July 21, 2019 since the correlation was highest with this data—indeed, Agoop introduced new software to collect location "waypoints" in 2019.[20]

Tuning Parameters

In the previous sections, we used the pre-determined set of three parameters to process the data: GPS accuracy, the margin added to a station-specific radius, and the beginning of the day. The first and second variables adjust the close cases for false positives (abstainers who were judged as casting a vote) and false negatives (actual voters who were judged as abstainers). The third variable affects the estimated locations of residences and corresponding designated polling stations. Following Chen et al. (2019), we adopted an agnostic approach in determining these parameters. That is, we prepared several options for each of these parameters, and then selected the set of parameters that showed the best performance for an evaluation criterion.

The panels in Figure 4 show how performance changes depending on the values of tuning parameters. Our evaluation criterion, presented in the vertical axis of each panel, is the election-day to reference-day difference in the correlation coefficient between the administrative turnout record and estimated voter turnout aggregated at the level of a voting precinct. Take Figure 3 for example, this quantity is calculated as $.60 - .22 = .38$. It turns out that among the combinations we prepared, the difference in the correlation coefficients is largest

[20]Waypoints are intermediate points on a given route between point A and point B, such as where a user changes direction in course.



(a) Margin to Radius = 10m (b) Beginning of Day = 6AM (c) Signal Accuracy = 100m

Figure 4: Results of parameter tuning

Note: Parameter tuning based on the margin added to building radius (left panel), the timing of the first record to define users' estimated address (middle panel), and signal accuracy in meters (right panel). The y-axis represents the difference in correlation between the election day data and reference day data. The sub-captions indicate the tuned values for each parameter.

when the margin is 10 meters, GPS accuracy is 100 meters, and the beginning of the day is set as 6:00 AM, and none of these was the minimum or maximum value for tuning. Among the possible combinations of these parameters, the difference in the performance between the best and worst results, or the improvement from the worst-case scenario due to tuning, was 0.054.

Application 1: Cost of Voting and Turnout

We first illustrate the substantive utility of our approach with an application based on the rational voter model (see Aldrich, 1993; Downs, 1957; Riker and Ordeshook, 1968). According to this model, an individual voter will decide to turn out on election day if the cost of voting (C) is outweighed by the expected benefit the voter will get from electing their preferred candidate (B) times the probability that their vote might be decisive (P), plus the intrinsic value (D) the voter might get from the civic act of participation: $C < PB + D$.[21]

In existing empirical tests of the rational voter model, the cost of voting has often been proxied by the distance to the voter's designated polling station (e.g., Bhatti, 2012; Cantoni, 2020; Garnett and Grogan, 2021; Gibson et al., 2013; Haspel and Knotts, 2005; Nishizawa, 1991). For this application, we focus only on the distance (in walking time) between voters' homes and their designated polling stations, as well as early voting stations. We do not attempt to measure the other components of the model (P, B, D). The basic expectation is simply that a greater distance to the polling station should reduce the likelihood of turning out to vote on election day (i.e., a cell-phone mobility record at the polling station during voting hours).

Existing empirical tests of the rational voter model face some methodological limitations based on how turnout and the cost of voting are measured. For example, if turnout data are aggregated to some unit (such as a ward or voting precinct), then the cost of voting must be measured as an average distance to a polling station across all households within the unit, masking important individual-level variation. In contrast, direct questions about individual-level turnout in surveys are subject to social desirability bias and over-reporting. The GPS-based measure we introduce provides an alternative approach that can overcome these challenges.

Key Explanatory Variables

Our key explanatory variable is the cost of voting at a designated polling station, measured in terms of walking time needed to travel the distance. We use walking time because a simple Euclidean distance between a home neighborhood and a designated polling station does not take into account other factors that affect the cost of voting, such as road alignments,

[21] There is an extensive literature, beyond the scope of this application, that further attempts to reconcile the paradox of voting in large electorates based on elite mobilization, social pressure, and group affiliations (e.g., Bond et al., 2012; Cox et al., 1998; Morton, 1991; Schachar and Nalebuff, 1999; Uhlaner, 1989).

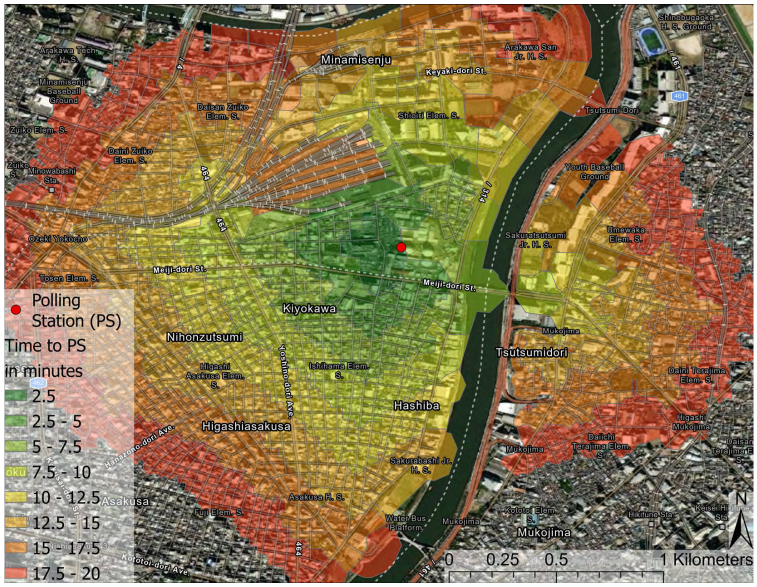


Figure 5: An example of the cost distance calculation
Note: Regions for which cost distance is over 20 minutes are not shown.

hills, and traffic conditions.[22] More specifically, we calculated the "cost distance" to each polling station based on the shortest weighted distance to each polling station (Environmental Systems Research Institute, 2021). Since this cost distance calculation involves multiple parameters that may be arbitrary, we rely on the service area analysis program of ArcGIS Network Analyst extension, a standard tool for cost distance analysis.[23]

Figure 5 shows the results of the cost distance calculation for a sample polling station area in Tokyo's Arakawa Ward, using a range of 0 to 20 minutes. As most voters in Tokyo's 23 wards go to polling stations on foot (and to a somewhat lesser extent by bicycle), we calculated the cost distance in walking time ranging from 0 to 30 minutes with cutoff values in increments of 2.5 minutes. The figure illustrates that the cost distances measured in walking time are different from a simple Euclidean distance.

Given a large proportion of early voters in the 2019 House of Councillors election, we also

[22]Note that walking time does not include indirect costs, such as the opportunity cost from missing work.

[23]The calculation was performed with the following options. Mode: Walking Time; Direction: Away from Facilities; Cutoffs: from 2.5 to 30 minutes in increments of 2.5; Date & Time: July 21, 2019, noon; Polygon Detail Level: High Precision; Boundary Type: Overlap; and Polygon Rings or Disks: Rings. A 0.5 unit of ArcGIS credit is required per cutoff value for each polling station.

collected data on early voting to control for its impact. Specifically, we geo-coded the list of all early-voting polling stations in Tokyo's 23 wards, provided by the Tokyo Metropolitan Government (SEACc, 2019). We then calculated the cost distances for the 205 locations in the same way as for the election-day polling stations.[24]

In the validation section, we estimated each user's home neighborhood at a 100-meter-grid scale. This allows us to measure the estimated walking time from the grid to a designated polling station for each voter. After filtering out irrelevant polling stations, 71 percent of grids have a unique cost distance value. For the remaining grids with multiple cost distance values, we employ sampling strategies explained in the following subsection. We drop the users whose estimated walking time to an election-day polling station is over 30 minutes,[25] and the remaining sample contains 92,400 unique daily users in the election day and 87,848 unique daily users in the reference day.

Estimation Model

We construct a two-day pooled cross-sectional data set consisting of the cellphone users on July 21, 2019 (the election day) and the reference day (one week later). We use a difference-in-differences (DID) estimation strategy to test whether a user's likelihood of going to a polling station on the election day decreases significantly compared to the counterpart estimate for the reference day as the estimated walking time to polling station increases (see, e.g., Nunn

[24]The same parameters were used except Date & Time, which was set at July 17, 2019, noon. Mail-in absentee voting is only permitted for a small number of voters with physical handicaps; all other early voters must vote in person. See Kitamura and Matsubayashi (2022) for an analysis of how weather conditions (precipitation) influence early voting decisions.

[25]In densely-populated Tokyo's 23 wards, it is unlikely for the estimated time to take more than 30 minutes, so such values are probably due to some errors in home neighborhood estimation or cost distance calculation.

and Qian, 2011). We estimate a linear probability model with the following form:

$$\begin{aligned} Vote_{igst} &= \sum_{k=1}^{11} \alpha_k I_{gs}^{min.{}_{2.5k}^{2.5(k+1)}} + \sum_{k=1}^{11} \beta_k \left(I_{gs}^{min.{}_{2.5k}^{2.5(k+1)}} \times I_t^{Elec} \right) \\ &+ \sum_{j=1}^{12} \gamma_j I_{gs}^{EV_{2.5j}^{2.5(j+1)}} + \sum_{j=1}^{12} \delta_j \left(I_{gs}^{EV_{2.5j}^{2.5(j+1)}} \times I_t^{Elec} \right) + \sum_s \lambda_s I_s + \rho I_t^{Elec} + \epsilon_{igst}, \end{aligned} \tag{1}$$

where $Vote_{igst}$ is whether a cellphone user i who lives in a grid g went to vote at their designated polling station s in day t. $I_{gs}^{min.{}_{2.5k}^{2.5(k+1)}}$ is a dummy variable that takes the value of 1 if an estimated walking time from a grid g to a polling station s is from $2.5k$ to $2.5(k+1)$ minutes. Similarly, $I_{gs}^{EV_{2.5j}^{2.5(j+1)}}$ is a dummy variable that takes the value of 1 if a walking time to the closest polling station for *early voting* is from $2.5k$ to $2.5(k+1)$ minutes. I_t^{Elec} is a dummy variable that takes the value of 1 on the election day, $\sum_s \lambda_s I_s$ is polling station fixed effects, and the final term in the equation (ϵ_{igst}) is a disturbance term.

The quantities of interest are $\sum_{k=1}^{11} \beta_k$ that represent the changes in the probability of voting for the voters whose walking time to their polling stations are from $2.5k$ to $2.5(k+1)$ minutes compared with the counterfactual situation of a walking time from 0 to 2.5 minutes (the baseline). These coefficients are expected to be negative, and their sizes are expected to become larger in absolute terms as the walking time to the polling station increases ($H_1 : \beta_{k+1} < \beta_k < 0$). We also consider the impact of the proximity to early-voting polling stations, represented by $\sum_{j=1}^{12} \delta_j$.[26] People are more likely to take the opportunity to vote early if the polling stations for early voting are nearby. Therefore, we expect that the proximity to polling stations for early voting will decrease the likelihood of voting on election day, so these coefficients are expected to be positive and become larger as walking time increases ($H_2 : \delta_{j+1} > \delta_j > 0$).

[26] For early-voting dummies, j takes the value up to $j = 12$ representing "More than 30 minutes" not "30 to 32.5 minutes." Since about one percent of the users fall in this category, we keep them in the sample.

Results

Figure 6 shows the estimated effects of walking time to polling station on voter turnout on election day. The point estimates and confidence intervals in red are for models based on distance to the polling station, whereas those in blue are for distance to an early-voting station. Point estimates and CIs are obtained from the sample in which the polling stations judged as having been voted at (based on our algorithm) are first selected as the designated polling stations, and the other polling stations were randomly selected when multiple designated polling stations exist for a would-be voter (cell-phone user). The red and blue rounded rectangles surrounding the main point estimates and CIs are collections of alternative point estimates obtained from alternative samples where observations (users) were randomly selected when multiple designated polling stations exist.

The results paint a picture that is consistent with the expectations of the rational voter model. As cost (walking distance to the polling station) increases, the likelihood of voting (visiting the polling station) on election day decreases. In addition, the probability of voting on election day is higher for voters whose designated early-voting polling stations are less convenient (further from home).[27] Compared to the baseline of 0 to 2.5 minutes of walking time, all of the coefficients for larger cost distances are negative and statistically significant. The substantive size of the effects is also large: compared to a voter with almost no travel distance (the baseline), the probability of turning out to vote on election day for a voter whose polling station is more than 10 minutes away decreases by more than 10 percentage points.

An additional point about the results in Figure 6 that merits attention is that, although the coefficients change with distance to the polling station, the rate of change slows as stations get further away. This suggests that there exists some threshold of exclusion for the relationship between convenience and voting that affects some voters (e.g., casual voters),

[27] One limitation is that we lack the ability to directly investigate the usage of early voting, since we do not have cell-phone mobility data for early-voting days in our panel. The coefficient of δ_{12}, the category omitted from the graph, is .119 (*s.e.* =.015).

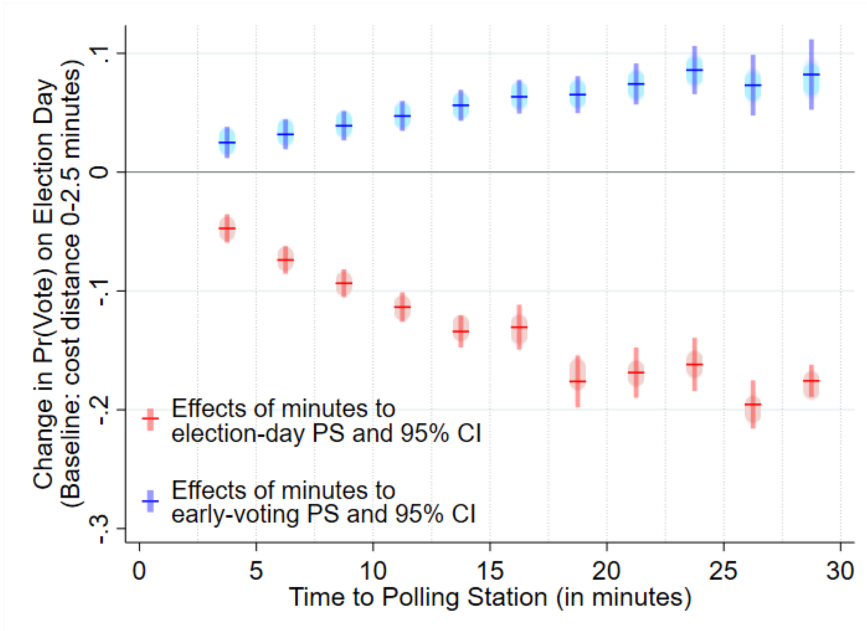


Figure 6: Effects of walking time to polling station on voter turnout
Note: Linear probability model with robust standard errors for 95% confidence intervals. The coefficients $\sum_{k=1}^{11} \beta_k$ and $\sum_{j=1}^{12} \delta_j$ in Equation 1 are on the vertical axis. See Table A.2 for the results in table format. The point estimates and CIs in red were obtained from the sample in which the polling stations judged as being voted in our algorithm are first selected as designated polling stations, and remaining polling stations were randomly selected when multiple designated polling stations exist for a voter. Blue rounded rectangles are the 1,000 point estimates that were obtained from alternative samples where observations, or voters, were randomly selected when multiple designated polling stations exist for a voter.

but not others (e.g., committed voters).

Sensitivity Analyses

Although the mobility data we use provide a powerful tool for studying questions like turnout, these data can potentially be fraught with measurement error. For example, measurement error in Y (turnout) can widen confidence intervals, and can also cause attenuation bias, which can also occur when measurement error exists in X (distance to a polling station).

In this subsection, we perform sensitivity analyses to gauge the impact of two major sources of measurement error. The first type of error originates from the fact that we can only observe users' noisy GPS signals to determine their locations. The tuning of parameters,

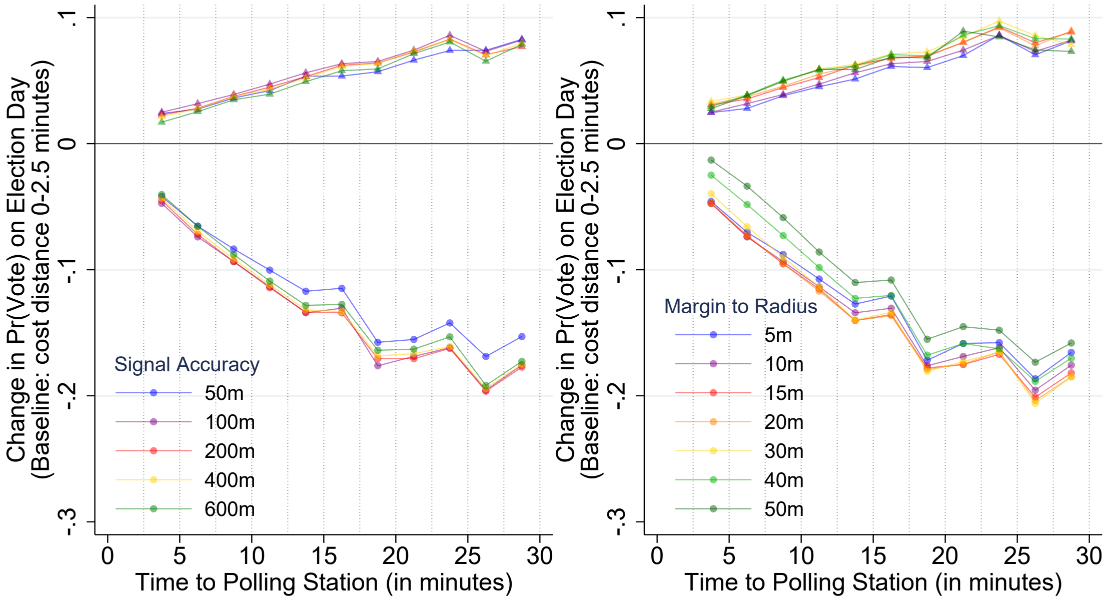


Figure 7: Sensitivity analyses based on signal accuracy and polling station radius
Note: The left panel replicates the results from Figure 6 with varying signal accuracy for the sample of included users; the right panel shows corresponding results when varying the margin of the radius around the polling station for coding a user as having voted (entered the polling station).

which we described earlier, helps to minimize noise by selecting users whose signals are more accurate and by changing the size of the radius surrounding each polling station used to determine whether a user voted (entered the polling station). To evaluate whether this effort sufficiently reduces noise, and whether a failure to do so poses a risk to our ability to make correct inferences, we perform the following sensitivity analysis.

The line plots in Figure 7 show the point estimates of the effect of walking time to polling station on election-day turnout estimated with the sample with varying signal accuracy (in the left panel) and margin to a polling-station specific radius (in the right panel). The signal accuracy and margin range from 50 meters to 600 meters and from 5 meters to 50 meters, respectively. In the left panel, the estimates for the sample of users with 50m signal accuracy (i.e., a worse signal accuracy) is closer to zero (reflecting attenuation bias). However, the basic pattern from the main results is apparent regardless of signal accuracy. Similarly, the larger the radius used for each polling station, the closer the coefficients are to zero, while

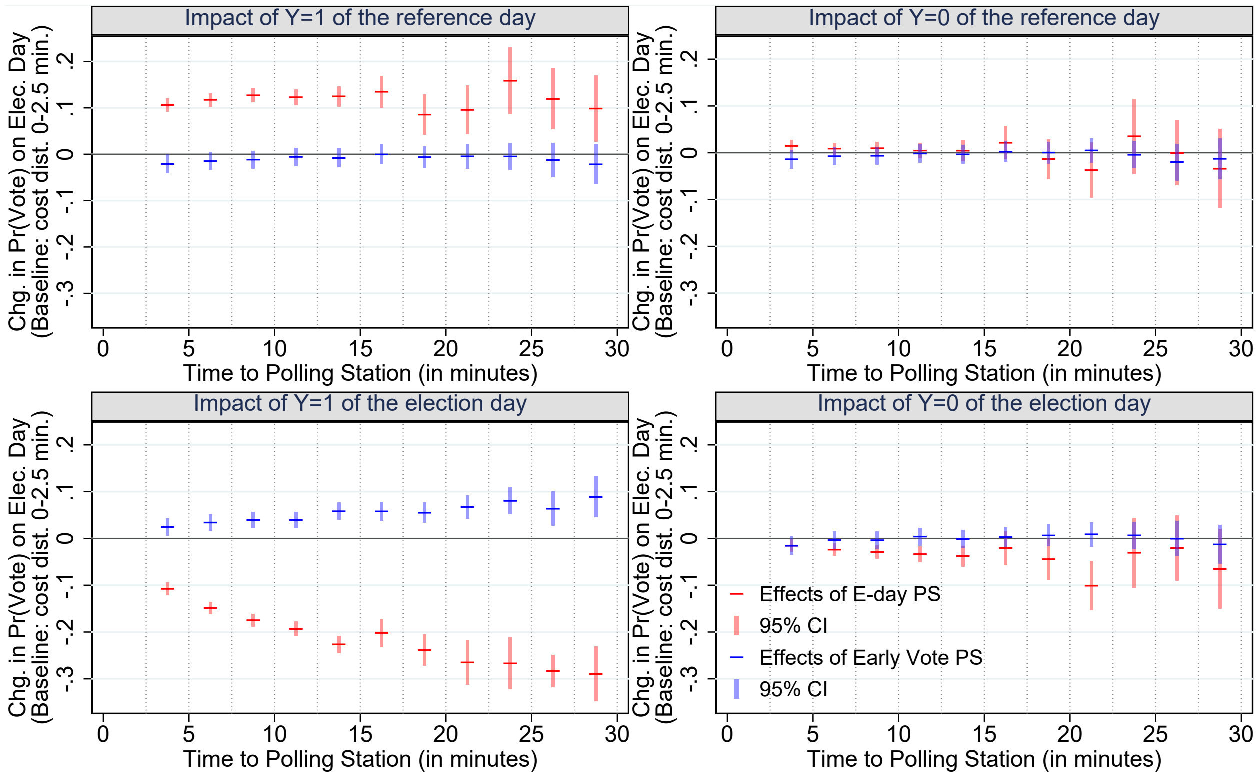


Figure 8: Sensitivity analysis based on decomposition of DID estimation
Note: The total effect ≈ lower left panel + lower right panel − upper left panel − upper left panel.

the basic pattern holds.

The second source of measurement error is the misclassification of foot traffic near polling stations as "voting," since many polling stations are located in places where people might gather anyway. This kind of misclassification can take the form of a false negative (coding non-voting users as having voted) or a false positive (coding voting users as having not-voted). While the extent of this type of measurement error is difficult to assess, the DID estimation using a reference day provides a way to parse this potential source of error. Specifically, we can decompose the observed DID coefficients into four parts and visualize what kinds of systematic errors occur, in which data-generating processes, to identify the type of assumptions necessary to accept our results.

Figure 8 presents this decomposition of the DID effects. Overall, users in the data can be grouped into four types: those coded as having voted on the election day, those coded as not having voted on the election day, and the two counterparts on the reference day. For each component, we keep the selected component unchanged while eliminating all systematic variation of the other three components by randomly assigning the outcome variable according to the flow chart shown in Figure A.2. When only the information of $Y = 1$ of election day is retained (lower left panel of Figure 8), we get a picture similar to Figure 6. This means that the majority of the estimated effects in the main results can be attributed to the users who were coded as having voted on election day.

Other panels in Figure 8 show evidence of potential confounding. For example, when only the information of $Y = 1$ on the reference day is retained, the coefficients are positive. Interpreting these results requires some caution. Because the impact of $Y = 1$ on election day is 0 in this case (due to randomization) and in the DID estimation the effect estimates are obtained by $Y_{elec} - Y_{ref}$, users who live further than 2.5 minutes away in terms of cost distance are about 10% less likely to be (mis)coded as having voted compared with users living within 0-2.5 minutes distance *on the reference day.* Equivalently, users living within 0-2.5 minutes distance on the reference day are about 10% more likely to be miscoded as having voted because of the proximity of their residence to polling stations.

We do not know whether users who were coded as having voted ($Y = 1$) on election day suffer from the same kind of bias. However, if we can assume they did, then the DID estimation rightly remove this bias and identifies the effects. If a researcher still worries about this kind of bias, a solution is to drop the 0-2.5 minutes (or closest) category of user, and instead apply the second-closest category as the baseline. In the lower right panel, $Y = 0$ on the election day sample shows limited but consistent effects because the cost distance does not affect most of the users with $Y = 0$ while there are somewhat more $Y = 0$ users for locations further from the polling station.

Overall, these sensitivity analyses are useful for adjudicating the validity of the GPS-based measure, ensuring that it is appropriately tuned to address the research question at hand, and identifying the sources of potential systematic bias. In addition, the sensitivity analyses help to clarify the necessary assumptions for the DID estimation design to identify the estimands of interest. The analyses we have performed here may need to be adapted or supplemented in other applications, as needed.

Application 2: Effects of Community-Level Destruction

Our second application to demonstrate the substantive utility of our approach uses data on neighborhood-level damages caused by the firebombing of Tokyo during World War II, from a recent study by Harada et al. (2021). A growing literature at the intersection of research on political violence and historical political economy considers how exposure to wartime violence and other forms of destruction or trauma influences social behavior, including participation in elections (for a review, see Walden and Zhukov, 2020). An influential argument in existing studies is that exposure to war violence might increase prosocial behavior, including participation in elections, due to the activation of collective action (e.g., Bellows and Miguel, 2009; Blattman, 2009; Gilligan et al., 2014; Kage, 2021; Lupu and Peisakhin, 2017).

However, empirical evidence is limited to a few cases, and large-scale destruction with fewer survivors might have long-term detrimental effects on turnout—due to, for example, the impact of wholesale destruction on neighborhood-level social capital and socioeconomic well-being. Indeed, Harada et al. (2021) find that the most heavily damaged neighborhoods have less-organized neighborhood associations (in indicator of geographically localized social capital), and that they exhibit lower socioeconomic well-being in terms of education, occupation, and residential stability. Each of these social conditions might lower the propensity to local residents to vote (e.g., Bond et al., 2012; Brady et al., 1995; Cox et al., 1998; Dowding et al., 2012; Wolfinger and Rosenstone, 1980).

Although the Harada et al. (2021) data set on damages from the Tokyo firebombing represents one of the most disaggregated data sets of war violence and large-scale destruction (cf. Kocher et al., 2011), the available outcome variables at the same unit of analysis (neighborhood) are limited to socioeconomic indicators. Combining the Tokyo firebombing data and our turnout estimates allows us to offer additional quantitative evidence that speaks to the ongoing debate in the literature regarding the long-term effects of exposure to violence and community-level voter behavior.

Additional Data Processing

The firebombing damages are measured at the neighborhood level based on georeferenced historical aerial photographs and remote-sensing techniques.[28] Our key variable of interest is the fraction of destroyed residential area relative to the overall residential area within each neighborhood (henceforth "damage ratio"), which is conceptually defined as follows:

$$Damage = \frac{\text{Destroyed residential area}}{\text{Overall residential area}} \tag{2}$$

Each ratio variable was assigned a value ranging from 0% to 100% in increments of 10%, for an eleven-unit scale. We treat the damage ratio as a continuous variable for our analysis. Figure 9 presents the damage ratios across neighborhoods of Tokyo's 23 wards, as reported by Harada et al. (2021).

Estimation Model and Variables

We construct the two-day pooled cross-sectional data set similar to the previous section and adopt a linear probability model with a difference-in-differences specification with a continuous treatment variable to test whether the amount of firebombing damage affects estimated individual turnout in contemporary Tokyo. Neighborhood-level population is used as a weight in the regressions. In addition, as shown in the previous section, there

[28] Specifically, Harada et al. (2021) measure damages at the level of the *chō-chōmoku*, which is the lowest-level administrative unit in Japan for measuring census-based socioeconomic variables, and roughly corresponds to the popular notion of a neighborhood. This is considerably smaller than the units used in other studies of long-term bombing effects (e.g., Brakman et al., 2004; Davis and Weinstein, 2002; Kocher et al., 2011; Lin, 2022).

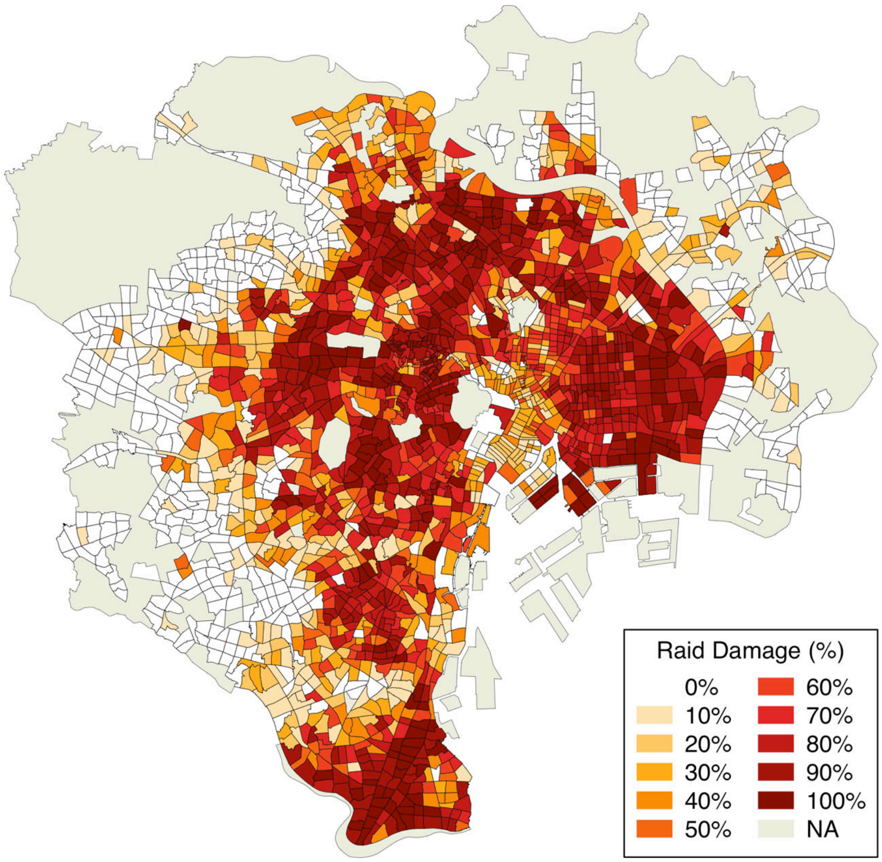


Figure 9: Distribution of the air raid damages across Tokyo's 23 wards
Note: Shading of the neighborhoods indicates the damage ratio in percentage scale. The neighborhoods excluded from our sample are left blank. Reproduced from Harada et al. (2021).

is a concentration of misclassified voters in areas that are close to the election-day polling stations—measurement error in this case would cause larger standard errors and attenuation bias. Therefore, we exclude the voters whose estimated address is located from 0 to 2.5 minutes in cost distance to their designated polling station.[29] Our estimation model is formally written as follows:

[29]We show the results including these voters in robustness checks.

$$
\begin{aligned}
Vote_{ignswt} &= \tau\left(Damage_n \times I_t^{Elec}\right) + \sum_{n=1}^{2{,}139} \lambda_n I_n + \rho I_t^{Elec} \qquad (3)\\
&+ \sum_{k=1}^{10} \alpha_k I_{gs}^{min._{2.5k}^{2.5(k+1)}} + \sum_{k=1}^{11} \beta_k \left(I_{gs}^{min._{2.5k}^{2.5(k+1)}} \times I_t^{Elec}\right)\\
&+ \sum_{j=1}^{12} \gamma_j I_{gs}^{EV_{2.5j}^{2.5(j+1)}} + \sum_{j=1}^{12} \delta_j \left(I_{gs}^{EV_{2.5j}^{2.5(j+1)}} \times I_t^{Elec}\right)\\
&+ \sum_{h=1}^{20} \theta_h \left(Lon._n^p \times Lat._n^q \times I_t^{Elec} : p+q \leq 5, 0 \leq \{p,q\}, \{p,q\} \in \mathbf{Z}\right)\\
&+ \sum_{m=1}^{5} \phi_m \left(IP_n^m \times I_t^{Elec}\right) + \sum_{m=1}^{5} \psi_m \left(AP_n^m \times I_t^{Elec}\right)\\
&+ \pi\left(\ln(Pop.Density1939)_n \times I_t^{Elec}\right) + \zeta\left(\%Residence_n \times I_t^{Elec}\right) + \epsilon_{ignswt},
\end{aligned}
$$

where $Vote_{ignswt}$ is the binary estimate of whether a cell-phone user i voted in grid g in neighborhood n at polling station s in prewar ward w in day t. These subscripts reflect the multi-level nature of our data. Our key treatment variable, $Damage_n \times I_t^{Elec}$, is measured as the interaction between $Damage_n$ and I_t^{Elec}, where I_t^{Elec} is a dummy variable that takes the value of 1 on the election day, and $Damage_n$ is the fraction of destroyed residential area relative to the overall residential area within each neighborhood. The effects of the bombing are measured by τ.

All other terms represent control variables and an error term.[30] $\sum_{n=1}^{2,139} \lambda_n I_n$ is neighborhood fixed effects, $\sum_{h=1}^{20} \theta_h(Lon._n^p \times Lat._n^q \times I_t^{Elec} : p+q \leq 5, 0 \leq \{p,q\}, \{p,q\} \in \mathbf{Z})$ is the set of polynomials between standardized longitude and latitude of neighborhood centroids (up to the fifth-order polynomial), interacted with the election day dummy. $\sum_{m=1}^{5} \phi_m(IP_n^m \times I_t^{Elec})$ and $\sum_{m=1}^{5} \psi_m(AP_n^m \times I_t^{Elec})$ is the set of polynomials of the standardized distances from the Imperial Palace and the nearest aiming point (the key treatment assignment mechanisms which must be controlled for) up to the fifth order polynomial, interacted with the election

[30]Non-interacted cost distance terms are not subsumed by neighborhood fixed effects because some neighborhoods are divided into a few smaller areas, each of which is assigned to a different designated polling stations.

day dummy, respectively. Finally, $\ln(Pop.Density1939)_n$ is the logged population density as of 1939 (in population per km^2), and $\%Residence_n$ is a human-coded prewar residential ratio. A description of other terms is provided in the previous section.[31]

Results

Table 3 reports the main regression estimates for the effect of the damage on our estimated voter turnout. Model 1 only includes the variables in the first line of Equation 2, while Model 2 includes all the control variables listed.

The results suggest a persistent and negative association between the extent of aerial bombing damages and voting behavior in the present day. Regardless of inclusion of control variables, $Damage \times Election\ Day$ is negative and statistically significant (Models 1–2). The negative coefficient estimates for the interaction of $Damage \times Election\ Day$ indicate a decrease in the probability of voters going to their polling stations on election days (July 21, 2019) in the neighborhoods with heavier bombing damages, relative to the baseline of non-election days (July 28, 2019).[32]

Putting the size of the estimated coefficient into perspective, a voter living in a neighborhood that had no damage would go to vote with a 2-percentage-point lower probability if this neighborhood had been completely destroyed in the firebombing based on Model 3 of Table 3. In all, the systematic negative association in Models 1–2 suggests a form of persistent legacies of the World War II air raids on voting behavior in Tokyo today, complementing the findings of Harada et al. (2021) with regard to social capital and socioeconomic outcomes.

[31] Since the treatment and control variables (except polling station fixed effects) are interaction terms, we can express this model using the Interaction Directed Acyclic Graph (IDAG; Nilsson et al., 2020) and be explicit about the necessary assumptions. See Appendix B for details.

[32] As Nunn and Qian (2011, 619) note, choosing a different reference period would result in changes in the point estimates and standard errors.

Table 3: Difference-in-differences regression of estimated election-day turnout on the level of damages from the Tokyo firebombing

Outcome Variable: Estimated Vote	(1)	(2)
Damage × Election Day	-0.0138^{**}	-0.0206^{**}
	(0.0054)	(0.0081)
Neighborhood FE (2,139 categories)	✓	✓
Day FE	✓	✓
Covariates		
Cost distance to election-day polling station		✓
Cost distance to early-voting polling station		✓
Covariates × Election Day		
Cost distance to election-day polling station		✓
Cost distance to early-voting polling station		✓
Fifth-ordered polynomials(longitude, latitude)		✓
Fifth-ordered distance to the Imperial Palace		✓
Fifth-ordered distance to the closest aiming point		✓
Logged neighborhood population density 1939		✓
Prewar Residenatial Ratio		✓
Observations	120,495	120,495
Within Polling Station R^2	0.0417	0.0573

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Linear probability model with population weights. Standard errors clustered by neighborhoods in parentheses. All models exclude the users whose cost distance is from 0 to 2.5 minutes. Difference-in-differences estimation is performed between the election day (July 21, 2019) and the reference day (July 28, 2019.) The set of *Covariates* includes: 10 dummy variables of cost distance to the closest early-voting polling stations; 12 dummy variables of cost distance to the closest election-day polling stations. The set of *Covariates* × *Election Day* includes: 11 dummy variables of cost distance to the closest early-voting polling stations; 12 dummy variables of cost distance to the closest election-day polling stations; 20 fifth-order polynomials of longitude and latitude including interactions; fifth-order polynomials of standardized distance to the Imperial Palace; fifth-order polynomials of standardized distance to the closest aiming points of the US bombing campaign in WWII; logged neighborhood-level population density in 1939 (population counts per km^2); ratio of prewar residential areas, each interacted with the election day dummy. See Table A.3 for the full results.

Robustness Checks

To check for the validity of the measure, we replicated the analysis in Table 3 using the estimated turnout and official turnout aggregated to the level of polling station as outcome variables. Estimated turnout, aggregated to the polling station level, was calculated using the estimated residences counted at each polling station as the denominator and the total number of those who approached the designated polling station as the numerator. The official turnout was calculated as the ratio of the number of voters to the number of voters on the election day and was set to 0 on the reference day. When aggregating the treatment and control variables to the polling station level, we used the arithmetic mean of the values of the variables assigned to each cell-phone user based on the estimated address.

Table 4 shows the results of estimating Equation 3 using these aggregated data. The two cost-distance variables are entered as continuous variables in this model. Models 1–2 use estimated turnout, while Models 3–4 use official turnout, with and without control variables, respectively. Table 4 highlights two important points. First, the coefficients obtained using the estimated and official turnout rates are similar. Indeed, the null hypothesis that the difference between the coefficients of Models 1 and 3 (or 2 and 4) is zero was not statistically significant ($p = 0.67$ and $p = 0.66$, respectively). This shows that our estimated turnout meets the necessary conditions to be used as a proxy for official turnout. Second, we see that the coefficients from the analyses using official voting rates are statistically significant and negative, and that their sizes do not differ significantly from those estimated from the individual cell-phone user data in Table 3. This result supports the idea that estimation using estimated turnout can lead to substantively identical conclusions as estimation using official turnout.

In Table 5, we summarize the results of several robustness checks. First, we re-estimate Model (2) in Table 3 including the voters whose estimated address is located from 0 to 2.5 minutes from the polling station in cost distance. Second, the model is re-estimated without using population weights. Third, we clustered the standard errors by Tokyo's prewar 35 wards instead of neighborhoods to provide a conservative estimate of standard errors against potential spatial correlations. Finally, we used the cost distance to the election-day polling stations as a *negative control* outcome, a class of falsification test (Arnold and Ercumen, 2016). The cost distance should not be correlated with raid damage but largely share the same construction process as the vote count. Therefore, a weak or null association between raid damages and our estimated cost distance provides some assurance that our estimates

Table 4: Difference-in-differences regression of estimated turnout (column 1–2) and official turnout (column 3–4) on the level of damages from the Tokyo firebombing using the data aggregated at the level of a voting precinct

Outcome Variable:	Estimated Turnout		Official Turnout	
	(1)	(2)	(3)	(4)
Damage × Election Day	−0.0101	−0.0151	-0.0066^{*}	-0.0202^{***}
	(0.0072)	(0.0112)	(0.0040)	(0.0050)
Neighborhood FE	✓	✓	✓	✓
Day FE	✓	✓	✓	✓
Covariates				
Cost-dist.to election-day polling st.		✓		✓
Cost-dist.to early-voting polling st.		✓		✓
Covariates × Election Day				
Cost-dist.to election-day polling st.		✓		✓
Cost-dist.to early-voting polling st.		✓		✓
5th-ordered polynomials(lon., lat.)		✓		✓
5th-ordered dist.to the Palace		✓		✓
5th-ordered dist.to aiming points		✓		✓
Logged pop.density 1939		✓		✓
Prewar Residenatial Ratio		✓		✓
Observations	1,878	1,878	1,878	1,878
Within Polling Station R^2	0.7208	0.7774	0.9944	0.9969

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. The number of voters is used as weights. Control variables are weighted averages of the values from the cell phone user's estimated address for each polling station. See Table 3 for the selection of days used in the analysis and the description of the covariates except cost distance variables and their interactions, which are entered into these models as continuous variables. See Tables A.4 and A.5 for the full results.

are unlikely to be attributable to systematic biases in our mobility data processing.

The results of Models 1–3 in Table 5 show that the estimated coefficients are negative and statistically significant, consistent with their counterparts in Table 3. A slightly smaller coefficient in model (1) indicates that the bias due to misclassification attenuates the negative impact of the bombing. The only slight increase in the standard errors in model (3) may indicate that clustering at the neighborhood level already takes into account much of any potential serial correlation within geographical units.

Importantly, $Damage_n \times I_t^{Elec}$ is not systematically associated with the cost distance estimates in Model (4). The lack of systematic association provides further confidence that the estimates in Models 1–3 capture the associations specific to the mobility patterns around polling stations on election days, rather than omitted confounding structures inflating the

Table 5: Summary of the robustness checks: estimation with observations very close to polling stations (column 1), without population weights (column 2), with SEs clustered by prewar 35 wards (column 3), and with cost distance to election-day polling station as a negative control outcome (column 4)

	Type of Robustness Checks:			
Outcome Variable: Estimated Vote (except Model(4))	Including observations very close to PS (1)	Estimated w/o population weights (2)	Standard errors clustered by prewar 35 wards (3)	Using cost-distance as an outcome (4)
Damage × Election Day	−0.0168** (0.0082)	−0.0137* (0.0075)	−0.0206** (0.0094)	−0.0330 (0.0541)
Neighborhood FE	✓	✓	✓	✓
Day FE	✓	✓	✓	✓
Covariates				
Cost-dist.to election-day polling st.	✓	✓	✓	
Cost-dist.to early-voting polling st.	✓	✓	✓	✓
Covariates × Election Day				
Cost-dist.to election-day polling st.	✓	✓	✓	
Cost-dist.to early-voting polling st.	✓	✓	✓	✓
5th-ordered polynomials(lon., lat.)	✓	✓	✓	✓
5th-ordered dist.to the Palace	✓	✓	✓	✓
5th-ordered dist.to aiming points	✓	✓	✓	✓
Logged pop.density 1939	✓	✓	✓	✓
Prewar Residenatial Ratio	✓	✓	✓	✓
Observations	137,403	120,495	120,495	120,495
Within Polling Station R^2	0.0854	0.0535	0.0573	0.0281

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Linear probability model except Model (4). In all models except Model 2, population weights are calculated from the neighborhood level population counts divided by the number of users in the sample. Standard errors in parentheses are clustered by neighborhoods in Models 1–3 and by prewar 35 wards in Model 4. All models except Model 1 exclude the users whose cost distance is from 0 to 2.5 minutes. See Table 3 for the selection of days used in the analysis and the description of covariates. Also, see Tables A.6 and A.7 for the full results.

GPS-based vote counts.

Discussion

Our finding that the firebombing of Tokyo lowered contemporary voter turnout may seem to contradict past studies based on the prosocial hypothesis, i.e., that exposure to violence leads to higher levels of political participation (e.g., Bauer et al., 2016). However, a careful evaluation of potential mechanisms proposed in these studies, and how they relate to the sociopolitical context of postwar Tokyo, paints a more nuanced picture. Here, we discuss how the context of postwar Tokyo is different from contexts in which the prosocial hypothesis has been developed, and that the mechanism supported by our finding is a destruction of social capital, which has not previously been extensively discussed.

First, 77 years since the air raids, the first-generation war victims are relatively small in number. Accordingly, post-traumatic growth, one of the main drivers of victims' prosocial behavior, can play only a very limited role in contemporary Tokyo. Second, the process of purging and coping among survivors (Gilligan et al., 2014; Hadzic et al., 2020), another mechanism supporting the prosocial hypothesis, only applies when residents in damaged and undamaged communities have different wartime experience. However, all residents in Tokyo's dense urban neighborhoods faced similar risk of death due to the indiscriminate nature of the bombing. They were also mobilized to communal activities such as procurement of metal and fire drill (Watanabe, 2013), which served as a building block of postwar civic society (Kage, 2010).

Third, the existing literature reports that intergenerational transmission of war memory, whether it is through family, community or government, decreases the support for perpetrators and related parties (Dinas et al., 2021; Lupu and Peisakhin, 2017; Rozenas et al., 2017), but does not affect turnout. Only when major existing parties either support or oppose perpetrators does war memory have an effect on turnout. This is unlikely to be the case in Japan, where political parties take ideological positions broad enough to capture various transmitted emotions and perspectives.

Finally, we argue that "destruction of civic associations" (Gilligan et al., 2014) is the plausible mechanism in contemporary Tokyo. Harada et al. (2021) show that the air raids lowered the organizational strength of neighborhood associations, a key barometer of a neighborhood's geographically-specific social capital. The role of norms and social capital gained through active civil participation and group membership in turnout decisions has been emphasized in previous studies (e.g., Bond et al., 2012; Cox et al., 1998; Dowding et al., 2012; Morton, 1991; Schachar and Nalebuff, 1999; Uhlaner, 1989), and the negative relationship between the damage and turnout coincides with this idea.[33]

Conclusion

Scholars concerned with substantive questions related to voting behavior—whether as explanatory factors or outcomes—have long been hamstrung by the crude availability of the most basic variable of interest: turnout. Although turnout is widely recognized as a marker of democratic health, individual-level data on turnout is often difficult to obtain, or is unreliable due to over-reporting by voters.

Turnout is also rarely studied at the level of small geographical units, such as communities. Although decades of research suggests that the axiom that "all politics is local" applies to voter mobilization, existing studies tend to focus on municipalities, counties, or larger administrative units (e.g., Fiva et al., 2021; Górecki and Marsh, 2012; Key, 1949; Meredith, 2013; Rice and Macht, 1987). This data limitation may prevent scholars from asking questions and obtaining answers pertaining to more fine-grained political processes, how networks function within communities, and the political impacts of localized events.

The approach we have described in this study provides a solution to this problem, using the increasingly available "big data" on cell-phone mobility. We have demonstrated how mobility data can provide reasonable proxies for voter participation at the individual level (which allows for aggregation to any unit of choice), and demonstrated the utility of our estimated measure of turnout for studying two substantive questions of interest to political science: whether the distance to a polling station reduces the likelihood of participating in elections; and whether wartime destruction of communities generates persistent negative effects on voter participation.

Naturally, there are also some limitations to our approach, including the cost of obtaining the data, and necessary safeguards to protect the privacy of cell-phone users. However,

[33] The security dilemma, referring to vicious cycle of growing mutual distrust in this context, can also lower the political participation of affected region (Gilligan et al., 2014). However, this only occurs in a situation where internal conflicts occur, which does not apply to wartime Tokyo.

the potential costs might be greater than the costs of (1) abandoning important questions due to insufficient or poor-quality data; or (2) obtaining inaccurate or imprecise answers to important questions due to the aggregation of turnout and other variables of interest at higher levels. Scholars of voting behavior should weigh these concerns carefully, and consider whether the approach we introduce is suitable for their research designs.[34]

We believe that our approach also opens up potential questions and applications in political science beyond the study of turnout. For example, how many constituents visit the campaign offices of men versus women candidates before and after elections? Where do politicians choose to stage campaign events (such as the street-level oratory sessions common in Japan and other democracies), and do these events help to inform, mobilize, or persuade voters who reside nearby? Numerous future research questions become possible when potential treatments and outcomes are measured at a sufficiently disaggregated level thanks to cell-phone mobility data.

References

Agoop Corporation (2020). Fluid population data of Tokyo [data file]. Detailed information available from https://www.agoop.co.jp/service/dynamic-population-data/.

Aldrich, John (1993). Rational choice and turnout. *American Journal of Political Science 37*(1), 246–278.

Ansolabehere, Stephen and Eitan Hersh (2012). Validation: What big data reveal about survey misreporting and the real electorate. *Political Analysis 20*(4), 437–459.

Anzia, Sarah F. (2014). *Timing and Turnout: How Off-Cycle Elections Favor Organized Groups*. University of Chicago Press.

Arnold, Benjamin F. and Ayse Ercumen (2016). Negative control outcomes: A tool to detect bias in randomized trials. *JAMA 316*(24), 2597–2598.

Bauer, Michal , Christopher Blattman, Julie Chytilová, Joseph Henrich, Edward Miguel, and Tamar Mitts (2016). Can war foster cooperation? *Journal of Economic Perspectives 30*(3), 249–74.

Bellows, John and Edward Miguel (2009). War and local collective action in Sierra Leone. *Journal of Public Economics 93*, 1144–1157.

Bhatti, Yosef (2012). Distance and voting: Evidence from Danish municipalities. *Scandinavian Political Studies 35*(3), 141–158.

Bhatti, Yosef , Kasper M. Hansen, and Hanna Wass (2012). The relationship between age and turnout: A roller-coaster ride. *Electoral Studies 31*(3), 588–593.

Blais, André (2006). What affects voter turnout? *Annual Review of Political Science 9*(1), 111–125.

[34]Appendix C provides a practical checklist for researchers interested in using cell-phone mobility data.

Blattman, Christopher (2009). From Violence to Voting: War and Political Participation in Uganda. *American Political Science Review 103*(2), 231–247.

Bond, Robert M. , Christopher J. Fariss, Jason J. Jones, Adam D. I. Kramer, Cameron Marlow, Jaime E. Settle, and James H. Fowler (2012). A 61-million-person experiment in social influence and political mobilization. *Nature 489*, 295–298.

Brady, Henry E. , Sidney Verba, and Kay Lehman Schlozman (1995). Beyond SES: A resource model of political participation. *American Political Science Review 89*(2), 271–294.

Brakman, Steven , Harry Garretsen, and Marc Schramm (2004). The strategic bombing of German cities during World War II and its impact on city growth. *Journal of Economic Geography 4*, 201–218.

Brockington, David and Jeffrey A. Karp (2005). Social desirability and response validity: A comparative analysis of overreporting voter turnout in five countries. *The Journal of Politics 67*(3), 825–840.

Cancela, João and Benny Geys (2016). Explaining voter turnout: A meta-analysis of national and subnational elections. *Electoral Studies 42*, 264–275.

Cantoni, Enrico (2020). A precinct too far: Turnout and voting costs. *American Economic Journal: Applied Economics 12*(1), 61–85.

Chen, M. Keith , Kareem Haggag, Devin G. Pope, and Ryne Rohla (2019). Racial disparities in voting wait times: Evidence from smartphone data. *Review of Economics and Statistics online first*, 1–27.

Clinton, Joshua , Jon Cohen, John Lapinski, and Marc Trussler (2021). Partisan pandemic: How partisanship and public health concerns affect individuals' social mobility during COVID-19. *Science Advances 7*(2), eabd7204.

Cox, Gary W. , Jon H. Fiva, and Daniel M. Smith (2016). The contraction effect: How proportional representation affects mobilization and turnout. *The Journal of Politics 78*(4), 1249–1263.

Cox, Gary W. , Frances M. Rosenbluth, and Michael F. Thies (1998). Mobilization, social networks, and turnout. *World Politics 50*(3), 447–474.

Davis, Donald R. and David E. Weinstein (2002). Bones, Bombs, and Break Points: The Geography of Economic Activity. *American Economic Review 92*(5), 1269–1289.

Dinas, Elias , Vasiliki Fouka, and Alain Schläpfer (2021). Family history and attitudes toward out-groups: Evidence from the European refugee crisis. *The Journal of Politics 83*(2), 647–661.

Dowding, Keith , Peter John, and Daniel Rubenson (2012). Geographic mobility, social connections and voter turnout. *Journal of Elections, Public Opinion and Parties 22*(2), 109–122.

Downs, Anthony (1957). *An Economic Theory of Democracy.* Harper and Row.

Environmental Systems Research Institute (2021). Arcgis pro 2.9.

Fiva, Jon H. , Askill H. Halse, and Daniel M. Smith (2021). Local representation and voter mobilization in closed-list proportional representation systems. *Quarterly Journal of Political Science 16*(2), 185–213.

Fotheringham, A Stewart and David WS Wong (1991). The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A 23*(7), 1025–1044.

Fowler, Anthony (2013). Electoral and policy consequences of voter turnout: Evidence from compulsory voting in Australia. *Quarterly Journal of Political Science 8*(2), 159–182.

Fowler, James H. , Laura A. Baker, and Christopher T. Dawes (2008). Genetic variation in political participation. *American Political Science Review 102*(2), 233–248.

Franklin, Mark N. (2004). *Voter Turnout and the Dynamics of Electoral Competition in Established Democracies since 1945.* Cambridge University Press.

Garnett, Holly Ann and Sean Grogan (2021). I came, I saw, I voted: Distance to polling locations and voter turnout in Ontario, Canada. *Canadian Journal of Political Science 54*(2), 316–334.

Gibson, John , Bonggeun Kim, Steven Stillman, and Geua Boe-Gibson (2013). Time to vote? *Public Choice 156*(3/4), 517–536.

Gilligan, Michael J. , Benjamin J. Pasquale, and Cyrus Samii (2014). Civil war and social cohesion: Lab-in-the-field evidence from Nepal. *American Journal of Political Science 58*(3), 604–619.

Google (2020). Google COVID-19 community mobility reports. https://www.google.com/covid19/mobility/. [Online; accessed 22-August-2021].

Górecki, Maciej A. and Michael Marsh (2012). Not just 'friends and neighbours': Canvassing, geographic proximity and voter choice. *European Journal of Political Research 51*(5), 563–582.

Hadzic, Dino , David Carlson, and Margit Tavits (2020). How exposure to violence affects ethnic voting. *British Journal of Political Science 50*(1), 345–362.

Hansford, Thomas G. and Brad T. Gomez (2010). Estimating the electoral effects of voter turnout. *American Political Science Review 104*(2), 268–288.

Harada, Masataka , Gaku Ito, and Daniel M. Smith (2021). Destruction from above: Long-term legacies of the Tokyo air raids. http://dx.doi.org/10.2139/ssrn.3471361.

Haspel, Moshe and H. Gibbs Knotts (2005). Location, location, location: Precinct placement and the costs of voting. *The Journal of Politics 67*(2), 560–573.

Hobbs, William and Nazita Lajevardi (2019). Effects of divisive political campaigns on the day-to-day segregation of Arab and Muslim Americans. *American Political Science Review 113*(1), 270–276.

Holbrook, Allyson L. and Jon A. Krosnick (2010). Social desirability bias in voter turnout reports: Tests using the item count technique. *Public Opinion Quarterly 74*(1), 37–67.

Jay, Jonathan , Jacob Bor, Elaine O. Nsoesie, Sarah K. Lipson, David K. Jones, Sandro Galea, and Julia Raifman (2020). Neighbourhood income and physical distancing during the COVID-19 pandemic in the United States. *Nature Human Behaviour 4*(12), 1294–1302.

Kage, Rieko (2010). *Civic Engagement in Postwar Japan: The Revival of a Defeated Society.* Cambridge University Press.

Kage, Rieko (2021). War, democratization, and generational cohort effects on participation in Japan. *Electoral Studies 73*, 102380.

Key, V.O., Jr. (1949). *Southern Politics in State and Nation.* Alfred A. Knopf.

King, Gary (2013). *A Solution to the Ecological Inference Problem.* Princeton University Press.

Kitamura, Shuhei and Tetsuya Matsubayashi (2022). Now or later?: The inter-temporal decision-making of electoral participation. *Political Behavior online first.*

Kocher, Matthew Adam , Thomas B. Pepinsky, and Stathis N. Kalyvas (2011). Aerial bombing and counterinsurgency in the Vietnam War. *American Journal of Political Science 55*(2), 201–218.

Lijphart, Arend (1997). Unequal participation: Democracy's unresolved dilemma. *American Political Science Review 91*(1), 1–14.

Lin, Erin (2022). How war changes land: Soil fertility, unexploded bombs, and the underdevelopment of Cambodia. *American Journal of Political Science 66*(1), 222–237.

Loewen, Peter and Daniel Rubenson (Eds.) (2019). *Duty and Choice: The Evolution of the Study of Voting and Voters.* Toronto University Press.

Lupu, Noam and Leonid Peisakhin (2017). The legacy of political violence across generations. *American Journal of Political Science 61*(4), 836–851.

Martikainen, Pekka , Tuomo Martikainen, and Hanna Wass (2005). The effect of socioeconomic factors on voter turnout in Finland: A register-based study of 2.9 million voters. *European Journal of Political Research 44*(5), 645–669.

Martinez, Michael D. and Jeff Gill (2005). The effects of turnout on partisan outcomes in U.S.

presidential elections 1960–2000. *The Journal of Politics 67*(4), 1248–1274.
Matsusaka, John G. and Filip Palda (1993). The Downsian voter meets the ecological fallacy. *Public Choice 77*, 855–878.
Meredith, Marc (2013). Exploiting friends-and-neighbors to estimate coattail effects. *American Political Science Review 107*(04), 742–765.
Moore, Ryan T. and Andrew Reeves (2020). Defining racial and ethnic context with geolocation data. *Political Science Research and Methods 8*(4), 780–794.
Morton, Rebecca B. (1991). Groups in rational turnout models. *American Journal of Political Science 35*(3), 758–776.
Nilsson, Anton , Carl Bonander, Ulf Strömberg, and Jonas Björk (2020). A directed acyclic graph for interactions. *International Journal of Epidemiology 50*(2), 613–619.
Nishizawa, Yoshitaka (1991). Chihō senkyo ni okeru tōhyōritsu: Gōri-teki yūkensha no tōhyō kōdō [turnout in local elections: Election behavior of rational voters]. *Senkyo Kenkyū [Electoral Studies] 82*(10), 27–44.
Nunn, Nathan and Nancy Qian (2011). The potato's contribution to population and urbanization: Evidence from a historical experiment. *The Quarterly Journal of Economics 126*(2), 593–650.
Nyhan, Brendan , Christopher Skovron, and Rocío Titiunik (2017). Differential registration bias in voter file data: A sensitivity analysis approach. *American Journal of Political Science 61*(3), 744–760.
Rice, Tom W. and Alisa A. Macht (1987). Friends and neighbors voting in statewide general elections. *American Journal of Political Science 31*(2), 448–452.
Riker, William H. and Peter C. Ordeshook (1968). A theory of the calculus of voting. *American Political Science Review 62*, 25–42.
Rotman, Assaf and Michael Shalev (2020). Using location data from mobile phones to study participation in mass protests. *Sociological Methods & Research online first*, 1–56.
Rozenas, Arturas , Sebastian Schutte, and Yuri Zhukov (2017). The political legacy of violence: The long-term impact of Stalin's repression in Ukraine. *The Journal of Politics 79*(4), 1147–1161.
Schachar, Ron and Barry Nalebuff (1999). Follow the leader: Theory and evidence on political participation. *American Economic Review 89*, 525–547.
Secretariat to Election Administration Commission of Tokyo (2019a). Kijitsu-mae tōhyō jōkyō (saishū kekka) [the results of the early in-person voting (final edition)]. https://www.senkyo.metro.tokyo.lg.jp/uploads/r1san_kizitsuzen_saishu.pdf. [Online; accessed 21-September-2021].
Secretariat to Election Administration Commission of Tokyo (2019b). Reiwa gannen 7 gatsu 21 nichi shikkou sangiin giin senkyo kijitsu-mae tōhyōjo ichiran [the list of polling stations for early voting for the upper house election held on July 21, 2019]. https://www.senkyo.metro.tokyo.lg.jp/uploads/02_r1sangiin-kizitsuzen-list.pdf. [Online; accessed 21-December-2021].
Secretariat to Election Administration Commission of Tokyo (2019c). Sangiin giin senkyo tōhyō kekka [the results of the House of Councillors election (for proportional representation)]. https://www.senkyo.metro.tokyo.lg.jp/election/sanngiin-all/sanngiin-sokuhou2019/. [Online; accessed 21-September-2021].
Sobolev, Anton , M. Keith Chen, Jungseock Joo, and Zachary C. Steinert-Threlkeld (2020). News and geolocated social media accurately measure protest size variation. *American Political Science Review 114*(4), 1343–1351.
Steinert-Threlkeld, Zachary C. (2017). Spontaneous collective action: Peripheral mobilization during the Arab Spring. *American Political Science Review 111*(2), 379–403.
Tingsten, Herbert (1937). *Political Behavior: Studies in Election Statistics*. P.S. King & Sons.
Uhlaner, Carole J. (1989). Rational turnout: The neglected role of groups. *American Journal of*

Political Science 33(2), 390–422.
Walden, Jacob and Yuri M. Zhukov (2020). Historical legacies of political violence. *Oxford Research Encyclopedia of Politics*.
Watanabe, Yōkichi (2013). *Senjika no Nihonjin to Tonarigumi Kaihō [Japanese and Neighborhood Associations During the War]*. Gentosha Renaissance.
Wolfinger, Raymond E. and Steven J. Rosenstone (1980). *Who Votes?* Yale University Press.

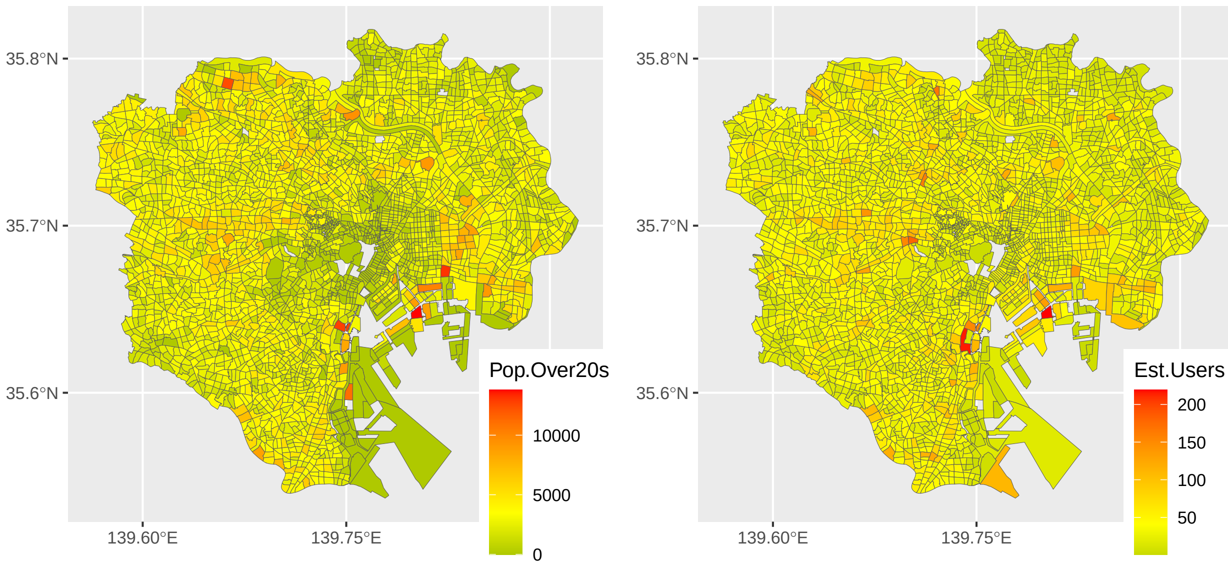
Appendices

Appendix A Supplementary Tables and Figures

Table A.1: Fictitious Sample of Track Record Data from Agoop

dailyid	year	month	day	hour	minute	latitude	longitude	os	accuracy	citycode	mesh100mid
899a93-omitted	2019	7	21	8	6	35.638845	139.766978	Android	14.695	13102	5339460000
899a93-omitted	2019	7	21	8	18	35.638845	139.766978	Android	14.695	13102	5339460000
899a93-omitted	2019	7	21	8	23	35.638845	139.766978	Android	14.695	13102	5339460000
899a93-omitted	2019	7	21	8	29	35.638845	139.766978	Android	14.695	13102	5339460000
899a93-omitted	2019	7	21	8	47	35.638845	139.766978	Android	14.695	13102	5339460000
899a93-omitted	2019	7	21	8	54	35.637945	139.767697	Android	14.69	13102	5339461111
899a93-omitted	2019	7	21	9	10	35.637945	139.767697	Android	14.69	13102	5339461111
899a93-omitted	2019	7	21	9	22	35.637945	139.767697	Android	14.69	13102	5339461111
899a93-omitted	2019	7	21	9	23	35.638945	139.767978	Android	14.895	13102	5339462222
899a93-omitted	2019	7	21	9	46	35.638945	139.767978	Android	14.895	13102	5339462222
899a93-omitted	2019	7	21	9	47	35.638945	139.767978	Android	14.895	13102	5339462222
899a93-omitted	2019	7	21	10	12	35.638945	139.767978	Android	14.895	13102	5339462222
899a93-omitted	2019	7	21	10	16	35.638945	139.767978	Android	14.895	13102	5339462222
899a93-omitted	2019	7	21	10	25	35.638927	139.767983	Android	15.251	13102	5339463333
899a93-omitted	2019	7	21	10	27	35.63893	139.767987	Android	19.322	13102	5339464444
899a93-omitted	2019	7	21	10	36	35.63893	139.767987	Android	15.322	13102	5339464444
899a93-omitted	2019	7	21	10	46	35.63893	139.767987	Android	15.322	13102	5339464444
899a93-omitted	2019	7	21	10	59	35.63893	139.767987	Android	15.322	13102	5339464444

Notes: This fictitious data sample was created to illustrate a set of data points from 8:00 to 11:00 for one user. The actual dailyid is a combination of 96 alphabetic characters from a to f and numbers.



(a) Census Population Counts

(b) GPS-based Population Counts

Figure A.1: Performance check based on population for the election day ($\rho = .69$)

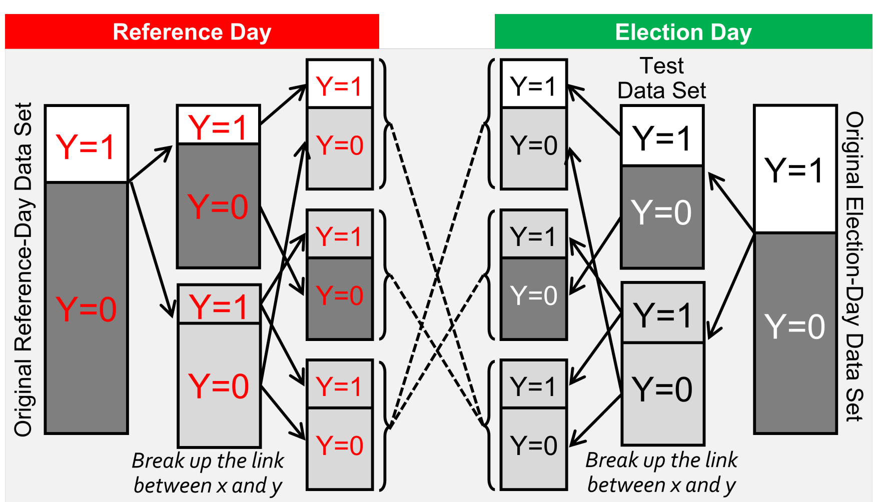


Figure A.2: Flow chart of constructing randomized data sets for the sensitivity analysis

Table A.2: Difference-in-differences regression of voting on election day on the estimated minutes to a polling station (full results)

Outcome Variable: Estimated Vote	(1)		(2)	
	b	s.e.	b	s.e.
Election day	0.157***	(0.006)	0.115***	(0.008)
Min. to PS: 2.5 - 5	−0.114***	(0.004)	−0.114***	(0.004)
Min. to PS: 5 - 7.5	−0.135***	(0.004)	−0.135***	(0.004)
Min. to PS: 7.5 - 10	−0.145***	(0.004)	−0.145***	(0.004)
Min. to PS: 10 - 12.5	−0.146***	(0.004)	−0.145***	(0.004)
Min. to PS: 12.5 - 15	−0.144***	(0.004)	−0.144***	(0.004)
Min. to PS: 15 - 17.5	−0.143***	(0.005)	−0.142***	(0.005)
Min. to PS: 17.5 - 20	−0.122***	(0.008)	−0.122***	(0.008)
Min. to PS: 20 - 22.5	−0.133***	(0.007)	−0.135***	(0.007)
Min. to PS: 22.5 - 25	−0.145***	(0.007)	−0.146***	(0.007)
Min. to PS: 25 - 27.5	−0.128***	(0.008)	−0.130***	(0.009)
Min. to PS: 27.5 - 30	−0.136***	(0.007)	−0.146***	(0.007)
Election day × Min. to PS: 2.5 - 5	−0.047***	(0.006)	−0.047***	(0.006)
Election day × Min. to PS: 5 - 7.5	−0.072***	(0.006)	−0.074***	(0.006)
Election day × Min. to PS: 7.5 - 10	−0.090***	(0.006)	−0.094***	(0.006)
Election day × Min. to PS: 10 - 12.5	−0.108***	(0.006)	−0.114***	(0.006)
Election day × Min. to PS: 12.5 - 15	−0.124***	(0.007)	−0.134***	(0.007)
Election day × Min. to PS: 15 - 17.5	−0.115***	(0.010)	−0.131***	(0.010)
Election day × Min. to PS: 17.5 - 20	−0.151***	(0.011)	−0.176***	(0.011)
Election day × Min. to PS: 20 - 22.5	−0.143***	(0.010)	−0.169***	(0.011)
Election day × Min. to PS: 22.5 - 25	−0.134***	(0.011)	−0.162***	(0.011)
Election day × Min. to PS: 25 - 27.5	−0.157***	(0.009)	−0.196***	(0.010)
Election day × Min. to PS: 27.5 - 30	−0.156***	(0.006)	−0.176***	(0.007)
Min. to Early PS: 2.5 - 5			0.008**	(0.004)
Min. to Early PS: 5 - 7.5			0.012***	(0.004)
Min. to Early PS: 7.5 - 10			0.015***	(0.004)
Min. to Early PS: 10 - 12.5			0.017***	(0.004)
Min. to Early PS: 12.5 - 15			0.014***	(0.004)
Min. to Early PS: 15 - 17.5			0.017***	(0.004)
Min. to Early PS: 17.5 - 20			0.016***	(0.005)
Min. to Early PS: 20 - 22.5			0.015***	(0.005)
Min. to Early PS: 22.5 - 25			0.015**	(0.006)
Min. to Early PS: 25 - 27.5			0.011	(0.008)
Min. to Early PS: 27.5 - 30			0.009	(0.009)
Min. to Early PS: Over 30			0.031***	(0.009)
Election day × Min. to Early PS: 2.5 - 5			0.025***	(0.007)
Election day × Min. to Early PS: 5 - 7.5			0.032***	(0.006)
Election day × Min. to Early PS: 7.5 - 10			0.039***	(0.006)
Election day × Min. to Early PS: 10 - 12.5			0.047***	(0.006)
Election day × Min. to Early PS: 12.5 - 15			0.056***	(0.007)
Election day × Min. to Early PS: 15 - 17.5			0.063***	(0.007)

Continued on next page

Table A.2 – continued from previous page

Outcome Variable: Estimated Vote	b	s.e.	b	s.e.
Election day × Min. to Early PS: 17.5 - 20			0.065***	(0.008)
Election day × Min. to Early PS: 20 - 22.5			0.074***	(0.009)
Election day × Min. to Early PS: 22.5 - 25			0.086***	(0.010)
Election day × Min. to Early PS: 25 - 27.5			0.073***	(0.013)
Election day × Min. to Early PS: 27.5 - 30			0.082***	(0.015)
Election day × Min. to Early PS: Over 30			0.119***	(0.015)
Constant	0.148***	(0.004)	0.135***	(0.005)
Election-Day Polling Station Fixed Effects	✓		✓	
Observations	180,248		180,248	
Within Polling Station R^2	0.101		0.103	

Notes: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$. Robust standard errors are in parentheses. Election day and reference day indicate July 21, 2019 and July 28, 2019, respectively. “Min.” and “PS” represent minutes and polling stations, respectively.

Table A.3: Difference-in-differences regression of estimated election-day turnout on the level of damages from the Tokyo firebombing (full results)

Outcome Variable: Estimated Vote	(1)		(2)	
	b	s.e.	b	s.e.
Damage × Election Day	−0.0138**	(0.0054)	−0.0206**	(0.0081)
Election Day	0.1135***	(0.0032)	0.7361*	(0.3921)
Min. to PS: 2.5 - 5			0.0674***	(0.0237)
Min. to PS: 5 - 7.5			0.0381	(0.0236)
Min. to PS: 7.5 - 10			0.0221	(0.0236)
Min. to PS: 10 - 12.5			0.0146	(0.0237)
Min. to PS: 12.5 - 15			0.0083	(0.0236)
Min. to PS: 15 - 17.5			−0.0048	(0.0249)
Min. to PS: 17.5 - 20			0.0295	(0.0263)
Min. to PS: 20 - 22.5			0.0007	(0.0274)
Min. to PS: 22.5 - 25			−0.0090	(0.0172)
Min. to PS: 25 - 27.5			−0.0036	(0.0205)
Min. to PS: 27.5 - 30			0.0000	(omitted)
Election Day × Min. to PS: 2.5 - 5			0.1112***	(0.0165)
Election Day × Min. to PS: 5 - 7.5			0.0810***	(0.0165)
Election Day × Min. to PS: 7.5 - 10			0.0581***	(0.0165)
Election Day × Min. to PS: 10 - 12.5			0.0207	(0.0170)
Election Day × Min. to PS: 12.5 - 15			0.0020	(0.0184)
Election Day × Min. to PS: 15 - 17.5			0.0284	(0.0269)
Election Day × Min. to PS: 17.5 - 20			−0.0242	(0.0230)

Continued on next page

Table A.3 – continued from previous page

Outcome Variable: Estimated Vote	b	s.e.	b	s.e.
Election Day × Min. to PS: 20 - 22.5			−0.0217	(0.0222)
Election Day × Min. to PS: 22.5 - 25			−0.0080	(0.0359)
Election Day × Min. to PS: 25 - 27.5			−0.0644**	(0.0293)
Election Day × Min. to PS: 27.5 - 30			0.0000	(omitted)
Min. to Early PS: 2.5 - 5			0.0013	(0.0049)
Min. to Early PS: 5 - 7.5			0.0043	(0.0053)
Min. to Early PS: 7.5 - 10			0.0048	(0.0056)
Min. to Early PS: 10 - 12.5			0.0101	(0.0062)
Min. to Early PS: 12.5 - 15			0.0131*	(0.0069)
Min. to Early PS: 15 - 17.5			0.0145*	(0.0082)
Min. to Early PS: 17.5 - 20			0.0175*	(0.0097)
Min. to Early PS: 20 - 22.5			0.0152	(0.0117)
Min. to Early PS: 22.5 - 25			0.0395***	(0.0139)
Min. to Early PS: 25 - 27.5			0.0471***	(0.0174)
Min. to Early PS: 27.5 - 30			0.0280	(0.0213)
Min. to Early PS: Over 30			0.0777***	(0.0264)
Election Day × Min. to Early PS: 2.5 - 5			0.0242***	(0.0094)
Election Day × Min. to Early PS: 5 - 7.5			0.0311***	(0.0090)
Election Day × Min. to Early PS: 7.5 - 10			0.0449***	(0.0092)
Election Day × Min. to Early PS: 10 - 12.5			0.0471***	(0.0093)
Election Day × Min. to Early PS: 12.5 - 15			0.0542***	(0.0098)
Election Day × Min. to Early PS: 15 - 17.5			0.0553***	(0.0108)
Election Day × Min. to Early PS: 17.5 - 20			0.0612***	(0.0132)
Election Day × Min. to Early PS: 20 - 22.5			0.0566***	(0.0151)
Election Day × Min. to Early PS: 22.5 - 25			0.0709***	(0.0200)
Election Day × Min. to Early PS: 25 - 27.5			0.0638**	(0.0290)
Election Day × Min. to Early PS: 27.5 - 30			0.0379	(0.0281)
Election Day × Min. to Early PS: Over 30			0.0269	(0.0271)
Election Day × Std. Longitude^1			0.2986*	(0.1806)
Election Day × Std. Longitude^2			−0.3862	(0.2398)
Election Day × Std. Longitude^3			−0.0461	(0.0306)
Election Day × Std. Longitude^4			0.0264	(0.0197)
Election Day × Std. Longitude^5			0.0012	(0.0014)
Election Day × Std. Latitude^1			0.0531***	(0.0186)
Election Day × Std. Latitude^2			−0.4504	(0.2743)
Election Day × Std. Latitude^3			−0.0424***	(0.0097)
Election Day × Std. Latitude^4			0.0383	(0.0242)
Election Day × Std. Latitude^5			0.0068***	(0.0018)
Election Day × Std. Longitude^1×Std. Latitude^1			−0.0036	(0.0117)
Election Day × Std. Longitude^2×Std. Latitude^1			−0.0110	(0.0142)
Election Day × Std. Longitude^3×Std. Latitude^1			−0.0012	(0.0041)
Election Day × Std. Longitude^4×Std. Latitude^1			−0.0001	(0.0032)
Election Day × Std. Longitude^1×Std. Latitude^2			−0.0147	(0.0375)
Election Day × Std. Longitude^2×Std. Latitude^2			0.0662	(0.0436)

Continued on next page

Table A.3 – continued from previous page

Outcome Variable: Estimated Vote	b	s.e.	b	s.e.
Election Day × Std. Longitude^3×Std. Latitude^2			0.0002	(0.0048)
Election Day × Std. Longitude^1×Std. Latitude^3			0.0036	(0.0045)
Election Day × Std. Longitude^2×Std. Latitude^3			0.0022	(0.0048)
Election Day × Std. Longitude^1×Std. Latitude^4			−0.0106***	(0.0038)
Election Day × Dist. to Imperial Palace^1			0.3978*	(0.2352)
Election Day × Dist. to Aiming Point^1			0.0042	(0.0095)
Election Day × Dist. to Imperial Palace^2			0.0000	(omitted)
Election Day × Dist. to Aiming Point^2			0.0098*	(0.0053)
Election Day × Dist. to Imperial Palace^3			−0.0329	(0.0213)
Election Day × Dist. to Aiming Point^3			0.0021	(0.0027)
Election Day × Dist. to Imperial Palace^4			0.0000	(omitted)
Election Day × Dist. to Aiming Point^4			−0.0009	(0.0020)
Election Day × Dist. to Imperial Palace^5			0.0036*	(0.0021)
Election Day × Dist. to Aiming Point^5			−0.0007	(0.0017)
Election Day × logged Prewar Pop. Density			−0.0091***	(0.0033)
Election Day × Ratio of Residential Area			−0.0157	(0.0106)
Constant	0.0232***	(0.0010)	−0.0283	(0.0244)
Election-Day Polling Station Fixed Effects	✓		✓	
Observations	120,495		120,495	
Within Polling Station R^2	0.0417		0.0573	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Linear probability model with population weights. Standard errors clustered by neighborhoods in parentheses. All models exclude the users whose cost distance is from 0 to 2.5 minutes. Difference-in-differences estimation is performed between the election day (July 21, 2019) and the reference day (July 28, 2019.) "Min.", "PS", "Dist." and "Pop." represent minutes, polling stations, distance and population, respectively.

Table A.4: Difference-in-differences regression of estimated turnout (Model 1) and official turnout (Model 3) on the level of damages from the Tokyo firebombing using the data aggregated at the level of a voting precinct without control variables (full results)

Outcome Variable:	Estimated Turnout (1)		Official Turnout (3)	
	b	s.e.	b	s.e.
Damage × Election Day	−0.0101	(0.0072)	−0.0066*	(0.0040)
Election Day	0.1235***	(0.0044)	0.5172***	(0.0022)
Constant	0.0382***	(0.0013)	−0.0000	(0.0007)
Election-Day Polling Station Fixed Effects	✓		✓	
Observations	1,878		1,878	

Continued on next page

Table A.4 – continued from previous page

	b	s.e.	b	s.e.
Within Polling Station R^2	0.7208		0.9944	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. The number of voters is used as weights. Difference-in-differences estimation is performed between the election day (July 21, 2019) and the reference day (July 28, 2019). Table 3 for the description of the covariates.

Table A.5: Difference-in-differences regression of estimated turnout (Model 2) and official turnout (Model 4) on the level of damages from the Tokyo firebombing using the data aggregated at the level of a voting precinct with control variables (full results)

Outcome Variable:	Estimated Turnout (2)		Official Turnout (4)	
	b	s.e.	b	s.e.
Damage × Election Day	−0.0151	(0.0112)	−0.0202***	(0.0050)
Election Day	0.2628	(4.9739)	0.4400	(2.5663)
Ave. Minutes to PS	−0.0320***	(0.0067)	−0.0047**	(0.0022)
Ave. Minutes to Early PS	0.0004	(0.0003)	0.0000	(0.0002)
Election Day × Ave. Minutes to PS	−0.0083***	(0.0012)	−0.0004	(0.0005)
Election Day × Ave. Minutes to Early PS	−0.0000	(0.0000)	−0.0000	(0.0000)
Election Day × Std. Longitude^1	1.8978	(2.8983)	0.4410	(1.4632)
Election Day × Std. Longitude^2	−2.5964	(4.2267)	−0.5475	(2.1568)
Election Day × Std. Longitude^3	−1.2908	(1.3813)	−0.2999	(0.5941)
Election Day × Std. Longitude^4	1.0268	(1.0943)	0.2322	(0.4703)
Election Day × Std. Longitude^5	−0.0015	(0.0024)	−0.0030**	(0.0012)
Election Day × Std. Latitude^1	−0.3744	(0.6358)	−0.0614	(0.3213)
Election Day × Std. Latitude^2	−2.7561	(4.1971)	−0.5878	(2.1205)
Election Day × Std. Latitude^3	0.2244	(0.2785)	0.0468	(0.1196)
Election Day × Std. Latitude^4	0.8738	(0.9238)	0.2070	(0.3970)
Election Day × Std. Latitude^5	0.0054***	(0.0019)	0.0039***	(0.0008)
Election Day × Std. Longitude^1×Std. Latitude^1	−0.1665	(0.1954)	0.0107	(0.0841)
Election Day × Std. Longitude^2×Std. Latitude^1	0.2684	(0.3047)	−0.0094	(0.1317)
Election Day × Std. Longitude^3×Std. Latitude^1	−0.0012	(0.0047)	−0.0090***	(0.0023)
Election Day × Std. Longitude^4×Std. Latitude^1	0.0026	(0.0047)	0.0172***	(0.0022)
Election Day × Std. Longitude^1×Std. Latitude^2	−1.1725	(1.2695)	−0.3050	(0.5453)
Election Day × Std. Longitude^2×Std. Latitude^2	1.8920	(2.0123)	0.4205	(0.8646)
Election Day × Std. Longitude^3×Std. Latitude^2	−0.0021	(0.0060)	0.0159***	(0.0027)
Election Day × Std. Longitude^1×Std. Latitude^3	−0.0017	(0.0059)	−0.0168***	(0.0025)
Election Day × Std. Longitude^2×Std. Latitude^3	0.0044	(0.0063)	0.0050*	(0.0026)
Election Day × Std. Longitude^1×Std. Latitude^4	−0.0082**	(0.0040)	−0.0042**	(0.0018)
Election Day × Dist. to Imperial Palace^1	−2.1408	(3.6907)	−0.5802	(1.6262)

Continued on next page

Table A.5 – continued from previous page

	b	s.e.	b	s.e.
Election Day × Dist. to Aiming Point^1	0.0153	(0.0119)	0.0488***	(0.0055)
Election Day × Dist. to Imperial Palace^2	−2.4917	(2.5571)	−0.5919	(1.0147)
Election Day × Dist. to Aiming Point^2	0.0131**	(0.0063)	0.0127***	(0.0026)
Election Day × Dist. to Imperial Palace^3	−0.9224	(0.9778)	−0.2109	(0.4200)
Election Day × Dist. to Aiming Point^3	0.0019	(0.0031)	−0.0063***	(0.0014)
Election Day × Dist. to Imperial Palace^4	−0.0976	(0.1088)	−0.0177	(0.0466)
Election Day × Dist. to Aiming Point^4	−0.0011	(0.0023)	−0.0006	(0.0010)
Election Day × Dist. to Imperial Palace^5	0.0025	(0.0025)	0.0002	(0.0013)
Election Day × Dist. to Aiming Point^5	0.0012	(0.0022)	0.0002	(0.0010)
Election Day × logged Prewar Pop. Density	−0.0091**	(0.0042)	−0.0004	(0.0016)
Election Day × Ratio of Residential Area	−0.0289*	(0.0152)	−0.0003	(0.0064)
Constant	0.2287***	(0.0408)	0.0280**	(0.0134)
Election-Day Polling Station Fixed Effects	✓		✓	
Observations	1,878		1,878	
Within Polling Station R^2	0.7774		0.9969	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors in parentheses. The number of voters is used as weights. Difference-in-differences estimation is performed between the election day (July 21, 2019) and the reference day (July 28, 2019.). Control variables are weighted averages of the values from the cell phone user's estimated address for each polling station. Table 3 for the description of the covariates except cost distance variables and their interactions, which are entered into these models as continuous variables. "Ave.", "PS", "Std.", "Dist." and "Pop." represent average, polling stations, standardized, distance and population, respectively.

Table A.6: Summary of the robustness checks: estimation with observations very close to polling stations (Model 1) and without population weights (Model 2) (full results)

Outcome Variable: Estimated Vote	(1)		(2)	
	b	s.e.	b	s.e.
Damage × Election Day	−0.0168**	(0.0082)	−0.0137*	(0.0075)
Election Day	1.0359***	(0.3945)	0.9120**	(0.3703)
Min. to PS: 2.5 - 5	−0.1312***	(0.0060)	0.0522***	(0.0197)
Min. to PS: 5 - 7.5	−0.1599***	(0.0064)	0.0253	(0.0196)
Min. to PS: 7.5 - 10	−0.1765***	(0.0066)	0.0109	(0.0196)
Min. to PS: 10 - 12.5	−0.1843***	(0.0073)	0.0041	(0.0198)
Min. to PS: 12.5 - 15	−0.1921***	(0.0088)	0.0027	(0.0198)
Min. to PS: 15 - 17.5	−0.2005***	(0.0117)	−0.0101	(0.0214)
Min. to PS: 17.5 - 20	−0.1618***	(0.0214)	0.0156	(0.0232)
Min. to PS: 20 - 22.5	−0.1982***	(0.0267)	0.0085	(0.0289)

Continued on next page

Table A.6 – continued from previous page

Outcome Variable: Estimated Vote	b	s.e.	b	s.e.
Min. to PS: 22.5 - 25	−0.2211***	(0.0219)	−0.0259	(0.0214)
Min. to PS: 25 - 27.5	−0.2103***	(0.0223)	0.0133	(0.0230)
Min. to PS: 27.5 - 30	−0.2036***	(0.0239)	0.0000	(omitted)
Election Day × Min. to PS: 2.5 - 5	−0.0421***	(0.0073)	0.1303***	(0.0094)
Election Day × Min. to PS: 5 - 7.5	−0.0725***	(0.0072)	0.1020***	(0.0093)
Election Day × Min. to PS: 7.5 - 10	−0.0948***	(0.0076)	0.0778***	(0.0095)
Election Day × Min. to PS: 10 - 12.5	−0.1321***	(0.0084)	0.0460***	(0.0100)
Election Day × Min. to PS: 12.5 - 15	−0.1507***	(0.0110)	0.0217*	(0.0113)
Election Day × Min. to PS: 15 - 17.5	−0.1242***	(0.0236)	0.0441**	(0.0216)
Election Day × Min. to PS: 17.5 - 20	−0.1789***	(0.0188)	−0.0029	(0.0169)
Election Day × Min. to PS: 20 - 22.5	−0.1745***	(0.0202)	0.0003	(0.0171)
Election Day × Min. to PS: 22.5 - 25	−0.1623***	(0.0387)	0.0447	(0.0364)
Election Day × Min. to PS: 25 - 27.5	−0.2210***	(0.0274)	−0.0312	(0.0198)
Election Day × Min. to PS: 27.5 - 30	−0.1597***	(0.0162)	0.0000	(omitted)
Min. to Early PS: 2.5 - 5	−0.0027	(0.0064)	0.0006	(0.0045)
Min. to Early PS: 5 - 7.5	0.0028	(0.0066)	0.0033	(0.0048)
Min. to Early PS: 7.5 - 10	0.0036	(0.0069)	0.0033	(0.0052)
Min. to Early PS: 10 - 12.5	0.0087	(0.0074)	0.0089	(0.0057)
Min. to Early PS: 12.5 - 15	0.0066	(0.0081)	0.0112*	(0.0064)
Min. to Early PS: 15 - 17.5	0.0052	(0.0094)	0.0114	(0.0076)
Min. to Early PS: 17.5 - 20	0.0072	(0.0112)	0.0169*	(0.0089)
Min. to Early PS: 20 - 22.5	0.0088	(0.0134)	0.0142	(0.0106)
Min. to Early PS: 22.5 - 25	0.0330**	(0.0151)	0.0398***	(0.0121)
Min. to Early PS: 25 - 27.5	0.0289	(0.0238)	0.0422***	(0.0151)
Min. to Early PS: 27.5 - 30	0.0100	(0.0231)	0.0372**	(0.0177)
Min. to Early PS: Over 30	0.0454*	(0.0268)	0.0676***	(0.0215)
Election Day × Min. to Early PS: 2.5 - 5	0.0350***	(0.0090)	0.0247***	(0.0081)
Election Day × Min. to Early PS: 5 - 7.5	0.0421***	(0.0086)	0.0337***	(0.0077)
Election Day × Min. to Early PS: 7.5 - 10	0.0500***	(0.0090)	0.0491***	(0.0080)
Election Day × Min. to Early PS: 10 - 12.5	0.0538***	(0.0091)	0.0511***	(0.0080)
Election Day × Min. to Early PS: 12.5 - 15	0.0632***	(0.0095)	0.0579***	(0.0086)
Election Day × Min. to Early PS: 15 - 17.5	0.0617***	(0.0105)	0.0604***	(0.0095)
Election Day × Min. to Early PS: 17.5 - 20	0.0720***	(0.0131)	0.0614***	(0.0116)
Election Day × Min. to Early PS: 20 - 22.5	0.0687***	(0.0145)	0.0631***	(0.0128)
Election Day × Min. to Early PS: 22.5 - 25	0.0790***	(0.0199)	0.0755***	(0.0169)
Election Day × Min. to Early PS: 25 - 27.5	0.0704**	(0.0276)	0.0645***	(0.0247)
Election Day × Min. to Early PS: 27.5 - 30	0.0510*	(0.0308)	0.0267	(0.0323)
Election Day × Min. to Early PS: Over 30	0.0482	(0.0309)	0.0382*	(0.0220)
Election Day × Std. Longitude^1	0.3672**	(0.1819)	0.4016**	(0.1706)
Election Day × Std. Longitude^2	−0.4827**	(0.2418)	−0.5192**	(0.2265)
Election Day × Std. Longitude^3	−0.0552*	(0.0307)	−0.0617**	(0.0290)
Election Day × Std. Longitude^4	0.0333*	(0.0198)	0.0357*	(0.0186)
Election Day × Std. Longitude^5	0.0012	(0.0014)	0.0014	(0.0014)
Election Day × Std. Latitude^1	0.0628***	(0.0188)	0.0590***	(0.0176)

Continued on next page

Table A.6 – continued from previous page

Outcome Variable: Estimated Vote	b	s.e.	b	s.e.
Election Day × Std. Latitude^2	−0.5601**	(0.2767)	−0.6045**	(0.2591)
Election Day × Std. Latitude^3	−0.0476***	(0.0099)	−0.0439***	(0.0092)
Election Day × Std. Latitude^4	0.0481**	(0.0244)	0.0523**	(0.0228)
Election Day × Std. Latitude^5	0.0075***	(0.0019)	0.0073***	(0.0018)
Election Day × Std. Longitude^1×Std. Latitude^1	0.0147	(0.0119)	0.0094	(0.0111)
Election Day × Std. Longitude^2×Std. Latitude^1	−0.0133	(0.0144)	−0.0133	(0.0137)
Election Day × Std. Longitude^3×Std. Latitude^1	−0.0038	(0.0042)	−0.0036	(0.0038)
Election Day × Std. Longitude^4×Std. Latitude^1	−0.0001	(0.0033)	0.0006	(0.0030)
Election Day × Std. Longitude^1×Std. Latitude^2	−0.0292	(0.0373)	−0.0423	(0.0358)
Election Day × Std. Longitude^2×Std. Latitude^2	0.0792*	(0.0442)	0.0843**	(0.0411)
Election Day × Std. Longitude^3×Std. Latitude^2	0.0022	(0.0047)	0.0030	(0.0046)
Election Day × Std. Longitude^1×Std. Latitude^3	−0.0017	(0.0044)	0.0009	(0.0043)
Election Day × Std. Longitude^2×Std. Latitude^3	0.0037	(0.0049)	0.0032	(0.0046)
Election Day × Std. Longitude^1×Std. Latitude^4	−0.0104***	(0.0038)	−0.0091**	(0.0036)
Election Day × Dist. to Imperial Palace^1	0.4918**	(0.2372)	0.5338**	(0.2221)
Election Day × Dist. to Aiming Point^1	0.0120	(0.0094)	0.0127	(0.0093)
Election Day × Dist. to Imperial Palace^2	0.0000	(omitted)	0.0000	(omitted)
Election Day × Dist. to Aiming Point^2	0.0119**	(0.0052)	0.0107**	(0.0051)
Election Day × Dist. to Imperial Palace^3	−0.0437**	(0.0215)	−0.0456**	(0.0202)
Election Day × Dist. to Aiming Point^3	0.0023	(0.0027)	0.0009	(0.0026)
Election Day × Dist. to Imperial Palace^4	0.0000	(omitted)	0.0000	(omitted)
Election Day × Dist. to Aiming Point^4	−0.0007	(0.0019)	−0.0013	(0.0018)
Election Day × Dist. to Imperial Palace^5	0.0040*	(0.0021)	0.0046**	(0.0020)
Election Day × Dist. to Aiming Point^5	−0.0005	(0.0017)	−0.0009	(0.0016)
Election Day × logged Prewar Pop. Density	−0.0092***	(0.0033)	−0.0078**	(0.0033)
Election Day × Ratio of Residential Area	−0.0140	(0.0107)	−0.0151	(0.0100)
Constant	0.1724***	(0.0083)	−0.0147	(0.0205)
Election-Day Polling Station Fixed Effects	✓		✓	
Observations	137,403		120,495	
Within Polling Station R^2	0.0854		0.0535	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Linear probability model. In Model 1, population weights are calculated from the neighborhood level population counts divided by the number of users in the sample. Standard errors in parentheses are clustered by neighborhoods. Model 2 exclude the users whose cost distance is from 0 to 2.5 minutes. Difference-in-differences estimation is performed between the election day (July 21, 2019) and the reference day (July 28, 2019.) See Table 3 for the description of covariates. “Min.”, “PS”, “Std.”, “Dist.” and “Pop.” represent minutes, polling stations, standardized, distance and population, respectively.

Table A.7: Summary of the robustness checks: estimation with SEs clustered by prewar 35 wards (Model 3), and with cost distance to election-day polling station as a negative control outcome (Model 4) (full results)

Outcome Variable:	Estimated Vote (3)		Cost-distance to Election-day PS (4)	
	b	s.e.	b	s.e.
Damage × Election Day	−0.0206**	(0.0094)	−0.0330	(0.0541)
Election Day	0.0222	(0.0328)	−0.4441	(2.4654)
Min. to PS: 2.5 - 5	0.0674***	(0.0184)		
Min. to PS: 5 - 7.5	0.0381**	(0.0181)		
Min. to PS: 7.5 - 10	0.0221	(0.0175)		
Min. to PS: 10 - 12.5	0.0146	(0.0170)		
Min. to PS: 12.5 - 15	0.0083	(0.0187)		
Min. to PS: 15 - 17.5	−0.0048	(0.0189)		
Min. to PS: 17.5 - 20	0.0295	(0.0186)		
Min. to PS: 20 - 22.5	0.0007	(0.0143)		
Min. to PS: 22.5 - 25	−0.0090	(0.0138)		
Min. to PS: 25 - 27.5	−0.0036	(0.0208)		
Min. to PS: 27.5 - 30	0.0000	(omitted)		
Election Day × Min. to PS: 2.5 - 5	0.1756***	(0.0263)		
Election Day × Min. to PS: 5 - 7.5	0.1454***	(0.0259)		
Election Day × Min. to PS: 7.5 - 10	0.1226***	(0.0267)		
Election Day × Min. to PS: 10 - 12.5	0.0851***	(0.0259)		
Election Day × Min. to PS: 12.5 - 15	0.0665**	(0.0269)		
Election Day × Min. to PS: 15 - 17.5	0.0928***	(0.0326)		
Election Day × Min. to PS: 17.5 - 20	0.0402	(0.0295)		
Election Day × Min. to PS: 20 - 22.5	0.0427	(0.0280)		
Election Day × Min. to PS: 22.5 - 25	0.0564*	(0.0323)		
Election Day × Min. to PS: 25 - 27.5	0.0000	(omitted)		
Election Day × Min. to PS: 27.5 - 30	0.0644**	(0.0306)		
Min. to Early PS: 2.5 - 5	0.0013	(0.0044)	−0.1621	(0.1141)
Min. to Early PS: 5 - 7.5	0.0043	(0.0052)	0.1982	(0.1374)
Min. to Early PS: 7.5 - 10	0.0048	(0.0057)	0.5080***	(0.1514)
Min. to Early PS: 10 - 12.5	0.0101	(0.0063)	0.8259***	(0.1706)
Min. to Early PS: 12.5 - 15	0.0131*	(0.0065)	1.1953***	(0.1957)
Min. to Early PS: 15 - 17.5	0.0145*	(0.0074)	1.7502***	(0.2385)
Min. to Early PS: 17.5 - 20	0.0175*	(0.0098)	2.2290***	(0.3130)
Min. to Early PS: 20 - 22.5	0.0152	(0.0098)	2.9262***	(0.4465)
Min. to Early PS: 22.5 - 25	0.0395***	(0.0091)	3.8809***	(0.6172)
Min. to Early PS: 25 - 27.5	0.0471***	(0.0139)	4.4436***	(0.9201)
Min. to Early PS: 27.5 - 30	0.0280**	(0.0132)	5.0894***	(1.0874)
Min. to Early PS: Over 30	0.0777**	(0.0295)	6.4062***	(2.0674)
Election Day × Min. to Early PS: 2.5 - 5	0.0242***	(0.0081)	0.0582	(0.0591)
Election Day × Min. to Early PS: 5 - 7.5	0.0311***	(0.0094)	0.0217	(0.0557)

Continued on next page

Table A.7 – continued from previous page

	b	s.e.	b	s.e.
Election Day × Min. to Early PS: 7.5 - 10	0.0449***	(0.0087)	0.0389	(0.0585)
Election Day × Min. to Early PS: 10 - 12.5	0.0471***	(0.0078)	−0.0031	(0.0588)
Election Day × Min. to Early PS: 12.5 - 15	0.0542***	(0.0086)	0.0085	(0.0639)
Election Day × Min. to Early PS: 15 - 17.5	0.0553***	(0.0101)	0.0224	(0.0699)
Election Day × Min. to Early PS: 17.5 - 20	0.0612***	(0.0120)	0.0068	(0.0906)
Election Day × Min. to Early PS: 20 - 22.5	0.0566***	(0.0143)	0.0391	(0.1089)
Election Day × Min. to Early PS: 22.5 - 25	0.0709***	(0.0194)	0.0180	(0.1350)
Election Day × Min. to Early PS: 25 - 27.5	0.0638***	(0.0180)	−0.0126	(0.2042)
Election Day × Min. to Early PS: 27.5 - 30	0.0379	(0.0287)	−0.1058	(0.2668)
Election Day × Min. to Early PS: Over 30	0.0269*	(0.0151)	−0.1723	(0.1887)
Election Day × Std. Longitude^1	0.0058	(0.0183)	0.0252	(1.1320)
Election Day × Std. Longitude^2	0.0097	(0.0084)	0.0367	(1.5002)
Election Day × Std. Longitude^3	−0.0024	(0.0111)	0.0036	(0.1922)
Election Day × Std. Longitude^4	−0.0043*	(0.0024)	−0.0096	(0.1218)
Election Day × Std. Longitude^5	0.0012	(0.0016)	−0.0007	(0.0102)
Election Day × Std. Latitude^1	0.0335**	(0.0129)	−0.1101	(0.1228)
Election Day × Std. Latitude^2	0.0000	(omitted)	0.0002	(1.7259)
Election Day × Std. Latitude^3	−0.0392***	(0.0087)	0.0994	(0.0627)
Election Day × Std. Latitude^4	0.0016*	(0.0009)	0.0157	(0.1528)
Election Day × Std. Latitude^5	0.0067***	(0.0017)	−0.0131	(0.0121)
Election Day × Std. Longitude^1×Std. Latitude^1	−0.0054	(0.0140)	0.0454	(0.0742)
Election Day × Std. Longitude^2×Std. Latitude^1	−0.0083	(0.0167)	0.0159	(0.1024)
Election Day × Std. Longitude^3×Std. Latitude^1	−0.0013	(0.0037)	−0.0280	(0.0226)
Election Day × Std. Longitude^4×Std. Latitude^1	−0.0001	(0.0033)	0.0106	(0.0232)
Election Day × Std. Longitude^1×Std. Latitude^2	0.0331	(0.0210)	−0.0947	(0.2236)
Election Day × Std. Longitude^2×Std. Latitude^2	−0.0010	(0.0070)	−0.0171	(0.2722)
Election Day × Std. Longitude^3×Std. Latitude^2	0.0002	(0.0047)	0.0078	(0.0296)
Election Day × Std. Longitude^1×Std. Latitude^3	0.0035	(0.0048)	−0.0067	(0.0283)
Election Day × Std. Longitude^2×Std. Latitude^3	0.0023	(0.0052)	−0.0068	(0.0291)
Election Day × Std. Longitude^1×Std. Latitude^4	−0.0106**	(0.0043)	0.0194	(0.0226)
Election Day × Dist. to Imperial Palace^1	0.0000	(omitted)	−0.0458	(1.4761)
Election Day × Dist. to Aiming Point^1	0.0042	(0.0094)	0.0744	(0.0642)
Election Day × Dist. to Imperial Palace^2	−0.0049	(0.0041)	0.0000	(omitted)
Election Day × Dist. to Aiming Point^2	0.0098*	(0.0053)	0.0002	(0.0331)
Election Day × Dist. to Imperial Palace^3	0.0030	(0.0032)	−0.0101	(0.1331)
Election Day × Dist. to Aiming Point^3	0.0021	(0.0024)	−0.0093	(0.0178)
Election Day × Dist. to Imperial Palace^4	0.0040*	(0.0020)	0.0000	(omitted)
Election Day × Dist. to Aiming Point^4	−0.0009	(0.0019)	0.0080	(0.0129)
Election Day × Dist. to Imperial Palace^5	0.0036**	(0.0015)	−0.0135	(0.0122)
Election Day × Dist. to Aiming Point^5	−0.0007	(0.0020)	0.0222**	(0.0111)
Election Day × logged Prewar Pop. Density	−0.0091**	(0.0039)	0.0328*	(0.0176)
Election Day × Ratio of Residential Area	−0.0157	(0.0109)	0.0426	(0.0623)
Constant	−0.0283	(0.0187)	5.7208***	(0.1498)

Continued on next page

Table A.7 – continued from previous page

	b	s.e.	b	s.e.
Election-Day Polling Station Fixed Effects	✓		✓	
Observations	120,495		120,495	
Within Polling Station R^2	0.0573		0.0281	

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Linear probability model in Model (3) and OLS in Model (4). Population weights are calculated from the neighborhood level population counts divided by the number of users in the sample. Standard errors in parentheses are clustered by neighborhoods in Model 3 and by prewar 35 wards in Model 4. Models exclude the users whose cost distance is from 0 to 2.5 minutes. Difference-in-differences estimation is performed between the election day (July 21, 2019) and the reference day (July 28, 2019.) See Table 3 for the description of covariates. “Min.”, “PS”, “Std.”, “Dist.” and “Pop.” represent minutes, polling stations, standardized, distance and population, respectively.

Appendix B Identification Assumptions Through Interaction DAG

This appendix discusses some of the assumptions underlying the estimation model used in Application 2 in the main text. To illustrate causal relationships in this and other similar studies, the Interaction Directed Acyclic Graph (IDAG) is instructive (Nilsson et al., 2020). IDAG is an extension of Directed Acyclic Graphs (DAG) (Pearl, 1995), graphical models for illustrating causal relationships. By using the causal effect of an explanatory variable on the outcome variable instead of the outcome variable as a node, IDAG allows for the explicit expression of the presence or absence of interaction terms, which was not possible with DAG.

With regard to Application 2 in the main text, since whether a given day is a reference day or an election day (I^{Elec}) has an effect on the voting decision ($Vote$), we denote this effect by the node $\Delta Vote_{Elec}$. The impact of the raid damages ($Damage$) on this node is the quantity of interest. With prewar/geographical control variables (X), the IDAG for our study can be illustrated as follows.[1]

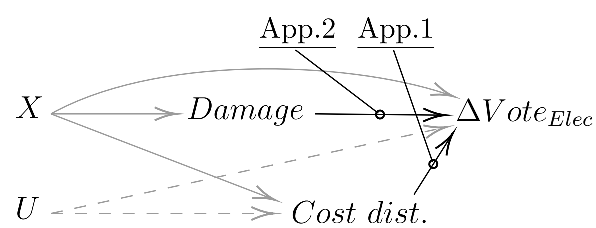


Figure B.1: Interaction DAG (IDAG) for two applications in this study

Note: Black solid arrows represent the causal relationships of interest. Gray solid arrows, gray dashed arrows, and gray dotted arrows represent the relationships between observed variables, those involving unobserved variables, and those that researchers assume to be non-existent, respectively.

In this IDAG, subscripts are omitted for simplicity. To make the necessary assumptions clear, we also added unobserved variables (U) to the IDAG. We assume that the unobserved variables may affect the decision to vote as well as the residential choice related to the cost of voting. Therefore, in the first application, we performed the sensitivity analysis to gauge the impact of unobserved confounding. For the second application, we assume that the unobserved variables do not affect the damage ratio, based on the extensive discussion of this question in Harada et al. (2021). Moreover, we do not assume that the damage ratio affects the cost of voting. Indeed, we later show the validity of this assumption in the main text, using the cost distance variable as a negative control in the robustness check for the second application.

[1]Since IDAG does not allow for the inclusion of fixed effects, this figure should be understood as representing the relationship between the demeaned variables with respect to polling station.

Appendix C Checklist for Using Cell-phone Mobility Data

This appendix offers some guidelines for researchers who use cell-phone mobility data to extract information beyond the presence or absence of users. Note that cell-phone mobility data is often provided in an aggregated form by location, but for the sake of simplicity, we will assume that the data are provided in a track-record format (as illustrated in the example in Table A.1).

1. Find data suitable for processing

First, the data used for applying our approach should ideally satisfy the following conditions:

1. A polling place for voters is assigned and voters have the ability to learn its location.
2. A correctly measured outcome variable with a sufficient sample size is available in a different form to validate the accuracy of the constructed data.
3. Some information (if appropriate to the case) is available to control for the influence of mail-in ballots and absentee ballots.

For reference, in Japan's national elections: condition 1 is satisfied because a voter can only vote on election day at the polling place designated by the resident's address, because voting takes place after matching the voter's identity with the electoral register. We have shown that addresses can be approximated from the geolocation of the first signal of the day. With regard to condition 2, polling station-level turnout data, organized by areas of about 3-5 neighborhoods, are available for most municipalities. For condition 3, mail-in ballots are rarely used because they are limited to the disabled, the war-injured, and those in need of nursing care, and the procedures for accessing this option are complicated.

Finally, an ideal data source would include identification numbers for users so that these could be linked across days in a panel data set for the purpose of including individual fixed effects in analyses. As far as we know, linking IDs across days is not possible in Japan, as it can easily identify individuals, so this approach is not illustrated in our study applications in the main text.

2. Set reference days

The next step is to establish a reference day and obtain the same data for the reference day as for the voting day. The ideal reference day is one that has the same characteristics as the voting day, except for the absence of voting. It is not possible to verify how well the chosen reference day meets this condition, but the reference day would need to be at least the same day of the week as the polling day.

If factors such as poor weather are present on the polling day that discourage people from going out to vote (e.g., Kitamura and Matsubayashi, 2022), a reference day should be chosen where the same conditions apply. Even if the weather conditions match, data from the previous or following year should be avoided because the composition of the population generating the data has changed. It is also generally preferable to set the reference day after the election, since overlap with the advance poll (early voting) should be avoided. In addition, it is important to check for other elections or large events that take place only on that day. In terms of data quality, days with active solar flares should be avoided because of reduced GPS accuracy.

If there is no reference day that satisfies all of these conditions, it is desirable to select multiple reference days to satisfy all of the conditions in a mutually complementary manner.

3. Construct pooled cross-section (or ideally panel) data

There are two advantages to setting a reference day. First, the data from the reference day can be used as the placebo outcome variable. Our study also took advantage of the fact that the reference day is theoretically uncorrelated with official election statistics—because no one goes to the polls on the reference day—and used it as a comparison for performance in inferring residence and voting status from track records.

Second, the theoretically uncorrelated nature of the reference day outcome variable allows us to construct a pooled cross-section data set, so that we can include fixed effects for any given region in our analysis. Because fixed effects generally control for things that are invariant within the observation period, the shorter the observation period, the more confounding factors can be controlled for because more things are invariant in that period. Cell-phone mobility data allows the researcher to take advantage of knowledge of the date of the event and divide the outcome variable into election day and other days on a daily basis. This makes it possible to construct pooled cross-section data in a very short period.

If individual IDs are available across days, one could construct panel data that allows for analysis with individual fixed effects. While the capability to use knowledge of exogenous events to clearly define data pre- and post-event is an advantage of natural experiments (Titiunik, 2021), we can see similar advantages in the analysis of cell-phone mobility data. Since the results of this method can vary greatly depending on which day is used as the reference day (Nunn and Qian, 2011), ideally, multiple reference days should be used to obtain consistent results.

4. Tune parameters and validate data

Parameters should be tuned during the data creation process so that the generated outcome variable reflects the behavior one wishes to measure with fewer errors. In general, anything that requires the researcher to determine values a priori during data processing is subject to tuning.

The purpose of parameter tuning includes not only the reduction of measurement error but also the elimination of arbitrariness in parameter selection, and the two objectives are not always compatible. For example, one method is to manually measure the radius of the polling place for each building, while another is to use a uniform radius of one hundred meters. The former reduces measurement error, but the latter is less arbitrary. This example also reveals another issue of tuning, which is that the impact of the choice of processing method, such as whether to set a radius for each building or a uniform radius, is far greater than the numerical choice, such as whether the radius should be 50 meters or 100 meters.

5. Validate data

In order to evaluate the tuning performance, the administrative data for the same outcome variable you wish to create should be aggregated at a level of granularity that ensures statistical power. For example, if voters in Tokyo's 23 wards were used as the population, the number of observations for validation would be 23, which is insufficient to test the tuning performance if the voter turnout was only available on a ward-by-ward basis. The result of validation is the result of tuning when the parameters are optimized. Therefore, tuning and validation are an inseparable pair of processes.

As shown in this study, it is important to use scatter plots as well as correlation coefficients for performance validation because scatter plots may reveal data-specific patterns that lead to the identification of bias. Obviously, the correlation coefficients between the created data and the administrative data on the election day should be as close to 1 as possible, and on the reference day, the closer those correlation coefficients are to 0, the better.

6. Estimate the regression model using a negative control

While tuning and validation can reveal relatively good parameter settings, they do not provide criteria for judging the accuracy of the variables created. This means that validation alone is not sufficient to ensure data accuracy. Robustness tests compensate for this deficiency. Two robustness checks are especially important: negative control and sensitivity analysis. A negative control is a kind of placebo outcome variable that is affected by the same covariates as the original outcome variable but is not supposed to be affected by the treatment variable (Arnold and Ercumen, 2016). In our study, unobserved confounders are likely to arise primarily in the process of creating outcome variables from track records. Therefore, the distance to the polling station, which was determined based on the estimated address from the track record, just like the outcome variable, was used as the negative control.

7. Perform sensitivity analysis

The second important robustness check is to conduct a sensitivity analysis. Sensitivity analysis is a systematic representation of how much the estimated coefficient of a treatment variable changes when the assumptions made in the analysis are relaxed in various ways. The type of sensitivity analysis to be performed should be selected based on the researcher's knowledge of the data generating process. However, when outcome variables are created by treating proximity to a polling station on the election day as a vote, the errors in determining the votes of those living near the polling station were found to be quite large. Therefore, it would always be necessary to check the extent to which the results might change if these voters were excluded from the sample.

Appendix D Limitation on the Release of Data

We are permitted to provide the analysis dataset used to produce the regression results in the manuscript, but are not permitted to upload/share the proprietary individual-level raw data from Agoop. The raw data are licensed for use only for this study and not for reuse by third parties. Other datasets, such as the measures of damages in the Tokyo firebombing, will be made available upon publication.

References

Arnold, Benjamin F. and Ayse Ercumen (2016). Negative control outcomes: A tool to detect bias in randomized trials. *JAMA 316*(24), 2597–2598.

Harada, Masataka , Gaku Ito, and Daniel M. Smith (2021). Destruction from above: Long-term legacies of the Tokyo air raids. http://dx.doi.org/10.2139/ssrn.3471361.

Kitamura, Shuhei and Tetsuya Matsubayashi (2022). Now or later?: The inter-temporal decision-making of electoral participation. *Political Behavior online first*.

Nilsson, Anton , Carl Bonander, Ulf Strömberg, and Jonas Björk (2020). A directed acyclic graph for interactions. *International Journal of Epidemiology 50*(2), 613–619.

Nunn, Nathan and Nancy Qian (2011). The potato's contribution to population and urbanization: Evidence from a historical experiment. *The Quarterly Journal of Economics 126*(2), 593–650.

Pearl, Judea (1995). Causal diagrams for empirical research. *Biometrika 82*(4), 669–688.

Titiunik, Rocío (2021). Natural experiments. In J. N. Druckman and D. P. Green (Eds.), *Advances in Experimental Political Science*, pp. 103–129. Cambridge University Press.