YOU'RE FIRED, MY LORD: EXPERIMENTAL EVIDENCE ON DEMOCRATIC NARRATIVES AND ACCOUNTABILITY*

Naila Shofia† & Piero Stanig‡

March 5, 2026

Abstract

We study how lay conceptions of democracy shape voter preferences and behavior using a large, multi-wave survey experiment conducted during Indonesia's 2024 general election. Using random utility models and a discrete choice experiment, we show that voters with different conceptions of democracy place systematically different weights on candidate characteristics. To provide causal evidence, we randomly assign an animated video treatment framing elections as instruments of accountability, where politicians are "employees" of the voter. The intervention significantly increases the weight voters assign to track record and policy alignment in candidate evaluation. Crucially, the treatment also affects voting behavior in the real-stakes presidential election by reducing last-minute switching to the poll-predicted winner. These effects are driven primarily by voters with limited prior understanding of electoral accountability. Extensive analyses support the interpretation that the intervention changes the lay understanding of democratic representation among those who did not initially think in terms of political agency. The results suggest that narrative-focused interventions aimed at "soft features" of democracy may offer a promising path to enhance democratic accountability.

JEL Codes: D72, D91, D83, C99, O43

*We thank Rema Hanna, Judd Kessler, Jamus Lim, Benjamin Marx, Steve Monroe, Rohini Pande, Vincent Pons, Pavithra Suryanarayan, and David Yang for insightful comments and constructive feedback. We are also grateful to seminar and workshop participants at Harvard, Boston University, the National University of Singapore, SIOE (University of Chicago), Monash Indonesia, GRIPS Japan, APSA, ESSEC Singapore, SMU, the University of Barcelona, and SEAREG 2025 for valuable suggestions. We thank Ayesha Syed, Katya Loviana, and Wan Hasjim Omar Jacob Saleh for excellent research assistance. All remaining errors are our own. Shofia acknowledges financial support from the Yale–NUS College Internal Grant (IG23-SG015). This project was pre-registered in the AEA RCT Registry (AEARCTR-0012872) and was approved by the Institutional Review Board at the National University of Singapore (Protocol: NUS-IRB-2023-971) and the University of Indonesia (Protocol: 038/UN2.F6.D2.LPM/PPM.KEP/2023).

†Department of Economics, National University of Singapore; Visiting scholar, Center for International Development, Harvard Kennedy School; formerly Yale-NUS College. Email: naila.shofia@nus.edu.sg

‡Department of Social and Political Science, Bocconi University. Email: piero.stanig@unibocconi.it

"*A republic,* ***if you can keep it****.*"

—Benjamin Franklin, in response to Elizabeth Powel's question, "Well, Doctor, what have we got, a republic or a monarchy?" (The Constitutional Convention, September 17^{th}, 1787)

Franklin's historical remark above underscores the conditional and fragile nature of democratic governance. It highlights the central role of citizens in remaining informed, engaged, and continuously vigilant against the misuse of power by governments—even when those governments are democratically elected. Crucially, for roughly half of the world's countries currently classified as liberal or electoral democracies (Nord *et al.*, 2025; Economist Intelligence Unit, 2025), the mere existence of democratic institutions is *not sufficient*; citizens ought to ensure that democracy works as intended. In this sense, citizens are the ultimate guardians of democratic institutions—institutions designed to enable self-governance. A necessary precondition for fulfilling this role is an accurate lay understanding of how democracy is meant to function, a concept we refer to here as *democratic narratives.*

Narratives shape voter expectations and behavior. Voters who perceive politicians as agents representing their interests—what we term an **agency** narrative—are more likely to reward competence and punish failure. By contrast, voters who view politicians as powerful patrons who dispense favors or protection—what we refer to as a **deferential** narrative—tend to see elected officials as rulers *over*, rather than representatives *of*, their constituents, and therefore as deserving of loyalty or gratitude rather than scrutiny.[1] This latter belief system reorients the purpose of voting away from selecting and disciplining agents toward expressing allegiance or gratitude. As a result, voters who adopt a deferential narrative may place less weight on performance and instead prioritize attributes such as popularity, power, or group identity—undermining democratic accountability and ultimately reducing voter welfare (Besley, 2006; Ferraz and Finan, 2025). This mechanism helps explain why, even under free and fair elections, democratic systems may fail to translate electoral competition into genuine improvements in public welfare.

In this paper, we use Indonesia—the world's third-largest democracy—as a testing ground to study the role of democratic narratives. We provide causal evidence that democratic narratives shape voting behavior by influencing the criteria voters use to evaluate political candidates, and we show that these narratives can be shifted through a targeted intervention. This shift, in turn, alters voters' preferences and behavior, including in a real-stakes election.

We administer a large, three-wave panel survey experiment—two waves before Indonesia's

[1]The terms *agency* vs. *deferential* narratives are conceptually close to the distinction between *nonpermissive* and *permissive* norms in Bidner and Francois (2013). In their framework, *nonpermissive* norms are held by voters who attach high disutility to transgressive acts such as corruption, whereas *permissive* norms are held by voters who regard such deviant acts as acceptable or routine features of political life.

2024 general election and a final wave after the election. Our study draws on a survey of more than 8,700 eligible voters at baseline and constitutes a nationally representative sample of voting-age Indonesians. The research design employs a discrete choice experiment to elicit voters preferences, where respondents evaluate fifteen pairs of hypothetical—but realistic—political candidates whose attributes are fully randomized, allowing clean identification, in a random-utility model of vote choice, of the weights that voters place on each attribute.

Prior to this task, we randomize exposure to a 75-second animated video that explains political agency using a lay metaphor, delivered in the second survey wave, 4–20 days before election day. Across waves, we measure baseline political views and positions, pre-election vote intentions, and post-election self-reported vote choice. Pre-treatment, we deploy a simple method, based on analogies, to measure how respondents conceptualize electoral democracy, which allows us to classify them as Deferential or Agency voters depending on how they score.[2]

We find that the weights voters place on candidate traits—i.e., their decision criteria—differ markedly between Agency and Deferential individuals. Agency respondents assign significantly greater weight to candidate valence and ideological proximity —roughly equivalent to increases in vote margins of about 4 and 5 percentage points, respectively— than Deferential respondents do. By contrast, the weight placed on identity markers does not differ statistically across the two types. Taken together, these patterns suggest that democratic narratives are tightly linked to the criteria voters apply when evaluating candidates.

Our randomized intervention induces a substantive, statistically significant reweighting of the evaluation of candidate characteristics. The treatment allows us to identify the causal effect of citizens' beliefs about electoral democracy on decision criteria, i.e., on the weights in the utility functions of voters. Treated individuals place greater weight on valence and ideological proximity, with effects concentrated among Deferential respondents who, pre-treatment, did not conceive of representation as an agency relationship. Among Deferential respondents, the treatment raises the weight on valence by roughly 19%, while no corresponding effect is detected among Agency respondents. The treatment also increases the weight on ideological congruence, particularly among the Deferential. By contrast, we find no treatment effect on the nominal weight placed on identity congruence, including within the Deferential group. The results suggest that Indonesian voter preferences are malleable in a particular way: Reframing elections as accountability mechanisms increases the weight

[2]Conducting the experiment during an active election campaign enhances external validity: voters are highly engaged with political choices and continuously exposed to campaign information, enabling us to capture preferences in a context of genuine electoral salience. Importantly, the panel structure further allows us to track the evolution of these preferences—from pre-election attitudes and stated intentions to self-reported vote choices immediately after the election.

voters place on valence (competence) and ideological proximity, but leaves ascriptive attachments largely intact. This suggests that while preferences can change, attachment to ethno-religious identity seems to be persistent and hard to dislodge.

Crucially, these preference shifts translate into observable changes in voting behavior in the actual presidential election. We document systematic differences in vote-switching between Agency and Deferential voters: Agency voters are less prone to the bandwagon effect—i.e., last-minute switching to the poll-predicted winning candidate—and more likely to stick with the choice stated in the pre-election wave, even when the candidate is polling as an underdog. The informational treatment significantly alters Deferential voters' behavior: post-treatment, their switching patterns converge toward those of Agency voters, with statistically significant reductions in bandwagoning and greater persistence of initial choices. This finding complements the evidence on voter preferences, showing that the narrative treatment induces Deferential voters to behave more like Agency voters.

Importantly, we also find—based on a subset of respondents for whom we measured democratic narratives in both the pre-election (pre-treatment) and post-election (post-treatment) waves—that the treatment significantly increased respondents' Agency scores measured in the post-election wave. This provides *direct evidence* that the treatment produces a shift in respondents' narratives and also implies that our brief and low-key intervention produced a persistent effect, at least several days or weeks after its administration. We also provide evidence that experimenter demand effects are not the explanation for the treatment effect we document, by showing that (i) the treatment effect is not homogeneous across Deferential and Agency respondents; (ii) the heterogeneity in treatment effect goes in the opposite direction for respondents with low *factual* political knowledge in the analysis of response time in the candidate-choice exercise; and (iii) unlike the treatment video, the placebo video—compared to a control sub-arm where respondents see no video—has null effect on individual decision criteria.

The effects identified with the main random utility model are corroborated through several analyses. First, we use response time, rather than stated choices, as the outcome. Analysis based on response times further corroborates our findings by showing that the effects remain identifiable even when using a minimally invasive measure. While respondents may exhibit social-desirability bias in stated choices—for instance, avoiding the appearance of favoring candidates from their own religious group—it is difficult to imagine that they strategically manipulate response times to project a more socially acceptable self-image as decision-makers. Second, we use a non-parametric estimator following Hainmueller *et al.* (2014) to re-estimate the weights and treatment effects. Beyond serving as a robustness test, the non-parametric approach estimates preferences more flexibly and thus, permits a

direct assessment of the *additive separability* assumption in the main model: valence and ideological proximity are rewarded equally whether a candidate belongs to the respondent's in-group or out-group. Finally, we conduct a sensitivity analysis to address respondent fatigue by re-estimating the utility weights using five-period rolling windows. Conclusions from these exercises closely align with those from the random utility model.

This paper contributes to the literature on the political economy of democracy. In the spirit of Bidner and Francois (2013), who emphasize the role of political norms in democratic consolidation, we provide the first empirical evidence that a *narrative* intervention is effective in shifting preferences and real-stakes voting behavior, opening a promising path for fostering a norm of electoral accountability. Given that democracy's ability to "deliver" helps make it self-sustaining (Acemoglu *et al.*, 2024), and institutional constraints are highly consequential, reforming political institutions remains essential for development. Shifting mass beliefs about democracy may substitute for institutional reform itself or complement it by enabling institutions to function as intended (Bidner and Francois, 2013). In practice, strengthening democratic institutions is likely best pursued incrementally, with each step informed by evidence on what works and what backfires (Banerjee and Duflo, 2011). Crucially, while randomizing institutional reforms is rarely feasible, belief-based interventions offer a more scalable lever.

This paper also contributes to the vast empirical literature on policies to improve accountability in developing democracies. A substantial body of work examines the electoral effects of candidate-specific information. For instance, performance report cards and the disclosure of audits or criminal records can increase support for high-performing or less corrupt politicians (Banerjee *et al.*, 2011; Ferraz and Finan, 2008; George *et al.*, 2019), particularly when the information is amplified by media exposure (Larreguy *et al.*, 2020). Other studies focus on exposure to candidate debates and campaign promises, showing that such interventions can enhance voter knowledge, increase the salience of policy and performance, and reduce identity-based or partisan voting (Wantchekon, 2003; Bowles and Larreguy, 2020; Kendall *et al.*, 2015; Cruz *et al.*, 2021, 2024). Finally, studies of legal or moral appeals to discourage vote buying find mixed effects (Vicente and Wantchekon, 2008; Vicente, 2014; Hicken *et al.*, 2015). Our study complements this literature in several ways. First, our approach differs from previous work as we focus on the "soft features" of political systems —political norms and attitudes— rather than providing "hard" factual information. This is a promising alternative in contexts where spreading candidate-specific information may be politically sensitive or polarizing, especially around actual elections. Second, our approach also avoids potential unintended effects, such as lower turnout (Chong *et al.*, 2015) or polarization (Baysan, 2022). This is because our intervention is not only neutral—avoiding partisan cues and the possible

perception of "black campaigning"—but also does not rely on candidate collaboration, nor does it alter campaign content. Therefore, it is particularly well-suited for real electoral settings where neutrality and scalability are essential. We rely on normative beliefs and a cognitively coherent rationale for electoral accountability, rather than appealing to legality or moral virtue, strategies that may lack credibility in low-enforcement contexts. Moreover, methodologically, our study randomizes treatment at the individual level instead of the cluster level (e.g., villages, constituencies, precincts, towns, or districts). We intentionally adopt this approach as it makes it possible to identify the micro-foundations of voter decision-making by inferring their utility function from their choices, rather than studying aggregate electoral outcomes.

Finally, this study is related to broader research in economics on the malleability of preferences. A large body of evidence shows that tastes and priorities shift with institutions, experiences, and information. Exposure to state socialism raised support for government intervention (Alesina and Fuchs-Schündeln, 2007), macro experiences altered financial risk-taking (Malmendier and Nagel, 2011), information about inequality and tax incidence moved preferences for redistribution (Kuziemko *et al.*, 2015), classroom curricula increased patience (Alan and Ertac, 2018), and curricular reforms shifted political attitudes (Cantoni *et al.*, 2017), consistent with theories of endogenous preference formation (Becker and Mulligan, 1997). Further studies on identity, norms, and narratives help explain how such change occurs: identity might enter utility directly (Akerlof and Kranton, 2000); social-image and norm-based incentives steer behavior (Bénabou and Tirole, 2006); contagious stories move behavior at scale (Shiller, 2017); and correcting misperceived norms shifts attitudes and actions (Bursztyn *et al.*, 2020). We contribute to this area of research by (i) showing that voter preferences are malleable, even when prompted by a brief yet illuminating narrative; (ii) demonstrating that this shift in narrative leads to downstream behavioral change in a real-stakes election; and (iii) identifying which group of voters is most malleable, thanks to the novel measurement strategy we devise to gauge voter narratives.

In what follows, Section 1 develops a conceptual framework linking democratic narratives to vote choice, and clientelism. Section 2 situates the study in Indonesia's political landscape and briefly reviews the 2024 general election. Section 3 describes the empirical design—survey methodology, measurement of respondents' conceptions of democracy, the randomized informational intervention, and the candidate-choice experiment used to elicit voter preferences. Section 4 presents the random utility model, our main approach for estimating the weights voters assign to candidate traits. Section 5 reports the estimated trait weights under the random utility model, the treatment effects on these weights, and the treatment effects on actual election choices. Section 6 offers a number of additional analyses

to bolster our confidence in the main findings. Section 7 examines the persistence of these effects and assesses experimenter-demand concerns. Section 8 concludes.

1 Lay conceptions of democracy, vote choices, and clientelistic linkages

In this paper, we focus on the ideational side: individuals differ in how they conceptualize the relationship between voters and elected politicians, and these lay conceptions (or "narratives") shape the criteria voters adopt to make choices. In normative theories of democracy, particularly in electoral control models (Barro, 1973; Ferejohn, 1986; Besley, 2006), elections establish a principal-agent relationship: voters, as principals, use the threat of non-reelection to incentivize politicians, their agents, to act in the public interest (Przeworski *et al.*, 1999). Do citizens view elections as instruments of accountability as in the agency model, or as ritualized contests embedded in entrenched social hierarchies, where ascriptive characteristics—such as ethnicity or religion— and a privileged position within patronage networks play a decisive role?

In many contexts—including Indonesia, the setting of our empirical study—voting might mostly reflect and validate existing social hierarchies, or it may resemble a transactional exchange in which political support is traded for private, excludable benefits, such as cash or favors—a logic well-documented in the literature on clientelism and vote buying (Stokes, 2011; Kitschelt and Wilkinson, 2007). Even with regular, free, and fair elections, politicians might prioritize the targeted distribution of state resources over broad-based policy performance and public good provision (Anderson *et al.*, 2015; Keefer and Vlaicu, 2008; Bardhan *et al.*, 2024). This can hinder development (Bardhan and Mookherjee, 2020; Pande, 2011) and entrench weak institutions, creating a self-reinforcing cycle of poor governance and low state capacity (Fergusson *et al.*, 2022), and low legitimacy (Acemoglu *et al.*, 2024).

Clientelism involves a two-sided commitment problem: voters may renege and fail to reciprocate, just as politicians may not deliver promised benefits (Robinson and Verdier, 2013). To function, clientelistic linkages require mechanisms to ensure compliance. Much of the literature, in economics and political science, emphasizes this enforcement side of clientelism—highlighting how brokers monitor voter behavior or deny services to non-supporters (Kitschelt, 2000; Stokes, 2005; Bardhan *et al.*, 2024). However, this perspective risks overlooking how clientelism can also be sustained by normative beliefs and social expectations, even in the absence of direct enforcement. In fact, hierarchical social norms and moral obligations of loyalty can substitute for monitoring: voters may feel a duty to support a patron

perceived as generous or powerful (Kitschelt, 2000; Auyero, 1999). In this view, clientelism is underpinned by a broader understanding of politics as a realm of deference and domination (Auyero, 2001): Deference-minded voters view politicians as powerful patrons whose favors deserve loyalty, not scrutiny, reorienting the purpose of voting—from selecting or disciplining agents to expressing allegiance or gratitude. Such views can be internalized and persistent, shaping how citizens engage with democratic institutions, possibly even without active reinforcement by political actors. [3]

This is where intervening on the ideational aspect becomes critical. Even in contexts where clientelistic linkages and identity cleavages are deeply entrenched, shifts in voters' mental models—whether in the form of new heuristics, evolving social norms, or redefined understandings of the purpose of democratic elections—can have meaningful effects on democratic accountability. Changing how voters think about elections may then be as consequential as institutional reforms. Theoretical work by Bidner and Francois (2013) highlights this potential, introducing the concept of "gripped" citizens—voters who are committed to punishing political transgressions even when it may not be individually rational due to the lack of better alternatives. These norm-enforcing voters help sustain an equilibrium where politicians, motivated by re-election incentives, refrain from misconduct. Crucially, even a minority of voters committed to such norms can reshape expectations and shift the broader political equilibrium toward greater accountability. In sum, when citizens begin to reinterpret the meaning of elections in agency terms, democratic institutions can function more effectively. Understanding, and shifting, these ideational foundations is therefore central to promoting programmatic politics and strengthening democratic governance.

2 Contextual Background

Indonesia is a presidential democracy that transitioned from authoritarian rule in 1998, following the collapse of Suharto's New Order regime after more than three decades in power. The democratic transition ushered in the Reformasi era, marked by sweeping political reforms including direct presidential elections, decentralization, the proliferation of political parties, and a strengthened role for civil society (Aspinall, 2010; Hadiwinata, 2003). Today, Indonesia is the world's third-largest democracy and is widely regarded as one of the most

[3] More generally, one should keep in mind that elections need not create a representation or accountability relationship at all. For instance, the Pope is elected by cardinals, but the vote serves to identify the most spiritually worthy candidate, not to delegate authority in a representative sense; similarly, a vote in a talent competition is a tool for aggregating information, not enforcing accountability. In the extreme, elections may function more like mechanisms for selecting figures with broad discretionary power—comparable to the selection of a feudal lord—rather than empowering public servants bound to citizen interests.

vibrant democracies in the Global South, though persistent institutional weaknesses and political practices continue to challenge democratic consolidation (Mietzner, 2015).

Two central features of Indonesian democratic politics are clientelism and identity-based mobilization. It is a widely accepted view in the literature that clientelistic linkages are deeply entrenched in the country's political life (Berenschot, 2018; Aspinall, 2014). On the supply side, the state's dominance in the economy—particularly through public employment and infrastructure spending—creates fertile ground for personalized exchanges between politicians and voters (Allen, 2015; Berenschot and Mulder, 2019). In many cases, electoral machines operate through informal networks of brokers and community leaders who distribute favors in exchange for political loyalty. Some have argued that these practices reflect a lingering "feudal" logic, in which hierarchical norms of patronage and loyalty shape political expectations and behavior (Komarudin and Pramuji, 2023).

Alongside clientelism, identity politics has taken on increasing prominence in Indonesia's democratic landscape (Pepinsky, 2019; Shofia and Pepinsky, 2019). The country is characterized by both religious and ethnic majorities—Muslim and Javanese, respectively—as well as significant minority groups, including Christians, Hindus, Buddhists, and ethnic Chinese Indonesians (Mietzner, 2020; Buehler, 2016). While the national ideology of Pancasila emphasizes pluralism and unity, in practice, electoral competition often reflects exclusionary tendencies. Discrimination typically flows from dominant groups toward religious and ethnic minorities, with implications for representation and civil rights. For example, Toha *et al.* (2021) document high levels of intolerance toward non-Muslims and Chinese Indonesians, measured through survey-based indicators of exclusionary political attitudes.

Identity politics also manifests in the form of in-group favoritism. Members of minority communities may display strong preferences for co-ethnic or co-religious candidates, sometimes regardless of qualifications or track records. A striking illustration is the 2017 Jakarta gubernatorial election, in which Basuki Tjahaja Purnama (Ahok), an ethnic Chinese Christian, was defeated despite strong approval ratings. Many observers view his loss—and subsequent imprisonment on a blasphemy charge—as a consequence of orchestrated identity-based attacks (Soedirgo, 2018; Fealy, 2016). More broadly, the near-total absence of class-based electoral discourse—due in part to the legacy of anti-communist violence and repression following the 1965–66 mass killings—means that identity categories frequently become the primary axis of political alignment (Robinson, 2018; Aspinall, 2010). Nonetheless, some evidence shows that elected officials skew in favor of wealthier voters (Warburton *et al.*, 2021), and that Islamist cues can convey economic positions in contexts where programmatic platforms are underdeveloped (Pepinsky *et al.*, 2012).

2.1 2024 General Election in Indonesia

Indonesia's 2024 general elections were held concurrently across all tiers of government—the presidency; the national legislature (DPR); the Regional Representative Council (DPD); and the provincial- and district-level DPRD. Approximately 204 million registered voters were eligible to cast ballots at more than 800,000 polling stations nationwide on the same day, making it widely described as the world's largest single-day election (CNN, 2024). The presidential race featured three candidates, marking a break from the two-candidate contests of the previous two cycles. Observers widely regarded the vote as a referendum on the outgoing administration's legacy and, more broadly, on Indonesia's democratic trajectory (Mietzner, 2024).

The campaign period was marked by intense political competition, widespread mobilization, and renewed concerns about the roles of patronage and identity-based appeals in shaping voter preferences. Despite formal improvements in electoral administration and transparency, observers noted the continued prevalence of vote buying, the mobilization of state resources for electoral gain, and appeals to ethnic and religious identity in both overt and coded forms (Buehler, 2023). Hence, the 2024 elections underscore the ongoing relevance of clientelism and identity politics in Indonesia's democratic practice, even as institutional reforms and civic engagement continue to evolve.

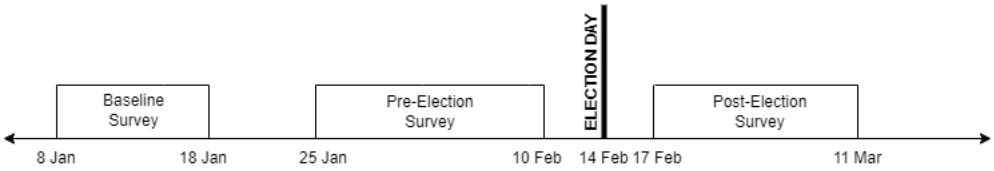
3 The Survey Experimental Design

We conduct a large, three-wave panel survey of Indonesian voters around the February 2024 Indonesian general election. The sample is nationally representative of eligible voters in various key dimensions, i.e., geographic location, gender, age group, rural/urban, and education.

Figure 1 depicts the timing of the survey waves. Wave 1 (baseline) was fielded 4–5 weeks before the election. Wave 2, conducted 1–3 weeks before the election, delivered the informational video treatment (or the control intervention). Wave 3 was fielded 3 days to 2 weeks after the election and collected reports of actual voting, along with the post-election candidate-choice task for a holdout group.

In the pre-analysis plan, we targeted approximately 3,600 respondents by the end of the third survey wave. Since we aimed to collect panel data, where we trace the respondents from the baseline to the last wave, and anticipating attrition across survey waves, we invited a substantially larger pool of respondents at baseline. In practice, about 8,700 respondents completed the baseline survey, of whom roughly 5,400 also participated in the pre-election

Figure 1: Survey Timeline



survey; approximately 4,800 respondents subsequently completed the post-election survey. In Appendix Table A.1, we report the baseline characteristics of respondents: there are no discernible differences in the respondents' characteristics across waves.

Following the pre-analysis plan, two-thirds of respondents in the pre-election wave were randomly assigned to complete the candidate-choice task prior to the election, while the remaining one-third were assigned to a different module and scheduled to complete the candidate-choice task in the post-election survey (wave 3). As a result, approximately 3,700 respondents were allocated to the pre-election candidate-choice exercise, while about 1,300 respondents were allocated to this module in the post-election survey.[4]

After matching respondents across waves and excluding incomplete responses, the final sample for the candidate-choice analysis consists of approximately 3,200 respondents who completed the pre-election candidate-choice task, about 650 of whom also completed the post-election candidate-choice task.[5] All respondents in this sample can be traced back to the baseline survey, successfully completed the pre-treatment *drag & drop* task used to elicit democratic narratives, received one of the experimental treatments, and completed the candidate-choice module as recorded on the Qualtrics platform. Appendix Table A.2 shows the distribution of candidate-choice data for these respondents.

The final sample for the analysis of vote choice in the actual presidential election consists of approximately 2,600 respondents (about half of whom are classified as Deferential) who can

[4]In the third wave, after reaching the target sample size for the candidate-choice exercise, additional respondents were reassigned to the alternative module, such that fewer than one-third of third-wave respondents completed the candidate-choice task.

[5]In practice, there was a deviation from the pre-registered protocol in that a subset of respondents completed the candidate-choice task in both the pre- and post-election waves. As a result, the candidate-choice sample does not include a pure "holdout" group. Crucially, this deviation does not threaten the validity of the design, since the treatment was administered only in the pre-election wave and was fully randomized. In fact, in retrospect, this overlap turns out to be a blessing in disguise, as it allows us to estimate both the direct effect of the treatment on narratives and the persistence of treatment effects over time. We present this analysis in Section 7. Moreover, consistent with ethical guidelines, respondents were explicitly informed that participation was voluntary and that they could stop the survey at any time or skip any question without penalty. This resulted in substantial attrition before respondents reached the final sample. In addition, as some respondents completed the candidate-choice task in both waves, they contributed more than 15 choice pairs to the analysis. We accordingly cluster standard errors at the respondent level throughout.

be linked to the baseline survey, successfully completed the pre-treatment *drag & drop* task used to measure lay narratives, received one of the experimental treatments, and reported both their pre-election vote intention and their post-election vote choice. Crucially, as shown in Appendix Table A.3, there is no evidence of selective attrition into the post-election survey conditional on participation in the pre-election wave. Similarly, inclusion in the vote-analysis sample is not predicted by treatment assignment, deferential beliefs, baseline vote intention, or baseline demographic characteristics. Finally, we estimate raking weights separately by survey wave to rebalance each sample to the most recent census marginals, following the procedure described in Appendix Section F.

The survey was distributed online in collaboration with a local marketing polling company which maintains a panel of respondents with an excellent reach of the population across islands in Indonesia. We collected the responses via a Qualtrics questionnaire designed and managed by our research team; the survey company only played the role of connecting their panel respondents to our survey links. We instructed the company regarding quotas by province of residence, gender, age group, rural-urban, and education conditional on age group (over or under-40), based on the 2010 census and the 2024 voter distribution published by the Indonesian Electoral Commission. The full questionnaire is available at https://doi.org/10.7910/DVN/2ESHR5. Here we briefly discuss the general structure of the questionnaire and a small subset of variables, the definitions of which are important in the following.

3.1 Questionnaire Structure

Baseline. After introducing the survey and obtaining informed consent, the baseline questionnaire collected respondents' demographic characteristics, including education, ethnicity, religion, and measures of religiosity. The survey then included modules measuring, among others, media consumption, factual political knowledge, political worldview (Islamism, civic-nationalist Pancasila, or a compromise between the two), and positions on five key ideological issues (e.g., traditional morality and the role of the military in politics). Information from this last module is used to construct measures of ideological distance between respondents and political candidates.

Candidate-choice module. This module is administered both in the pre-election and the post-election waves. After introducing the survey and obtaining informed consent, respondents were first asked to identify their electoral district.[6] In the pre-election wave, the questionnaire also elicited, pre-treatment, respondents' electoral stances, including their in-

[6]This information is then used to define which province-specific candidates' ethnic distribution will be shown in the candidate-choice exercises below.

tention to turn out and their intended presidential candidate. In the post-election survey, this section was replaced with questions on realized electoral behavior, including whether respondents voted in the 2024 general election and their presidential vote choice.

Respondents then completed the *drag & drop* task used to measure the lay narrative held by respondents. Random assignment to treatment in the pre-election wave was triggered when respondents clicked "Next" after completing this task. Respondents were assigned to the treatment video with probability one-half, and to either the control video or the no-video condition with probability one-quarter each. Treatment assignment was therefore implemented "live" during the survey using the Qualtrics interface. As a result, treatment assignment is balanced by design. The balance in the realized sample can be observed from the descriptives reported in Appendix Table A.4. Treatments were administered only in the pre-election wave.

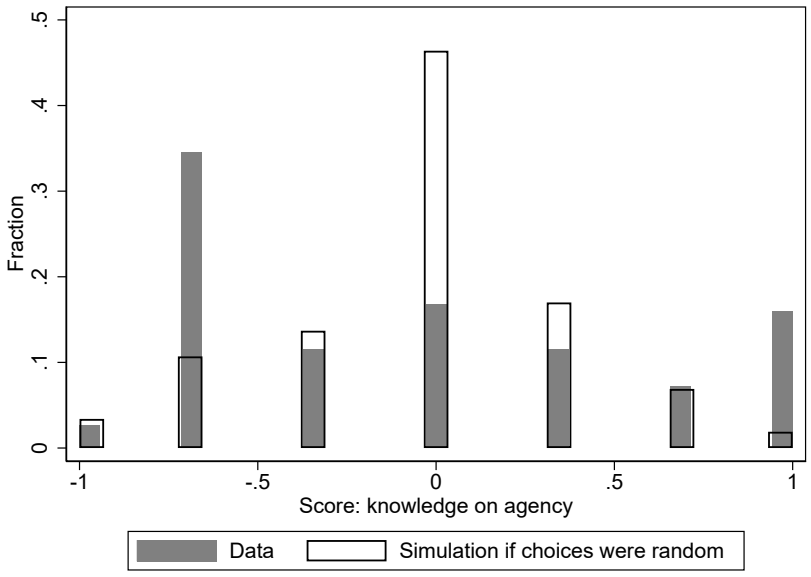
The subsequent section consisted of the candidate-choice exercise, in which respondents evaluated fifteen pairs of hypothetical candidates. Across all survey waves, respondents were informed that participation was voluntary and that they could stop at any time or skip any questions without penalty. Multiple attention checks were embedded throughout the survey to ensure response quality.

3.2 Measuring Voter Narrative: Agency vs. Deferential

We distinguish between Agency individuals, who view democratic elections as establishing a principal–agent relationship in which voters hold politicians accountable, and Deferential individuals, who perceive elections as a form of "electoral feudalism," whereby voting legitimizes hierarchical relationships in which elected officials rule over, rather than represent, their constituents. We measure adherence to these competing views using a *drag & drop* categorization game in which respondents sort objects into boxes according to the category they consider most appropriate. The task includes a practice question followed by two substantive questions; in each, respondents assign one object to each box. Illustrations and full instructions are provided in Appendix Section B. All the scoring rules were pre-specified in detail in the pre-analysis plan, available at https://www.socialscienceregistry.org/trials/12872.

We then classify respondents as Deferential or Agency-minded based on a median split of the scores, which is realized at zero, and categorize respondents who scored < 0 as Deferential voters, and those who scored ≥ 0 as Agency voters. Figure 2 displays the distribution of the scores, compared to those that would have been obtained under the same scoring rules if respondents chose each of their answers randomly. One can notice that both very high and

Figure 2: Distribution of Scores on Agency Knowledge



Note: The graph presents a histogram comparison between the distribution of actual scores in our data and the simulated data if the respondents were responding randomly in the *drag & drop* categorization game.

very low scores are significantly more frequent than random, hence they plausibly reflect a deliberate choice.[7]

[7]One could be concerned that those who give a deferential response in the *drag & drop* exercise, rather than being ill-informed about how the roles of voters and representatives are conceived in standard democratic theory, are simply manifesting cynicism or disillusionment. Later in the survey, we also ask respondents whether (and why) voters can or cannot punish politicians with their vote. We can classify as *normatively* hierarchical those respondents who respond "No, because society must obey political leaders" and "No, because political leaders have higher power than society." Respondents who hold these hierarchical normative views score more than one quarter of a point lower on average than respondents who do not hold these views. Similarly, we can classify as tolerating the self-interest of politicians those respondents who *normatively* think that politicians can pursue their own self-interest. These are few (around 2% of the sample) but score around half of a point lower on the metric. As these questions were asked after the video treatment, we restrict the analysis here to respondents in the control arm. In addition, if cynicism were driving the responses, we should expect more educated voters, with higher general cognitive sophistication, to engage in cynicism more. Yet there is a clear monotone educational gradient, where respondents with university and graduate degrees score highest on average, and respondents with elementary and middle school score lowest. Finally, even conditional on education and other demographics, respondents with better factual knowledge about the political system score higher. Table A.5 in the Appendix reports regression results documenting these patterns. Following the interpretation we adopted in the pre-registration documents, we treat the score based on these survey items as a measure of agency vs. deferential narratives.

3.3 The Treatment

The main experimental intervention is an informational video explaining the agency view of electoral democracy. In the second survey wave, respondents were randomly assigned with equal probability to either the treatment or a control arm. The treatment video uses animation to illustrate electoral accountability by drawing an analogy between a shop owner who hires—and can fire—a misbehaving store manager and voters who elect—and can remove—politicians who misuse their authority.

Within the control arm, half of respondents viewed a placebo video of identical length and visual style that presented factual information about the electoral process (e.g., voter eligibility, registration procedures, and election administration), while the remaining half received no video. Full video scripts (in Bahasa Indonesia and English translation) are provided in Appendix D, and the videos are available at https://doi.org/10.7910/DVN/2ESHR5.

Our experimental manipulation aims to exogenously alter the normative beliefs of voters about democratic elections, allowing us to estimate, at the individual level, its effect on voter decisions. Often, studies in this area adopt designs where randomization and treatment administration take place at the cluster (village, constituency, precinct, town, or district) level (Bidwell *et al.*, 2020; Bowles and Larreguy, 2020; Cruz *et al.*, 2021, 2024; Fujiwara and Wantchekon, 2013; Kendall *et al.*, 2015; López-Moctezuma *et al.*, 2022; Platas and Raffler, 2021; Vicente, 2014; Wantchekon, 2003; de Albuquerque *et al.*, 2025). While arguably high in terms of "naturalism," such aggregate interventions hinder the study of the micro-foundations of electoral decision-making. In particular, if a given location (e.g., village) receives a specific type of campaign message (e.g., promises of broad provision of public goods) as a cluster-assigned treatment, the overall dynamics within that location might be affected: informal political discussion between citizens or responses of local political elites (unobserved to the researcher) create feedback effects that make it harder to pinpoint the mechanism, and even individual-level outcome data can only provide insights that are inclusive of any context-level feedback effect.

3.4 Preference Elicitation Strategy

Conceptually, our measure of interest is the importance that voters assign to various traits of candidates when making electoral choices. Rather than directly asking people what features they value in a politician, something they might not be correctly cognizant of, or might want to misrepresent for social desirability reasons, we estimate the importance voters assign to different candidate features via a *discrete choice experiment* in which respondents pick their

favorite among pairs of candidate profiles with randomized traits.

Survey-based preference elicitation with hypothetical scenarios and randomized attributes is widespread and increasingly common in economics and the social sciences. Many studies in economics present *paired* scenarios in each round and ask respondents to choose one option, mirroring our design, e.g., Benjamin *et al.* (2014, 2012); Wiswall and Zafar (2018); Maestas *et al.* (2023). Crucially, Lewandowski *et al.* (2026) demonstrate that this strategy strongly predicts actual counterpart behavior, even three years after the initial elicitation. A common variant instead shows a *single* hypothetical profile per round and elicits the likelihood of a specific action (e.g., hire, approve a loan, donate an amount), as in Fong and Luttmer (2009, 2011); Kessler *et al.* (2019); Low (2024); Macchi (2023); Colonnelli *et al.* (2025). Related approaches are widely used in political science, e.g., Bansak *et al.* (2022); Hainmueller *et al.* (2015); Svolik (2019); Hainmueller *et al.* (2014); in psychology, e.g., Caruso *et al.* (2009); Mogilski *et al.* (2014); Wade and Mogilski (2018); Anik and Hauser (2020); Neuman and Neuman (2008); Mattmann *et al.* (2019), and have a longer history in marketing research (Green and Rao, 1971). A key advantage of these designs is their ability to *recover causal, unbiased preferences* over multidimensional traits—thanks to the orthogonality induced by randomized attributes—within a compact survey format. With multiple rounds of choice tasks, this method also enables *within-respondent identification* of how different traits affect the probability of being chosen (i.e., voted for).[8]

In practice, the data collection in our study involves showing respondents fifteen pairs of hypothetical candidates for the national legislature; a sample of the candidate pairs in the choice exercise is available in Figure A.1 of the Appendix. The candidate traits are fully randomized, and hence orthogonal to each other. Candidates differ on the following characteristics: gender, age, religion, ethnicity, education, marital status, occupation, position on a contentious issue, and valence. The characteristics are presented in this order to all respondents. Importantly, we do not mention the party affiliation of the candidates. These profiles contain what one can consider "high-quality" information regarding track record, stances, and ethno-religious membership. Crucially, the information about candidates in actual legislative races in Indonesia might not be of such high quality, and might not be organized in the tidy format that we show to our respondents. Yet, our evidence shows that voters, in the presence of high-quality information, actually use it quite competently: hence it should also be in the interest of candidates and parties to spread better information.

Candidate demographic traits are drawn from distributions designed to approximate

[8] Across disciplines, the term "conjoint experiment" is often used for this class of methods, though Louviere *et al.* (2010) argues that conjoint analysis and discrete choice experiments are not identical; nevertheless, they conclude that most studies using randomized hypothetical scenarios are essentially estimated as discrete choice experiments.

those of candidates who ran in the 2019 general election, as well as province-specific ethnic compositions based on the national census. This design enhances the realism of the candidate-choice exercise (Hainmueller *et al.*, 2015). Appendix Table A.6 reports all candidate attributes, presenting both the benchmark distributions specified in the pre-analysis plan and the realized empirical distributions. The close correspondence between the two indicates successful implementation of the intended randomization.

Candidate religion is drawn from the following distribution: Islam (0.70), Protestant Christianity (0.10), Catholicism (0.10), Hinduism (0.05), and Buddhism (0.05). While these probabilities broadly reflect Indonesia's religious composition, minority religions are intentionally oversampled to ensure sufficient statistical power to study voter responses to minority candidates.

Candidate ethnicity is randomized using province-specific distributions benchmarked to census data. Respondents first report the province in which they vote, and candidate ethnicity is then drawn accordingly. To ensure comparability across provinces, the probability that a candidate belongs to the largest ethnic group in the province ("titular" ethnicity) is fixed at 0.50; candidates from other native ethnic groups ("non-titular," typically the second to fifth largest groups) are drawn with probability 0.40, and Chinese Indonesians (*Tionghoa*)—a politically salient minority present nationwide—with probability 0.10. Although specific ethnic labels (e.g., "Minang," "Balinese") are shown to respondents, the analysis classifies candidates into three categories: titular, non-titular, and Chinese.[9]

Finally, religion and ethnicity are randomized independently. This is motivated by evidence that religious and ethnic (anti-Chinese) prejudice in Indonesia are largely orthogonal (Mietzner and Muhtadi, 2019). Independent randomization therefore allows us to separately identify the effects of religion and ethnicity, as well as their interaction, including the electoral penalty associated with being a "double minority."

Valence. Candidates also differ in their valence (quality), which can be high (the candidate has a positive track record) or low. The valence issues are by definition constructed to be non-divisive. The valence item for each candidate is chosen from one of several issue areas: corruption; education and health; the economy; the environment; and the rights of women and children. Appendix Table E.1 reports the full list of statements that can be associated with a candidate. The valence item can be positive (high-valence) or negative (low-valence). For instance, high-valence descriptions for the issue area of corruption are "No track record

[9]We exclude a small number of highly implausible ethnicity–religion combinations to avoid distracting respondents. These combinations correspond to religion–ethnicity pairings with conditional probabilities below 0.0005 in the 2010 census.

of corruption" or "Successfully reduced the level of corruption in local government"; low-valence descriptions are sentences like "Previously involved in corruption cases" or "Involved in bribery cases for business permits."

For each profile, first one of the five issues is selected (with equal probability), then the polarity is selected (with equal probability), and finally a wording within the chosen issue-polarity combination is selected (with equal probability). The number of alternative items (specific wording) within a given issue-polarity combination is not balanced, yet the probability of selecting any one issue is constant at 1/5 and the probability of having a positive or negative valence trait is constant at one half. We use several alternative wordings for a given issue to increase the natural appearance of different descriptions across candidate profiles within any one choice. Being presented with two identical descriptions across the two profiles could be potentially confusing or distracting, and detract from the realism of the choice exercise. With this design, ex ante, the probability of facing two profiles with exactly the same description is quite small.[10]

Ideological position on contentious issues. We assign to candidates a position on one of several policy issues. These are chosen based on "divisive" issues that were discussed in the Indonesian public sphere in the recent past. The policy positions map directly to the ideological positions of respondents as measured in the baseline survey. Appendix Table E.2 reports the exact wording used. Each candidate is assigned a position on only one of the items, while during the baseline survey we ask respondents their positions on five issues. The policy areas covered are: the role of the Indonesian National Armed Forces; enforcement of traditional morality; freedom of association; freedom of speech; and immigration/economic nationalism. Traditional morality is measured via one of several sub-issues. For each profile, one of the issues is selected randomly, and then the polarity of the statement ("progressive" vs. "conservative") is selected with equal probability.

Recoding distances and creating aggregate traits. As previously mentioned, the stances on divisive issues for candidates in the candidate-choice exercise are assigned on the same dimensions on which respondent preferences were measured in the baseline survey. Each candidate is assigned a position on one of the five dimensions, while we observe respondents' positions on all five. In the analysis, we recode the traits on the divisive issues in terms of distance or congruence between the position of the respondent and that of the candidate.

[10] In the experiment as actually implemented, only around 5% of the pairs of profiles feature two candidates with the exact same description of the valence items, hence the modal respondent does not encounter choices with this type of redundancy.

How these distance-based variables are defined in practice differs slightly across the two approaches (random utility and non-parametric approaches) that we use to analyze the candidate-choice data, as per the pre-registration documents. In the random utility estimation, we use a "continuous" measure of distance: respondent and candidate have zero distance if they are both, for instance, supportive of traditional morality policies, and have high distance if the respondent is supportive of traditional morality and the candidate is progressive on this issue, or vice versa. The position of the candidate can be coded as -1 or 1, and the position of the respondent is coded so that "strongly agree" is coded -1, "agree" -.5, "disagree" .5, and "strongly disagree" 1. In the non-parametric estimation, we simply consider that candidates can be congruent with, or divergent from, the position stated by the respondent in the baseline. To this end, "strongly agree" and "agree", and "strongly disagree" and "disagree", are recoded into a binary agree/disagree feature. We then create binary indicator variables for when a profile is congruent with or divergent from the respondent's stance.

4 Empirical Framework

We analyze candidate choice in the utility maximization framework initially proposed by Ansolabehere and Snyder Jr (2000) and Londregan and Romer (1993), who model voter utility as determined by candidate valence and ideological distance between candidate and voter. This class of utility functions is adopted, for instance, in Kartik and McAfee (2007); Bernhardt *et al.* (2011); Kendall *et al.* (2015) and Kawai and Sunada (2025). As previously described, we study the choices made by respondents within pairs of hypothetical candidates for elected office where the features of the profiles are fully randomized. We elicit binary choices (rather than ratings of individual profiles or bipolar Likert scales) for two reasons: first, vote choices have this discrete all-or-nothing feature where voters can choose one candidate as opposed to communicating their rating; second, the binary choice data make it possible to directly fit the theoretical random utility model that is generally employed in the study of voting, e.g., Louviere *et al.* (2010). The utility framework has the advantage of permitting—conditional on the assumed model of decision-making—a direct interpretation of the estimates in terms of weights assigned to features of candidates in the utility function of the voter, something that might not be fully warranted in other frameworks (Abramson *et al.*, 2022).

We can write the utility that respondent i receives from candidate j based on ascriptive features (ethno-religious identity) on dimension n, valence issues m, and positional features on dimension k as

$$U_{ij} = \sum_{n} \delta_{in} \mathbb{1}(A_{in} = \tilde{A}_{jn}) + \sum_{m} \beta_{im} V_{jm} - \sum_{k} \gamma_{ik} (x_i^k - \tilde{x}_j^k)^2 + \epsilon_{ij} \quad (1)$$

where δ_{in}, β_{im}, and γ_{ik} are respectively the weights that respondent i assigns to ascriptive feature n, valence characteristic m, and positional issue k. Individual i has ascriptive feature A_i, while the candidate j has ascriptive feature $\tilde{A}_j$, and $\mathbb{1}$ is the indicator function. Individual i's observed position x_i^k on issue k is compared to the (randomly assigned) candidate position $\tilde{x}_j^k$. The difference in voter-candidate ideological position is squared, following the practice in the literature. This is convenient as it provides a function that is concave in ideological proximity and symmetric about zero, implying that the additional utility from closer alignment is larger when the ideological distance is further apart than when it is already closely aligned. The last term, ϵ_{ij}, is a random shock to the utility of candidate j for individual i.

Notice that the crucial distinction is between valence items, which all voters should value in the same way, and ideological stances, where citizens disagree (Ansolabehere and Snyder Jr, 2000; Londregan and Romer, 1993). Notice also that in this formulation, the relevant ascriptive feature of the candidate is congruence (sharing ethnicity or religion) with the voter. One could also consider religion and ethnicity as pure valence items, where for instance being a member of a given religion confers a "bonus" (or "malus") regardless of the religion of the voter, or on average across the electorate. We address this issue in Section C of the Appendix.[11]

In the main analysis, we proceed by assuming that the utility weights δ, β, and γ are constant across respondents. We then explore their heterogeneity across groups of respondents. The formulation of the utility function, following standard practice in the literature, embeds the assumption of additive separability of valence, identity congruence, and ideology. It is worth noting that departures from additive separability have observable counterparts that we assess below, finding no evidence of departures. Existing work using an approach close to ours (Kendall *et al.*, 2015; Cruz *et al.*, 2024) also finds little evidence that modeling the interactive elements of the utility function is consequential.

We observe, for a pair of candidates $j = 1, 2$, whether $U_{i1} \geq U_{i2}$ based on the stated choice across the pair of profiles. We can take the difference between U_{i1} and U_{i2}, as in Equation 1, for a pair of candidates (1,2). Define $\sigma_{ik} = -(x_i^k - \tilde{x}_1^k)^2 + (x_i^k - \tilde{x}_2^k)^2$, the difference in the ideological distance utility between the two candidates. This term ranges from -4, when candidate 1 is on the opposite side of the ideological dimension from the respondent, and

[11] For instance, candidates from a majority group may be perceived as better connected, leading even minority voters to prefer their district to be represented by a member of the majority group.

candidate 2 has a position that coincides with the voter bliss point, to 4 in the opposite situation; it takes a value of zero when both candidates are equally distant, or equally close, to the voter. As an aside, it is important to keep in mind that the information about the ideal points of respondents (the x_i^k) comes from the baseline survey, carried out weeks before the candidate-choice experiment.

Similarly, define $\xi_{in}^A = \mathbb{1}(A_i = \tilde{A}_1) - \mathbb{1}(A_i = \tilde{A}_2)$ as the difference between the ethnic (religious) congruence between the respondent and candidates. Notice $\xi_i \in \{-1, 0, 1\}$ indicating respectively a non-congruent candidate 1 vs. a congruent candidate 2, two equally (non-) congruent candidates, and a congruent candidate 1 vs. a non-congruent candidate 2.

The difference in valence $V_{1m} - V_{2m}$ takes values from -2 to 2, corresponding respectively to a low-quality candidate 1 matched with a high-quality candidate 2 and a high-quality candidate 1 matched with a low-quality candidate 2. In our exercise, differences of -1 and 1 are possible when one candidate is low quality on dimension m and the other has no information on that dimension (hence is imputed the expected value 0). A difference of zero implies that either the two candidates have the same level of valence on dimension m, or neither of the candidates has information about dimension m.

Given that in each profile, only one of the positional policy issues and one of the valence topics are displayed, we impute zeros to the position of candidates on the issues—for valence and for ideology—that are not displayed in the profile. Due to randomization (with equal probability) of "high" and "low" positions, zero is the expected value. In other words, if for instance a profile does not mention the candidate position on military intervention (which, if displayed, can take two values, -1 and 1) we treat the candidate as having an intermediate position 0 on that issue. This corresponds to the expected value (taken over the randomization of profiles) of the position of candidates on military intervention.[12]

We assume that the random components of the utility $\epsilon_{ij'}$ and $\epsilon_{ij''}$ are independently distributed for a pair (j', j''). This yields an error term ϵ_i for the difference in utilities. The predictors are based on the difference between the characteristics of the candidates in the pair, where the various elements of the utility of candidate 2 for respondent i are subtracted from their counterparts for candidate 1. Assuming constant weights across respondents, this yields an equation for the utility differential between candidates in a given pair as

$$\Delta U_i = \sum_n \delta_n \xi_{in}^A + \sum_m \beta_m (V_{1m} - V_{2m}) + \sum_k \gamma_k \sigma_{ik} + \epsilon_i \tag{2}$$

where $\epsilon_i = \epsilon_{ij'} - \epsilon_{ij''}$. Notice also that, following the pre-analysis plan, we do not include

[12] In the analysis with the non-parametric approach detailed in Section C of the Appendix, we investigate whether positive and negative features are actually symmetric in terms of their effect on choices.

in the random utility models the additional information (about age, occupation, education, marital status) that the profiles contain. Under independent randomization of the profile traits, these are orthogonal to the included features and therefore they can be comfortably relegated to the error term without introducing bias; in the AMCE analysis in Section C of the Appendix, we also study these additional candidate features.

The outcome variable we observe is whether the respondent chooses candidate 1 or 2, hence the sign of ΔU_i, and under the assumption of normality of the residual random part of the utility, this defines a probit model. Therefore, utility weights can be recovered by estimating a probit model where the outcome is an indicator variable for choosing candidate 1, equal to zero if the respondent chooses candidate 2. Notice that weights are identified up to the scale of the utility, which is fixed by choosing a variance for the latent random component of the utility. Fixing the variance at one yields the standard probit model. Given that the amount of randomness—the magnitude, i.e., variance, of the random shocks—depends on the specification, and this in turn defines the scale of the utility, one cannot compare absolute magnitudes of the coefficients across specifications; but it is still legitimate to compare ratios of coefficients. In the main analysis we fix the specification to include only the traits in Equation 1. Each observation is a respondent-pair combination, and we cluster standard errors at the respondent level.

Based on this general model, we run the main specification on the "aggregated" traits specification where a candidate is either ideologically close or distant, regardless of the specific policy issue mentioned in the profile; high- or low-valence, regardless of the topic; and congruent both in terms of ethnicity and religion, or different in at least one. Ultimately, this is equivalent to estimating a probit model with three predictors: difference in ethno-religious congruence between the two candidates, difference in valence of the two candidates, and difference in ideological congruence for the two candidates. The difference in ideological distances is given by comparing the distance between candidate 1 and the respondent on the dimension k_1 observed for candidate 1, and the distance between candidate 2 and the respondent on the dimension k_2 observed for candidate 2. These need not be the same: for instance, we calculate the distance between respondent and candidate 1 on the role of the military, and the distance between respondent and candidate 2 on traditional morality, and then treat these distances as homogeneous. Importantly, the difference in valence in this case takes on only the values -1, 0, and 1, as we treat the valence as high or low for a given candidate, regardless of the specific topic mentioned. Hence, either one candidate is high-valence and one is low-valence (hence values -1 and 1), or they both have the same valence (hence value 0). For convenience of interpretation of the estimates, the distance metric—which can range from -4 to 4, is rescaled by dividing by 4, and the topic-wise valence

differences are divided by 2, so that their range is the same as the ethnic and religious congruence and the aggregated valence differential. We also run the specification with full multidimensional traits, where we estimate each dimension of the three aggregated traits in the main model following the general random utility model previously discussed.

5 Results: Random Utility Model

Results for the main specification are reported in Table 1 for the aggregated traits and Figure 3 for the disaggregated, full multidimensional traits.[13] They show that, even in the control group—that is, in the absence of the treatment—the evaluation criteria, as reflected in the weights assigned to candidate attributes, differ sharply across voter types. Agency-minded voters place significantly greater weight on valence and ideology, despite no significant difference in the weight assigned to identity. In fact, for ideology, the estimated weight is indistinguishable from zero among Deferential voters but positive and significant among Agency voters.

In substantive terms, the bonus of having a high performance record is around four percentage points *larger* when facing an Agency-minded voter. Moreovoer, among initially indifferent Agency voters, ideological disagreement implies a vote loss of six percentage points; on the other hand, ideological disagreement is inconsequential (around 1 percentage point, but not statistically significant) among Deferential voters. The punishment for ideological disagreement is therefore around five percentage points larger when facing an Agency-minded voter. These patterns align with our prediction that narratives shape the weights voters assign to candidate attributes, a mechanism that we show translates into distinct voting behavior in the next section. Importantly, although identity receives a larger weight among Deferential voters, the between-group difference is not statistically significant when using the aggregate trait which define congruence both in ethnicity and religion. As we discuss below, however, a significant difference emerges when ethnicity and religion are examined separately. In general terms, this result suggests that while weighted less than valence, identity-based voting is present across different types of voters, including those who already hold an agency view of democracy.

When we perform the analysis with disaggregated traits, the message is analogous, but

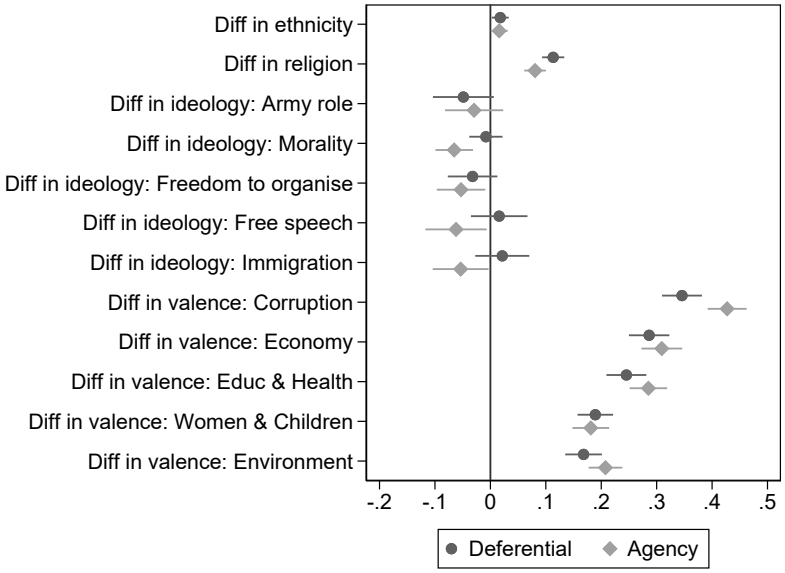
[13] Estimates from the unweighted sample are reported in the middle panel of Appendix Table A.8, where control group is used as the baseline. The results are highly similar to those obtained using survey weights and do not alter our conclusions. Moreover, we also run a robustness analysis where we use the continuous version of the heterogeneity based on the raw Agency score instead of the median-split dummy: results are presented in the first column of Appendix Table A.9. Moreover, the corresponding regression table for the disaggregated analysis is available in Appendix Table A.7

Table 1: Relative Importance of Traits – Control Arms

	$\boldsymbol{P(voted_1 = 1)}$			
Traits	Combined	Deferential	Agency	Difference
Identity	0.160***	0.174***	0.144***	**0.031**
	(0.019)	(0.026)	(0.028)	(0.038)
	[0.056]	[0.062]	[0.049]	
Valence	0.756***	0.692***	0.833***	**-0.141*** **
	(0.025)	(0.034)	(0.037)	(0.050)
	[0.263]	[0.246]	[0.282]	
Ideology	-0.092***	-0.024	-0.170***	**0.149*** **
	(0.024)	(0.032)	(0.034)	(0.047)
	[-0.032]	[-0.008]	[-0.058]	
Observations	26,378	12,248	14,130	26,378

Note: The table reports Probit estimates using the untreated (control) sample. The outcome is an indicator of choosing candidate 1 (left profile in the pair) and the regressors are the candidate traits, in aggregate form, following the specification in Equation (2). *Deferential* includes voters with Agency scores < 0; *Agency* includes those with scores ≥ 0; *Combined* pools both types. Column *Difference* reports the difference in weights assigned by Deferential and Agency voters. Robust standard errors clustered at the respondent level are in parentheses; marginal effects are in square brackets. *Identity* aggregates religion and ethnicity; *Valence* captures the candidate's track record on a key governance issue; *Ideology* is the respondent–candidate distance on a key ideological dimension. Survey weights are applied to ensure representativeness of Indonesia's 2024 vote-eligible population. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure 3: Weight Assigned to Traits – Disaggregated, Control Arms



Note: The graph presents the weight assigned to disaggregated candidate traits separately for Agency and Deferential voters. Each dot represents the marginal value of the probit regression estimates following the model in Equation 2. The bars represent 95% confidence intervals.

some interesting variation across traits and policy areas emerges. In particular, congruence on religion is much more important than on ethnicity. In addition, there is a statistically significant difference in the weight assigned to religious congruence by Agency and Deferential respondents: religious congruence confers an advantage that is around three percentage points larger when facing a Deferential respondent. [14] Among the valence traits, corruption carries the most weight, followed by the economy, and education and health; valence on environmental topics, and on women's and children's rights, rank last and have around half of the weight of corruption. For the ideological traits, the role of the army, traditional morality, and freedom of association have approximately the same weight, while free speech and economic nationalism carry less weight.

Moreover, the marginal effect estimates indicate that, when a voter is otherwise indifferent, having higher valence than one's rival will increase the probability of being chosen by approximately 26 percentage points. Sharing identity with the voter—when the other candidate does not—and being ideologically aligned with the voter—when the other does not—increases the probability of being chosen by around 6 and 3 percentage points respec-

[14]Notice that here we treat ethnic congruence as a binary trait: e.g., a respondent of the titular ethnicity is classified as facing a non-congruent candidate both if the candidate is of Chinese ancestry or a member of a non-titular ethnicity.

tively. In other words, since the weight assigned by voters at baseline to valence is much higher than the weight assigned to identity, this implies that respondents would prefer to vote for a high-valence candidate from an ethno-religious out-group than a low-valence candidate from their own ethno-religious group.[15] However, holding valence differential constant, respondents display a (small) preference for ethno-religiously congruent candidates with whom they disagree about divisive issues over ethno-religiously incongruent candidates with whom they agree on divisive issues.[16]

5.1 Video Treatment Effects

Table 2 reports the estimates of treatment effects.[17] Treated respondents place greater weight on valence, and separate analysis for Deferential and Agency voters reveals that the effect is concentrated among voters who did not initially frame elections in agency terms: within this group, treatment increases the weight on valence by about 19 percent. In contrast, among voters who already thought about elections in agency terms, the treatment effect is statistically indistinguishable from zero. In addition, the treatment has a significant effect on the weight on ideology, and again, this is driven mostly by Deferential voters. We do not detect any effect of the treatment on the weights assigned to identity considerations.[18]

Moreover, the difference in treatment effects across Deferential and Agency voters is statistically significant for valence but not for ideology (see Subsection 5.2 for formal tests). This pattern provides a more nuanced understanding of how the treatment operates: Since Agency and Deferential voters differ substantially in the importance they attach to valence; the treatment effectively closes this gap. Specifically, the treatment does not alter Agency voters' valuation of valence, but it significantly increases the weight placed on valence by

[15]The literature has suggested that voters may apply different standards (Landa and Duell, 2015) or tolerate poor governance from in-group candidates —even when they have criminal backgrounds or track records of corruption (Vaishnav, 2017; Klašnja *et al.*, 2018). Voters might also support parties that work against their economic interests if they align with their identity group (Banerjee and Pande, 2007; Suryanarayan, 2019). Some studies show that voters do penalize in-group candidates with poor records (Banerjee *et al.*, 2014; Chauchard, 2016).

[16]This result points to a less prominent role played by ideology in the Indonesian context, which may deviate from what obtains in advanced democracies.

[17]Estimates from the unweighted sample are reported in Appendix Table A.8. The results are highly similar to those obtained using survey weights and do not alter our conclusions. Moreover, we also run a robustness analysis where we use the continuous version of the heterogeneity based on the Agency score instead of the median-split dummy: results are presented in the first column of Appendix Table A.9.

[18]From the last two rows of Table 2, we can conclude that the randomization of profiles *and* of the treatment works well: First, the treatment has a zero effect on the probability of Profile 1 (the left profile) to be chosen by the respondents. Similarly, the constant is indistinguishable from zero, meaning that the base probability of choosing the left-hand side of the profile pair is one half—the probability of choosing the left profile is exactly half when the value of the Probit predictor is zero.

Table 2: Treatment Effects

	$P(voted_1 = 1)$			
Traits & Treatment	Combined	Deferential	Agency	
Identity × Treated	-0.006	-0.021	0.012	
	(0.027)	(0.036)	(0.040)	
Valence × Treated	**0.062***	**0.133*****	-0.023	
	(0.035)	(0.048)	(0.052)	
Ideology × Treated	**-0.090*****	**-0.112****	-0.070	
	(0.034)	(0.047)	(0.049)	
				Deferential (treated)
Identity	0.158***	0.172***	0.143***	0.151***
	(0.019)	(0.026)	(0.028)	(0.0253)
Valence	0.759***	0.693***	0.836***	0.827***
	(0.025)	(0.034)	(0.037)	(0.0335)
Ideology	-0.095***	-0.029	-0.171***	-0.140***
	(0.024)	(0.033)	(0.034)	(0.0332)
Treated = 1	0.005	0.014	-0.005	
	(0.019)	(0.026)	(0.028)	
Constant	-0.021	-0.028	-0.012	
	(0.013)	(0.018)	(0.018)	
Observations	52,136	25,020	27,116	

Note: The table reports treatment effects on the relative importance (weights) assigned to candidate traits. *Deferential* includes voters with Agency scores < 0; *Agency* includes voters with scores ≥ 0; *Combined* pools both types. The last column reports the "total effect" of each trait for Deferential voters under treatment. Robust standard errors clustered at the respondent level are in parentheses. *Identity* aggregates religion and ethnicity; *Valence* captures the candidate's track record on a key governance issue; *Ideology* is the respondent–candidate distance on a key ideological dimension. Survey weights are applied to ensure representativeness of Indonesia's 2024 vote-eligible population. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Deferential voters. Meanwhile, with respect to ideology, the treatment increases the weight placed on ideological proximity for *both* Agency and Deferential voters, with a somewhat larger and significant effect for the latter; as a result, the difference in treatment effects across voter types is not statistically significant for ideology. By contrast, the two types of voters do not differ in the weight they assign to identity, the treatment has no effect along this dimension, as there is no gap to be closed in the first place.

The last column of Table 2 formally tests whether the informational treatment narrows the gap in attribute weights between voter types. Among Deferential voters, the treatment raises the weights assigned to valence and ideology toward Agency-level benchmarks, while leaving the weight on identity essentially unchanged. The statistically significant treatment effects and the resulting convergence in preferences between Deferential and Agency voters provide causal evidence that lay conceptions of democracy are malleable and that shifts in narratives can reshape the criteria voters use to evaluate political candidates. From a policy perspective, these results are encouraging, as they suggest that if Agency-type preferences are normatively desirable, a simple and scalable informational intervention can improve the quality of electoral selection.

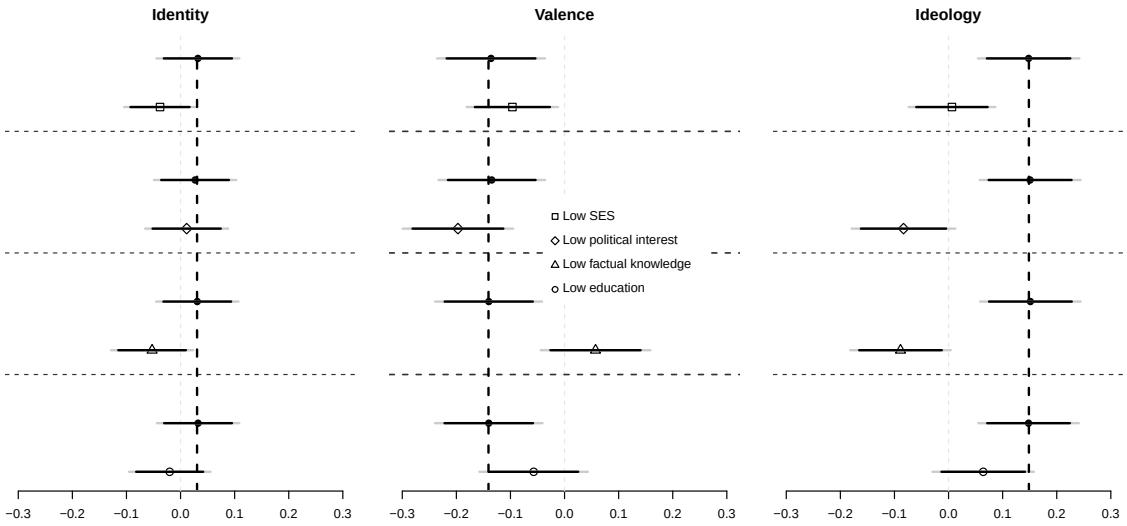
5.2 Evidence on specific mechanism

The evidence presented in the previous subsections documents two main patterns. First, there is substantial heterogeneity—predicted by the distinction between deferential and agency narratives—in the weights assigned to candidate valence and ideological distance. Second, the treatment has a statistically significant effect on the weights placed on valence and ideological distance, and this effect is driven primarily by respondents classified as Deferential. Crucially, following treatment, the behavior of these respondents converges toward that of respondents who were initially classified as Agency.

Our proposed interpretation is that the treatment operates by intervening on individuals' underlying narratives of democratic representation. First, the treatment *explicitly* explains the agency view in layperson terms, and thus directly targets the individual conception of democracy. We provide evidence on this in Section 7, by showing a direct, causal effect of the treatment on the narrative score.

Second, the fact that the intervention has a statistically significant effect among respondents who held a deferential narrative prior to treatment—but no detectable effect among those with an Agency view—strongly suggests that the treatment operates at the margin of changing individuals' conceptualization of democratic representation. Respondents who already subscribed to an Agency view are unaffected because the intervention does not convey

Figure 4: Heterogeneity Among Controls: The "Horse Race" Model



Notes: The figure compares differences in the weights assigned to candidate traits across alternative dimensions of heterogeneity. Each symbol represents the coefficient from a Probit regression on the interaction between candidate traits and a given heterogeneity dimension. Coefficients within each group—separated by horizontal dashed lines—are estimated from a single regression, each comparing heterogeneity based on Deference with an alternative dimension. Bars indicate 95% and 90% confidence intervals. The solid dot symbol corresponds to heterogeneity based on the Agency score (dummy for below-median values), while the hollow symbols represent alternative dimensions, as indicated in the legend (also coded as below-median dummies). The vertical dashed line denotes the heterogeneity from a regression that includes Deference as the sole heterogeneity dimension.

information or perspectives that are novel relative to their pre-existing beliefs.

Nevertheless, one might be concerned that the heterogeneity in preferences and treatment effects documented above is driven by other individual characteristics that are correlated with respondents' narratives, rather than by narratives themselves. In particular, respondents classified as Deferential may differ systematically along dimensions such as education, socioeconomic status, political knowledge, or political interest, and variation along these dimensions—rather than narrative beliefs per se—could be responsible for the observed heterogeneity.[19]

Here we provide additional evidence that strongly supports our interpretation of the results. We first show that, in the control group, the variation in weights assigned to valence and ideology driven by narrative is robust to the inclusion of measures of other variables

[19]It is important to note that these characteristics are indeed significant predictors of the Agency score, as shown in Appendix Table A.5. However, while these factors explain some variation in the Agency score, they fall far short of accounting for it fully. A substantial share of variation remains unexplained, as reflected in the very low R-squared values of these models. Our interpretation is that this residual component captures meaningful differences in individuals' specific understanding of agency in representative democracy.

related to alternative interpretations. Specifically, we estimate "horse-race" models in which each trait variable is interacted both with a dummy for Deferential respondents and with a variable capturing an alternative dimension of heterogeneity. This specification allows us to assess whether the inclusion of these alternative dimensions as controls affects the main estimates presented earlier, and to evaluate the relative importance of different sources of heterogeneity in explaining the observed variation.[20]

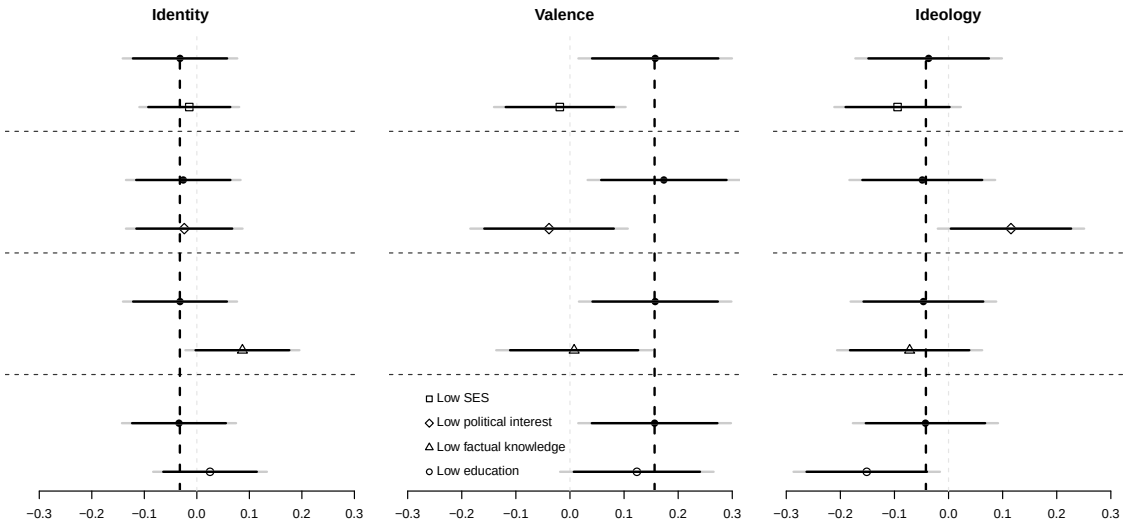
Figure 4 presents the main coefficients from these regressions (with the full set of estimates reported in Appendix Table A.10). The solid dot markers—which correspond to heterogeneity driven by Deference—are closely aligned across alternative specifications, indicating that the estimated variation driven by Deference is highly stable regardless of which alternative heterogeneity dimension is included. Moreover, these estimates are very similar to those obtained from specifications in which Deference is the sole source of heterogeneity, as presented in the main results. This pattern holds consistently across all traits considered in the analysis: identity, valence, and ideology. In contrast, the estimates associated with alternative dimensions of heterogeneity vary substantially across specifications and, in some cases, even differ in sign. Taken together, these findings strongly suggest that the differences in the weights assigned to valence and ideology by Agency and Deferential respondents are not simply a spurious reflection of correlated socioeconomic or political characteristics, but instead reflect a distinct and robust role of narrative-based heterogeneity.

We can also perform a similar analysis on the heterogeneity of the treatment effects, i.e., estimate how much the lay conception of democracy drives treatment effect heterogeneity while accounting for the role of another individual characteristic in driving treatment effect heterogeneity.[21] Figure 5 reports these estimates (detailed estimates are available in Appendix Table A.11). Again, we observe that the treatment effect heterogeneity driven by Deference is always stable across models and is identical to the one estimated using Deference as the sole heterogeneity dimension. In contrast, the estimates associated with alternative dimensions of heterogeneity vary substantially across specifications and, in some cases, even

[20]The alternative dimensions we examine are, in turn: not having university-level education; having below-median factual knowledge of Indonesian politics; having below-median interest in politics, measured by self-reported media consumption at baseline; and having low socioeconomic status (SES). Moreover, Table A.9 in the Appendix replicates these results with the continuous counterparts of these variables.

[21]In practice, we estimate models that include three-way interactions between each profile characteristic (identity, valence, and ideology), the treatment indicator, and a dummy for Deferential respondents, as well as three-way interactions between each profile characteristic, the treatment indicator, and a dummy for each alternative individual characteristic considered in the previous exercise. All corresponding two-way interactions among the constituent terms are also included. The coefficients on the three-way interactions capture how the treatment-induced changes in the weights assigned to identity, valence, and ideology vary jointly by narrative type and by individual characteristics. This specification therefore allows us to assess the extent to which treatment-effect heterogeneity is explained by narrative, net of the heterogeneity accounted for by other individual characteristics.

Figure 5: Heterogeneity of Treatment: The "Horse Race" Model



Notes: The figure compares differences in the treatment effects across alternative dimensions of heterogeneity. Each symbol represents the coefficient from a Probit regression on the three-way interaction between candidate traits, treatment status, and a given heterogeneity dimension. Coefficients within each group—separated by horizontal dashed lines—are estimated from a single regression, each comparing heterogeneity based on Deference with an alternative dimension. Bars indicate 95% and 90% confidence intervals. The solid dot symbol corresponds to heterogeneity based on the Agency score (dummy for below-median values), while the hollow symbols represent alternative dimensions, as indicated in the legend (also coded as below-median dummies). The vertical dashed line denotes the coefficient on the three-way interaction from a regression that includes Deference as the sole heterogeneity dimension.

differ in sign.

5.3 Treatment Effects on Real-world Votes

We next provide evidence that the treatment effect extends beyond stated preferences in the candidate-choice experiment and translates into actual voting behavior. Specifically, we show that, following treatment, the voting behavior of Deferential respondents converges toward that of Agency respondents. This convergence indicates that the intervention affects not only how respondents evaluate hypothetical candidates, but also how they behave in real electoral settings.

A natural—though somewhat naïve—benchmark would be to test whether the treatment unconditionally increases the vote share of higher-valence candidates. Unfortunately, such an exercise is neither feasible nor informative in our empirical context. The voting data we analyze pertain to presidential elections, which feature a small number of highly visible

candidates with long political careers.[22] Entry into these races reflects extensive prior selection based on experience, perceived electability, and popularity, leaving little meaningful variation in observable performance or track records across candidates.[23] Moreover, the race does not include a clear incumbent, as the president was term-limited after two terms in office. At the same time, not one but two of the candidates could present themselves as "continuity" tickets with the incumbent administration, given that one had the son of the incumbent president as a vice-presidential candidate and the other was the official ticket of the party under which the incumbent president ran in 2019 and 2014. Also, taking into account the non-partisan nature of the intervention, which does not provide new information about the candidates themselves, one would not expect the treatment to shift aggregate support toward any particular candidate in absolute terms, as respondents would have formed their (however imprecise) beliefs about valence and ideology earlier in the campaign.

Accordingly, we examine realized voting behavior conditional on respondents' pre-treatment vote intentions. This approach allows us to isolate changes in choice behavior that are plausibly attributable to the treatment, rather than to underlying partisan preferences or candidate characteristics. The key purpose of this exercise is therefore not to demonstrate electoral reallocation across candidates per se, but to establish that the treatment generates real-world behavioral consequences that are consistent with our experimental findings, in particular in the sense of making the choices of Deferential respondents closer to those of Agency respondents.

The election occurred 4 to 20 days after the respondents received the video treatment. In the pre-election survey, before the treatment, we elicited presidential vote intention; in a post-election follow-up, we recorded self-reported presidential vote choice. Based on national polls and media coverage, the presidential election of 2024 was a three-way race with a clear favorability ranking: a favored ticket ("front-runner"), a serious contender ("contender"), and an "underdog" ticket.[24]

[22] This analysis is not pre-registered.

[23] This type of test would be more feasible in low-level legislative races with many candidates of limited visibility or shorter political histories, where distinctions in candidate valence could be more clearly identified.

[24] This last ticket involved the party with which the incumbent president had run in 2019, but was not endorsed by the president himself.

Table 3: Presidential Votes: Intention vs. Actual

Intention	**Actual votes**				
	Control - Agency (N=672)				
	Contender	Underdog	Front-runner	No vote	Total
Contender	89.31	0.00	5.66	5.03	100.00
Underdog	7.25	**66.67**	26.09	0.00	100.00
Front-runner	1.95	2.28	92.51	3.26	100.00
No vote	14.29	0.00	28.57	57.14	100.00
Don't know	**24.39**	17.07	**47.15**	11.38	100.00
Total	27.53	11.01	55.51	5.95	100.00
	Control - Deferential (N=590)				
Contender	86.32	2.11	6.32	5.26	100.00
Underdog	10.67	**52.00**	32.00	5.33	100.00
Front-runner	3.41	2.73	92.15	1.71	100.00
No vote	20.00	0.00	20.00	60.00	100.00
Don't know	**17.86**	10.71	**62.50**	8.93	100.00
Total	20.85	10.34	63.22	5.59	100.00
	Treated - Deferential (N=652)				
Contender	83.33	1.52	10.61	4.55	100.00
Underdog	5.41	**75.68**	18.92	0.00	100.00
Front-runner	3.28	0.98	91.80	3.93	100.00
No vote	10.53	0.00	26.32	63.16	100.00
Don't know	**31.97**	9.84	**52.46**	5.74	100.00
Total	25.31	11.20	57.82	5.67	100.00
	Treated - Agency (N=625)				
Contender	88.28	2.34	7.03	2.34	100.00
Underdog	7.35	**67.65**	22.06	2.94	100.00
Front-runner	2.42	1.38	91.35	4.84	100.00
No vote	25.00	8.33	8.33	58.33	100.00
Don't know	**29.69**	11.72	**51.56**	7.03	100.00
Total	26.56	11.04	56.80	5.60	100.00

Notes: The table presents the distribution of the actual votes (self-declared during the post-election survey) conditional on the vote intention declared in the pre-election survey. Each number in this table is the percentage of voters who initially declared the intention to vote for the candidate in the row (Intention) and eventually voted for the candidate in the column (Actual votes), e.g., the first cell of the table shows that 89.31% of voters who initially intended to vote for the contender eventually voted for the contender, while the cell below it shows that 7.25% of voters who initially intended to vote for the underdog eventually voted for the contender.

In our pre-election survey, around 20 percent of respondents stated they were still undecided, and eleven percent intended to vote for the underdog. Of particular interest is understanding how these voters ultimately behave: how they make up their mind if they were initially undecided, and whether they stick to their initial choice of the underdog as it becomes clearer, in the final phase of the election campaign, that it is a hopeless ticket. The top two panels of Table 3 report the transition matrix for the Agency and Deferential voters, respectively, in the control arm. Each entry is the conditional probability of supporting one of the options in the columns, conditional on the vote intention in the pre-election survey

(in the rows), i.e., the fraction of respondents in a row that ultimately voted for one of the options in a column.

In the control group, there are clear differences in the vote-switching patterns of Agency and Deferential voters.[25] Take the initial supporters of the underdog. Two-thirds of the Agency-minded stick with their initial choice, while only around half of the Deferential do; this difference is statistically significant. When it comes to the initially undecided, six out of ten of the Deferential choose the front-runner, compared with less than half of the Agency-minded, and this difference is statistically significant.[26]

Most importantly, the experiment allows us to estimate how inducing an agency narrative shifts the final choices of Deferential voters. The bottom panel of Table 3 reports the transition matrix for treated Deferential voters. The shares of previously undecided voters who chose the contender and the front-runner move toward the levels observed among Agency voters; likewise, roughly three-quarters of treated voters who initially backed the underdog remain with that choice. In short, the informational treatment brings Deferential voters' revealed choices closer to those of Agency-minded voters, in line with the convergence pattern observed in the discrete-choice exercise.

To formally test the patterns described above, we regress post-election vote choices on pre-election vote intentions, interacting initial intention with treatment assignment. The results are reported in Table 4. Among respondents who initially supported the underdog, those assigned to the treatment are significantly more likely to remain with that choice and less likely to switch to the front-runner. Similarly, among initially undecided respondents,

[25] Full regression analysis showing the difference in voting behavior across the two types of voters is provided in Appendix Table A.12.

[26] Because the primary objective of this exercise is to establish that the behavior of Deferential voters converges toward that of Agency voters, we do not focus on disentangling the underlying reasons for differences in voting behavior across voter types. Nonetheless, we identify several closely related mechanisms that are consistent with—and help interpret—the observed patterns: First, as argued by Hicken *et al.* (2018), vote-switching may reflect direct clientelistic inducements received during the campaign period, such as outright vote buying or promises of rewards conditional on electoral support. Agency-minded voters should be substantially less susceptible to such inducements, as they are more likely to evaluate candidates based on performance and accountability considerations rather than contingent benefits. Second, when elections are perceived primarily as popularity contests rather than as mechanisms for selecting agents, campaign-period "horse-race" information becomes more influential. In such contexts, information about which candidate is leading shortly before election day may induce vote-switching if voters' objective is to support the eventual winner rather than to select the most competent representative (Lanoue and Bowler, 1998; Mutz, 1998; Rothschild and Malhotra, 2014). Third, evidence from the candidate-choice experiment indicates that ideology and track record play a substantially larger role in the evaluations of Agency voters. To the extent that the ideology and track record of presidential candidates—particularly those with long political careers—can be assessed with reasonable precision well before the election, the preferences of Agency voters are likely to be formed earlier in the campaign based on verified information. As a result, their choices should be comparatively stable and less responsive to short-term campaign dynamics in the period immediately preceding the election.

Table 4: Treatment Effects on Actual Votes—Deferential Voters

	Actual votes (self-declared)			
Intention & Treatment	Front-runner	Underdog	Contender	No vote
intent: Underdog x Treated	**-0.174****	**0.243*****	-0.023	-0.046
	(0.080)	(0.079)	(0.066)	(0.039)
intent: Front-runner x Treated	-0.046	-0.012	0.028	0.029
	(0.043)	(0.021)	(0.050)	(0.032)
intent: No Vote x Treated	0.020	0.006	-0.065	0.039
	(0.150)	(0.018)	(0.134)	(0.171)
intent: Undecided x Treated	**-0.143***	-0.003	**0.171****	-0.025
	(0.074)	(0.044)	(0.074)	(0.045)
Treated = 1	0.043	-0.006	-0.030	-0.007
	(0.037)	(0.018)	(0.048)	(0.029)
intent: Underdog	0.257***	0.499***	-0.756***	0.001
	(0.060)	(0.060)	(0.050)	(0.035)
intent: Front-runner	0.858***	0.006	-0.829***	-0.036
	(0.030)	(0.018)	(0.037)	(0.024)
intent: No vote	0.137	-0.021	-0.663***	0.547***
	(0.107)	(0.015)	(0.110)	(0.129)
intent: Don't know	0.562***	0.086***	-0.685***	0.037
	(0.052)	(0.033)	(0.051)	(0.036)
Base Intent: Contender	0.063**	0.021	0.863***	0.053**
	(0.025)	(0.015)	(0.035)	(0.023)
Observations	1,242	1,242	1,242	1,242
R-squared	0.486	0.427	0.516	0.176

Note: The table presents the linear probability model of actual presidential votes on the respondents' own vote intention and the treatment indicator. The analysis in this table is restricted to respondents classified as Deferential voters. Actual votes are self-declared during the post-election survey, while vote intention is declared during the pre-election survey. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

treated individuals are significantly less likely to gravitate toward the front-runner and more likely to vote for the contender. These effects are statistically significant and consistent with the transition patterns previously documented.

6 Additional Analysis

This section presents a set of analyses that corroborate the main findings from the random utility model. First, we examine response time as an alternative outcome to stated choice to assess whether the main patterns persist. Second, we analyze the candidate-choice data using a non-parametric Average Marginal Component Effect (AMCE) framework. Finally, we re-estimate candidate trait weights using subsets of profile pairs to verify that the estimated weights are stable and not sensitive to the number or the order in which profile pairs are presented to respondents. Across all exercises, the results closely align with the random utility estimates and support the same substantive conclusions. We discuss each in turn below.

6.1 Response Time Analysis

Rather than relying on revealed preferences from the choice task, we use *response time on candidate pairs* as the outcome. Although response time has a less direct interpretation, it provides an alternative way to infer the importance of traits and is plausibly less susceptible to social desirability bias.

In the candidate-choice exercise, on top of recording the individual choices made by the respondents over the pairs of candidates shown, we also record the response time (in milliseconds) for each pair of profiles. Importantly, response time is a completely non-obtrusive measure of the decision process: while the respondents might suspect that their response time is being recorded, they are not necessarily cognizant of the fact that this is in itself an outcome studied. Hence, much more than in the case of profile choice, demand effects here are going to be minimal. While a respondent might choose candidates as if religion did not matter for them for reasons of social acceptability (e.g., it is not respectable to favor candidates of one's own religion, hence respondents try to choose at least some religious out-group candidates in the course of the exercise), it is hard to imagine that they would strategically delay their response to project a more socially acceptable image of themselves as decision-makers.

The response time data allow us to further shed light on the importance respondents assign to various traits of candidates, in a fixed-effects (within-subject) design. Arguably,

longer response time indicates a harder choice for the respondent to make. In particular, if respondents base their decision on a particular trait of the candidate, when facing two candidates with the same trait (e.g., both are of the same religion) they have to scan the profile to look for other characteristics that make it possible to choose one over the other. This strategy, to our knowledge, has not been used in political economy studies, but it is a relatively well-established approach in market research (Haaijer *et al.*, 2000; Moskowitz and Silcher, 2006). We model log response time in the control group as a function of the similarity in candidate traits in a given pair, with indicators that take the value of one if the profiles share that feature, and zero otherwise. If religion is an important consideration for a respondent, showing two profiles with the exact same religion will delay the time to choose, as the respondent needs to closely inspect some other pair of traits in order to make a choice. We estimate linear regressions with individual fixed effects, with standard errors clustered by respondent. A positive coefficient on the indicator for pairs with the same ethnicity, or on the indicator for pairs with the same valence, would indicate that choosing between profiles with the same ethnicity is harder and analogously that choosing between profiles that differ in terms of valence is easier. Given the full randomization of the candidate traits, all these coefficients have a causal interpretation. Individual fixed effects remove variation across individuals in reading speed, cognitive ability, or impulsiveness, increasing precision. The log-linear specification ultimately models response delay as a proportional increase in response time.

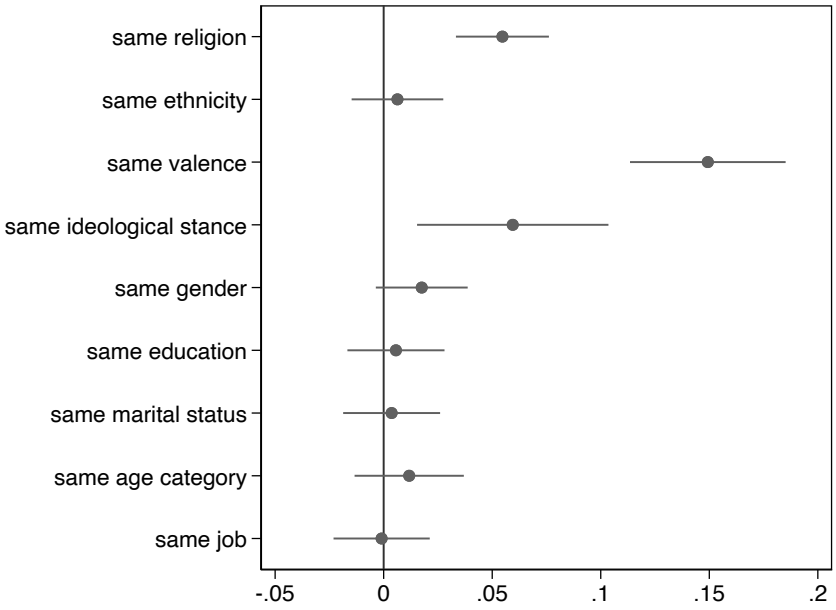
The log time that respondent i spent on choice pair k is modeled as

$$\log(t_{ik}) = \alpha_i + \theta_1 \mathbb{1}(R_1 = R_2) + \theta_2 \mathbb{1}(E_1 = E_2) + \theta_3 \mathbb{1}(V_1 = V_2) + \theta_4 \mathbb{1}(\tilde{X}_1 = \tilde{X}_2) + \eta_{ik} \quad (3)$$

where R_j and E_j respectively denote the candidate's religious and ethnic affiliations, $V_j \in \{\text{high}, \text{low}\}$ is the candidate's valence, $\mathbb{1}(\cdot)$ is an indicator function, and α_i is a respondent fixed effect. Valence is considered identical if both candidates have high performance or if both have low performance, regardless of the policy area. Meanwhile, we define $\tilde{X}_j = \tilde{x}_j^m$ if ideological trait m was presented for candidate j, meaning they are considered identical if both position and policy dimension are identical, e.g., both have a conservative position on traditional morality. The coefficients θ capture the (causal) effect on response time of the two candidates sharing an identical trait. Because these regressors are same-trait indicators, all θ are on a common scale and can be compared directly. In the log-response-time specification, $100(\exp(\theta) - 1)$ gives the percentage change in response time when the two candidates share the corresponding characteristic.

Results are reported in Figure 6, with detailed regression results in Appendix Table

Figure 6: Weight Assigned to Traits – Response Time, Control Arms



Note: The graph presents the weight assigned to traits of candidates using response time as the outcome in the control arm only. Each dot represents the OLS regression estimates following the model in Equation 3. Survey weights are applied to ensure representativeness of Indonesia's 2024 vote-eligible population. The bars represent 95% confidence intervals.

A.13. We can observe that the coefficient on valence is again the largest: when faced with two candidates with the same valence characteristic, respondents take about 15% longer to choose. The next largest coefficients are the ones on ideology and religion: respondents take 5.1% longer to choose if faced with candidates of the same religion, and 6.7% longer when the two candidates have the same stance on divisive positions. The remainder of the coefficients are not statistically distinguishable from zero. Overall, the message is in line with the random utility estimates (and the non-parametric estimates discussed in the next subsection) in substantive terms: respondents choose based first and foremost on valence traits, but also assign importance to religion of the candidates and stances on divisive issues.

We also find cross-sectional evidence that the treatment slows response times among Deferential respondents, making their behavior more similar to Agency voters. Table 5 reports regressions of log response time on an indicator for treatment status, an indicator for Deferential respondents, and pre-treatment demographics to account for systematic differences in response speed. Because treatment was randomly assigned, these controls serve primarily to improve precision; they also enable a more targeted interpretation of the heterogeneity that we document.

Conditional on demographics, we find that Deferential voters in the control group spend significantly less time—about 17% less—choosing between profiles. The treatment has no statistically significant effect on response time for Agency-minded voters. By contrast, among Deferential voters, the treatment increases response time by roughly 15%. Calculating the total effect, we find that treated Deferential respondents are not statistically different from Agency-minded voters in the control group. The results in Column (2), which includes a broader set of covariates, are nearly identical. Overall, the evidence indicates that the informational intervention makes Deferential voters behave more like Agency-minded voters, plausibly by shifting their narrative of democracy in line with the video content.

A plausible alternative interpretation is that the treatment simply makes less-informed (Deferential) voters spend more time on pair choices. While possible in principle, the evidence does not support this view. Columns (3)–(4) replicate the specifications in Columns (1)–(2) but add a measure of actual political knowledge—estimated from answers to questions about Indonesian politics—and an interaction between treatment and political knowledge. These specifications provide ancillary evidence against a simplistic interpretation of the treatment effect heterogeneity based on deferential vs. agency narratives as a "treatment slows down the less informed" mechanism.

Specifically, in the control condition both *low-knowledge* (based on answers to factual political questions) and *Deferential* respondents spend less time on candidate pairs on average. However, the treatment effects for these two groups go in opposite directions: the treatment

Table 5: Treatment Effects on Response Time

	Log(time spent on pairs)			
	(1)	(2)	(3)	(4)
Treated = 1	-0.062	-0.057	0.009	0.018
	(0.044)	(0.044)	(0.052)	(0.052)
Deferential = 1	-0.182***	-0.178***	-0.178***	-0.174***
	(0.042)	(0.042)	(0.042)	(0.042)
Treated × Deferential	0.139**	0.130**	0.158***	0.149**
	(0.061)	(0.061)	(0.061)	(0.061)
Factual knowledge low = 1			-0.101**	-0.103**
			(0.042)	(0.041)
Treated × Factual knowledge low			-0.132**	-0.137**
			(0.061)	(0.060)
Demographic controls	✓	✓	✓	✓
Religion & worldview controls		✓		✓
Observations	52,134	52,134	52,134	52,134
R-squared	0.027	0.030	0.036	0.039

Note: The table reports OLS regressions of the logarithm of time spent on candidate pairs on indicators for treatment status, a Deferential respondent indicator, and their interaction. All specifications control for pre-treatment demographic covariates, including gender, age group, college education, ethnicity, and socio-economic status. Religion/worldview controls include a Muslim indicator, an index of religiosity (constructed from baseline prayer frequency and perceived importance of religion), and an indicator for an Islamist versus civic-nationalist (Pancasila) worldview. Columns (3)–(4) additionally include a binary measure of factual knowledge of Indonesian politics (below/above median) and its interaction with treatment. Robust standard errors clustered at the respondent level are reported in parentheses. Survey weights are applied to ensure representativeness of Indonesia's 2024 vote-eligible population. *** p<0.01, ** p<0.05, * p<0.10.

increases the response time of Deferential voters (rendering them statistically indistinguishable from Agency voters in the control group), whereas *low-knowledge* voters spend even less time under treatment.[27] Taken together, these patterns suggest that the treatment effect is more consistent with a shift in democratic narratives—and the choice criteria—among Deferential voters.

6.2 Non-parametric approach

We re-estimate candidate trait weights and treatment effects using a non-parametric Average Marginal Component Effect (AMCE) framework. As detailed in Appendix Section C, the results are largely consistent with—and closely replicate—those from the main random utility model, strengthening confidence in our overall findings.

Beyond serving as a robustness check, the AMCE approach offers greater flexibility in parameterization and thus has several additional advantages. First, it allows us to assess the validity of the assumptions and modeling choices underlying the main specification. In particular, as discussed in Appendix Subsection C.1 and shown in Appendix Table C.2, we find that the additive separability assumption embedded in the primary random utility model is satisfied.[28]

Second, the AMCE framework allows us to estimate the component effects of the original candidate traits directly, rather than only in congruence terms. As shown in Appendix Figure C.1, the effects of identity, valence, and ideological distance are substantially larger than those of education, age, gender, or previous occupation. This further justifies the choice of core traits in the random utility model, as these characteristics are precisely those that matter most for voter choice.

Moreover, this approach allows us to examine asymmetric effects of high versus low valence across policy areas. Specifically, we estimate separate effects of having a positive versus a negative track record in each of the five policy domains. As shown in Appendix Figure C.2, the effects are approximately symmetric for education, the environment, and the rights of women and children. In contrast, for corruption, the *malus* associated with a poor record is larger in absolute value than the *bonus* from a good record, whereas for the economy, the *bonus* from a strong track record exceeds in absolute value the *malus* from a

[27] This last result also obtains if we omit the Deferential indicator and the treatment×Deferential interaction, so it is not an artifact of over-conditioning.

[28] These results suggest that sharing an identity with a candidate does not alter how voters value valence attributes. This is reassuring and strengthens a claim we advanced cautiously earlier: when confronted with a corrupt in-group candidate and a clean out-group candidate, voters tend to choose the non-corrupt candidate despite religious differences. Similarly, the weight placed on ideological congruence does not depend on whether voters share the candidate's identity.

weak one.

6.3 Fatigue effect

We also consider the possibility that our earlier results may be sensitive to the number of profile pairs presented to respondents or to the order in which candidates are shown. To address this concern, we re-estimate the weights assigned to candidate traits using rolling windows of five profile pairs. This approach allows us to assess whether the estimated weights are stable across different presentation orders and when based on a substantially smaller number of comparisons. The results are reported in Appendix Figure A.2. We find that the estimated weights are remarkably stable across different orderings—from the first to the last set of five pairs. This stability holds for all traits examined, including identity, valence, and ideology.

7 Direct Effect on Narrative, Persistence, and Experimenter Demand

A natural question is whether the effects of a brief, lightweight video intervention persist beyond immediate exposure. We leverage an unplanned, minor protocol deviation: a small subset of respondents also played the *drag & drop* categorization game in the post-election (wave 3), despite already playing it in the pre-election (wave 2). For 664 respondents (347 treated and 317 controls), we observe post-treatment game responses a few weeks after wave 2. This does not affect our main results—respondents received the treatment only once, in wave 2—but it allows a direct, measurable assessment of whether the intervention shifts voters' narratives over time. Conceptually, this "narrative shift" can be viewed as a *first stage* in an instrumental variables framework, where the ultimate outcomes are voters' trait weights and (real-world) vote choice, and the endogenous regressor is the extent of agency in voter narrative (the Agency score in the *drag & drop* categorization game) induced by the video, which acts as the instrument.

Table 6 reports linear regressions of the post-election agency score on a treatment indicator and its interaction with the pre-treatment score and other relevant controls, e.g., a baseline index of factual political knowledge. Unsurprisingly, we find that the scores in the *drag & drop* exercise are strongly correlated across waves: the pre-treatment score predicts the score in the post-treatment. Column (1) shows a positive and statistically significant average treatment effect on the post-treatment score: on average, treatment significantly increases the score. Moreover, consistent with our framework, learning effects should be

Table 6: Persistence: Treatment (Direct) Effects on Narrative Post-Election

	$\boldsymbol{Score_{POST}}$			
	(1)	(2)	(3)	(4)
$Score_{PRE}$	0.570***	0.643***	0.644***	0.644***
	(0.031)	(0.040)	(0.039)	(0.039)
Treated = 1	0.129***	0.122***	0.132***	0.128***
	(0.043)	(0.043)	(0.042)	(0.043)
$Score_{PRE}$ × Treated		-0.144**	-0.159**	-0.162**
		(0.062)	(0.062)	(0.063)
Factual knowledge			0.085**	0.067
			(0.033)	(0.042)
Factual knowledge × Treated				0.034
				(0.066)
Constant	0.108***	0.110***	0.095***	0.098***
	(0.029)	(0.028)	(0.029)	(0.029)
Observations	664	664	664	664
R-squared	0.298	0.303	0.309	0.309

Notes: The table reports OLS regressions of the post-election *drag & drop* categorization game score on its pre-election value and the interaction of the pre-election score with the treatment indicator. *Factual knowledge* is the pre-registered index of factual political knowledge measured at baseline. *Treated* indicates assignment to the information treatment at the end of the pre-election survey. Standard errors are clustered at the respondent level and reported in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

concentrated among respondents who did *not* hold an agency view, pre-treatment. This is indeed what we find. Estimating a model with an interaction between the pre-treatment score and the treatment indicator documents a much larger effect—about 0.27 ($t = 3.4$)—for those who scored the lowest, pre-treatment (-1), while predictably, for respondents at the highest pre-treatment score (1), the treatment effect is statistically indistinguishable from zero ($t = -0.3$) and negligible in magnitude. Columns (3)–(4) further show that factual political knowledge is a significant, albeit less robust, predictor of the agency score; the treatment effect remains robust to including this control, and there is no evidence of an interaction between treatment and factual knowledge. Taken together, these models provide *direct evidence* that the treatment shifted the narrative among respondents who initially held a deferential view of democracy—and, crucially, that *this shift persisted through the post-election wave.*

We also find persistent effects of the treatment on utility weights when we restrict the analysis to the post-election wave: the pre-election treatment significantly increases the weight placed on valence and ideology in the post-election wave, with magnitudes closely matching the main estimates in Table 2. We relegate the full-sample and post-election estimates to Appendix Figure A.3 due to space limitations.[29]

7.1 Experimenter Demand Effect

A potential concern is that the video conveys a normative cue about how voters "ought" to evaluate candidates, inducing *experimenter demand* (de Quidt *et al.*, 2018; Orne, 1962; Zizzo, 2010). Three patterns mitigate this concern. First, referring to Table 1, we find no discernible effect for Agency voters, contrary to the idea that respondents uniformly move toward perceived researcher expectations. Second, treatment effects are heterogeneous in ways inconsistent with generic compliance: referring to Table 6, response times increase for Deferential voters but decrease for low-knowledge voters, whereas a demand-induced "try harder" channel would predict shifts in the same direction across groups. Third, the treatment effect on the agency score persists in the post-election wave—and only among Deferential voters—which is hard to reconcile with a short-lived, uniform, experimenter demand artifact.

Appendix Table A.14 shows that the placebo video—which does not discuss democratic narratives and electoral accountability, but still deals with the general topic of elections—does not replicate the treatment effects. Recall that the control group includes respondents assigned to (i) no video and (ii) the placebo video. The estimated weights closely track those

[29]To preserve power, we use the broadest sample and retain respondents with missing agency scores.

in Table 1, including the Deferential–Agency heterogeneity. Crucially, unlike the treatment video, which yields large and statistically significant effects in the main tables, the placebo effects are statistically indistinguishable from zero (and occasionally flip sign across sample splits). These findings support an interpretation of shifts in underlying narratives rather than experimenter demand.

8 Implications and Conclusions

Drawing on a large, nationally representative sample of Indonesian voters during the 2024 general election campaign, we provide causal evidence that how individuals conceptualize the relationship between voters and elected officials fundamentally shapes their preferences and their electoral choices. Compared to Agency-minded respondents, individuals who hold a Deferential view of democracy place less weight on the valence traits of candidates, and no weight on ideological disagreement. A randomized intervention, which explains the role of elections as creating a principal-agent relationship, leads to a marked increase in the weight Deferential voters assign to candidate performance and ideological disagreement. We also document how the intervention affects choices in the actual presidential election. Across these two settings, the intervention makes the choices of Deferential voters more similar to those of Agency-minded voters.

The implications of these findings are far-reaching. At the most immediate level, they reveal the presence of frictions in electoral control: voters may not instinctively view elections as opportunities to hold politicians accountable or select public-oriented types, particularly in settings where clientelism is prevalent and elite dominance is normalized. Yet, once prompted to adopt an accountability frame, voters recalibrate their priorities. This indicates that performance-based voting may be latent rather than absent—constrained not by voters' preferences, but by how they have been socialized to understand democratic practice. By shifting voters' mental models, even a brief, non-partisan informational cue can enhance their ability to reward competent politicians and, conversely, penalize underperformance. In standard models of democratic politics, failures of accountability lead to a reduction in voter welfare (Besley, 2006; Ferraz and Finan, 2025). Therefore, interventions that enhance vote choice criteria by promoting an understanding of elections as establishing a principal-agent relationship can have important consequences for governance and development.

Our findings have direct implications for the scalability and cost-effectiveness of civic education efforts, in particular, when they spread a message that is not bound to be considered particularly controversial, at least in public debate. In contrast to more resource-intensive or enforcement-driven institutional reforms, ideational interventions offer a politically feasible

and sustainable approach to improving democratic accountability at scale. This is particularly relevant in developing democracies where institutional constraints and politician self-interest often preclude more ambitious governance reforms. The potential to deliver such interventions through media campaigns, public service announcements, or voter outreach initiatives makes this strategy both practical and impactful.

The governance implications are equally significant. When voters become more attuned to performance, political elites are likely to adjust their strategies. Rather than relying primarily on identity appeals or clientelistic inducements, politicians operating in such environments will face increasing pressure to demonstrate tangible results. In this sense, our findings speak to the endogenous nature of clientelism: it persists not simply because of elite manipulation, but because voters have come to expect and reward it. Changing voter expectations therefore has the potential to shift the political equilibrium itself—encouraging greater programmatic competition, reducing distortionary patronage, and ultimately fostering more effective policy and responsive public institutions.

Finally, this research offers a new lens through which to interpret some of the most consequential events in Indonesia's recent political history. The 2017 blasphemy trial of former Jakarta governor Basuki Tjahaja Purnama—a double minority with a strong performance record—has often been cited as evidence of the primacy of identity in Indonesian politics. Yet our results suggest a more nuanced picture: Purnama's popularity likely stemmed from his performance in office, and it was precisely because identity-based attacks alone could not erode his support that opponents resorted to mobilization, prosecution, and legal exclusion. This interpretation reinforces our broader point: when performance matters, the political stakes are raised, and so too are the tactics used to suppress viable challengers.

In sum, this study highlights the normative dimension of democratic accountability. Voter beliefs about what elections are for—whether they are acts of loyalty, expressions of group identity, or instruments of oversight—profoundly shape political outcomes. By demonstrating that these beliefs are malleable, and that even a short intervention can shift them toward a more accountability-oriented frame, we point to a promising and underutilized lever for democratic reform. In contexts where formal institutions are weak, cultivating the norms and expectations that underpin accountability may be one of the most effective routes to strengthening democratic governance and advancing inclusive development.

References

Abramson, S. F., Koçak, K. and Magazinnik, A. (2022). What do we learn about voter preferences from conjoint experiments? *American Journal of Political Science*, **66** (4),

1008–1020.

Acemoglu, D., Ajzenman, N., Aksoy, C. G., Fiszbein, M. and Molina, C. (2024). (successful) democracies breed their own support. *Review of Economic Studies*, p. rdae051.

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, **115** (3), 715–753.

Alan, S. and Ertac, S. (2018). Fostering patience in the classroom: Results from a randomized educational intervention. *Journal of Political Economy*, **126** (5), 1865–1911.

Alesina, A. and Fuchs-Schündeln, N. (2007). Good-bye lenin (or not?): The effect of communism on people's preferences. *American Economic Review*, **97** (4), 1507–1528.

Allen, N. W. (2015). Clientelism and the personal vote in indonesia. *Electoral Studies*, **37**, 73–85.

Anderson, S., Francois, P. and Kotwal, A. (2015). Clientelism in indian villages. *American Economic Review*, **105** (6), 1780–1816.

Anik, L. and Hauser, R. (2020). One of a kind: The strong and complex preference for unique treatment from romantic partners. *Journal of Experimental Social Psychology*, **86**, 103899.

Ansolabehere, S. and Snyder Jr, J. M. (2000). Valence politics and equilibrium in spatial election models. *Public choice*, **103** (3), 327–336.

Aspinall, E. (2010). Indonesia: The democracy agenda is eclipsed. *Journal of Democracy*, **21** (2), 20–34.

— (2014). When brokers betray: Clientelism, social networks, and electoral politics in indonesia. *Critical Asian Studies*, **46** (4), 545–570.

Auyero, J. (1999). " from the client's point (s) of view": How poor people perceive and evaluate political clientelism. *Theory and society*, **28** (2), 297–334.

— (2001). *Poor people's politics: Peronist survival networks and the legacy of Evita*. Duke University Press.

Banerjee, A., Green, D. P., McManus, J. and Pande, R. (2014). Are poor voters indifferent to whether elected leaders are criminal or corrupt? a vignette experiment in rural india. *Political Communication*, **31** (3), 391–407.

—, Kumar, S., Pande, R. and Su, F. (2011). Do informed voters make better choices? experimental evidence from urban india. *Unpublished manuscript*.

Banerjee, A. V. and Duflo, E. (2011). *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. New York: PublicAffairs, accessed 7 Aug. 2025.

— and Pande, R. (2007). Parochial politics: Ethnic preferences and politician corruption.

Bansak, K., Hainmueller, J., Hopkins, D. J. and Yamamoto, T. (2022). Using conjoint experiments to analyze election outcomes: The essential role of the average marginal component effect. *Political Analysis*, **31** (4), 500 – 518.

Bardhan, P., Mitra, S., Mookherjee, D. and Nath, A. (2024). How do voters respond to welfare vis-à-vis public good programs? theory and evidence of political clientelism. *Quantitative Economics*, **15** (3), 655–697.

— and Mookherjee, D. (2020). Clientelistic politics and economic development. *The Handbook of Economic Development and Institutions, Edited by Jean-Marie Baland, François Bourguignon, Jean-Philippe Platteau, and Thierry Verdier*, pp. 84–102.

Barro, R. J. (1973). The control of politicians: An economic model. *Public Choice*, **14**, 19–42.

Baysan, C. (2022). Persistent polarizing effects of persuasion: Experimental evidence from turkey. *American Economic Review*, **112** (11), 3528–3546.

Becker, G. S. and Mulligan, C. B. (1997). The endogenous determination of time preference. *The Quarterly Journal of Economics*, **112** (3), 729–758.

Bénabou, R. and Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review*, **96** (5), 1652–1678.

Benjamin, D. J., Heffetz, O., Kimball, M. S. and Rees-Jones, A. (2012). What do you think would make you happier? what do you think you would choose? *American Economic Review*, **102** (5), 2083–2110.

—, —, — and Szembrot, N. (2014). Beyond happiness and satisfaction: Toward well-being indices based on stated preference. *American Economic Review*, **104** (9), 2698–2735.

Berenschot, W. (2018). The political economy of clientelism: A comparative study of indonesia's patronage democracy. *Comparative political studies*, **51** (12), 1563–1593.

— and Mulder, P. (2019). Explaining regional variation in local governance: Clientelism and state-dependency in indonesia. *World Development*, **122**, 233–244.

Bernhardt, D., Câmara, O. and Squintani, F. (2011). Competence and ideology. *The Review of Economic Studies*, **78** (2), 487–522.

Besley, T. (2006). *Principled agents?: The political economy of good government*. Oxford University Press, USA.

Bidner, C. and Francois, P. (2013). The emergence of political accountability. *The Quarterly Journal of Economics*, **128** (3), 1397–1448.

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

Bowles, J. and Larreguy, H. (2020). Who debates, who wins? at-scale experimental evidence on the supply of policy information in a liberian election. *American Political Science Review*, pp. 1–20.

Buehler, M. (2016). Islam and democracy in indonesia. *Asian Journal of Political Science*, **24** (1), 1–17.

— (2023). Indonesia's 2024 elections: continuity amid competition. *New Mandala*, online commentary.

Bursztyn, L., González, A. L. and Yanagizawa-Drott, D. (2020). Misperceived social norms: Women working outside the home in saudi arabia. *American Economic Review*, **110** (10), 2997–3029.

Cantoni, D., Chen, Y., Yang, D. Y., Yuchtman, N. and Zhang, Y. J. (2017). Curriculum and ideology. *Journal of Political Economy*, **125** (2), 338–392.

Caruso, E. M., Rahnev, D. A. and Banaji, M. R. (2009). Using conjoint analysis to detect discrimination: Revealing covert preferences from overt choices. *Social Cognition*, **27** (1), 128–137.

Chauchard, S. (2016). Unpacking ethnic preferences: Theory and micro-level evidence from north india. *Comparative Political Studies*, **49** (2), 253–284.

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

CNN (2024). Indonesia general election explainer. Accessed: 2025-10-03.

Colonnelli, E., Neto, V. P. and Teso, E. (2025). Politics at work. *American Economic Review*, **115** (10), 3367–3414.

Cruz, C., Keefer, P. and Labonne, J. (2021). Buying informed voters: New effects of information on voters and candidates. *The Economic Journal*, **131** (635), 1105–1134.

—, —, — and Trebbi, F. (2024). Making policies matter: Voter responses to campaign promises. *The Economic Journal*, **134** (661), 1875–1913.

de Albuquerque, A., Finan, F., Jha, A., Karpuska, L. and Trebbi, F. (2025). *Decoupling Taste-Based versus Statistical Discrimination in Elections*. NBER Working Paper 33859, National Bureau of Economic Research, Cambridge, MA.

de Quidt, J., Haushofer, J. and Roth, C. (2018). Measuring and bounding experimenter demand. *American Economic Review*, **108** (11), 3266–3302.

Economist Intelligence Unit (2025). *Democracy Index 2024: What's wrong with representative democracy?* Tech. rep., The Economist Intelligence Unit.

Fealy, G. (2016). Bigger than ahok: explaining the 2 december mass rally. *Indonesia at Melbourne*, online article.

Ferejohn, J. (1986). Incumbent performance and electoral control. *Public choice*, pp. 5–25.

Fergusson, L., Molina, C. A. and Robinson, J. A. (2022). The weak state trap. *Economica*, **89** (354), 293–331.

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

— and — (2025). *Malfunctioning Democracies: Understanding Accountability Failures in Developing Countries*. Working Paper 34198, National Bureau of Economic Research.

Fong, C. M. and Luttmer, E. F. (2009). What determines giving to hurricane katrina victims? experimental evidence on racial group loyalty. *American Economic Journal: Applied Economics*, **1** (2), 64–87.

— and — (2011). Do fairness and race matter in generosity? evidence from a nationally representative charity experiment. *Journal of Public Economics*, **85** (5-6), 372–394.

Fujiwara, T. and Wantchekon, L. (2013). Can informed public deliberation overcome clientelism? experimental evidence from benin. *American Economic Journal: Applied Economics*, **5** (4), 241–255.

George, S., Gupta, S. and Neggers, Y. (2019). *Can we text criminal politicians out of office? Evidence from a mobile experiment in India*. Tech. rep., Working paper.

Green, P. E. and Rao, V. R. (1971). Conjoint measurement for quantifying judgmental data. *Journal of Marketing*, **8**, 355–63.

Haaijer, R., Kamakura, W. and Wedel, M. (2000). Response latencies in the analysis of conjoint choice experiments. *Journal of Marketing Research*, **37** (3), 376–382.

Hadiwinata, B. S. (2003). *The Politics of NGOs in Indonesia: Developing Democracy and Managing a Movement*. Routledge.

Hainmueller, J., Hangartner, D. and Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences*, **112** (8), 2395–2400.

—, Hopkins, D. J. and Yamamoto, T. (2014). Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. *Political analysis*, **22** (1), 1–30.

Hicken, A., Leider, S., Ravanilla, N. and Yang, D. (2015). Measuring vote-selling: Field evidence from the philippines. *American Economic Review*, **105** (5), 352–356.

—, —, — and — (2018). Temptation in vote-selling: Evidence from a field experiment in the philippines. *Journal of Development Economics*, **131**, 1–14.

Kartik, N. and McAfee, R. P. (2007). Signaling character in electoral competition. *American Economic Review*, **97** (3), 852–870.

Kawai, K. and Sunada, T. (2025). Estimating candidate valence. *Econometrica*, **93** (2), 463–501.

Keefer, P. and Vlaicu, R. (2008). Democracy, credibility, and clientelism. *The Journal of Law, Economics, & Organization*, **24** (2), 371–406.

Kendall, C., Nannicini, T. and Trebbi, F. (2015). How do voters respond to information? evidence from a randomized campaign. *American Economic Review*, **105** (1), 322–353.

Kessler, J. B., Low, C. and Sullivan, C. D. (2019). Incentivized resume rating: Eliciting employer preferences without deception. *American Economic Review*, **109** (11), 3713–44.

Kitschelt, H. (2000). Linkages between citizens and politicians in democratic polities. *Comparative political studies*, **33** (6-7), 845–879.

— and Wilkinson, S. I. (2007). Citizen-politician linkages: An introduction. In H. Kitschelt and S. I. Wilkinson (eds.), *Patrons, Clients and Policies: Patterns of Democratic Accountability and Political Competition*, Cambridge: Cambridge University Press, pp. 1–49.

Klašnja, M., Little, A. T. and Tucker, J. A. (2018). Political corruption traps. *Political Science Research and Methods*, **6** (3), 413–428.

Komarudin, U. and Pramuji, P. (2023). Between clientelism and patrimonialism: Local politics of the philippines and indonesia. *Jurnal Wacana Politik Volume*, **8** (1).

Kuziemko, I., Norton, M. I., Saez, E. and Stantcheva, S. (2015). How elastic are preferences for redistribution? evidence from randomized survey experiments. *American Economic Review*, **105** (4), 1478–1508.

Landa, D. and Duell, D. (2015). Social identity and electoral accountability. *American Journal of Political Science*, **59** (3), 671–689.

Lanoue, D. J. and Bowler, S. (1998). Picking the winners: perceptions of party viability and their impact on voting behavior. *Social Science Quarterly*, pp. 361–377.

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

Lewandowski, P., Lipowska, K. and Smoter, M. (2026). Mismatch in preferences for working from home: Evidence from discrete choice experiments with workers and employers. *ILR Review*, **79** (1), 142–172.

Londregan, J. and Romer, T. (1993). Polarization, incumbency, and the personal vote. *Political economy: Institutions, competition, and representation*, pp. 355–377.

López-Moctezuma, G., Wantchekon, L., Rubenson, D., Fujiwara, T. and Pe Lero, C. (2022). Policy deliberation and voter persuasion: Experimental evidence from an election in the philippines. *American Journal of Political Science*, **66** (1), 59–74.

Louviere, J. J., Flynn, T. N. and Carson, R. T. (2010). Discrete choice experiments are not conjoint analysis. *Journal of choice modelling*, **3** (3), 57–72.

Low, C. (2024). Pricing the biological clock: The marriage market costs of aging to women. *Journal of Labor Economics*, **42** (2), 395–426.

Macchi, E. (2023). Worth your weight: Experimental evidence on the benefits of obesity in low-income countries. *American Economic Review*, **113** (9), 2287–2322.

Maestas, N., Mullen, K. J., Powell, D., von Wachter, T. and Wenger, J. B. (2023). The value of working conditions in the united states and the implications for the structure of wages. *American Economic Review*, **113** (7), 2007–2047.

Malmendier, U. and Nagel, S. (2011). Depression babies: Do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics*, **126** (1), 373–416.

Mattmann, M., Logar, I. and Brouwer, R. (2019). Choice certainty, consistency, and monotonicity in discrete choice experiments. *Journal of Environmental Economics and Policy*, **8** (2), 109–127.

Mietzner, M. (2015). Reinventing asian democracy: The democratic turn in indonesia. *Journal of Democracy*, **26** (2), 110–123.

— (2020). Populist anti-scientism, religious conservatism, and pandemic management in indonesia. *East Asia Forum Quarterly*, **12** (2), 4–8.

— (2024). Jokowi's pyrrhic victory: Indonesia's 2024 elections and the political reinvention of prabowo subianto. *Contemporary Southeast Asia: A Journal of International and Strategic Affairs*, **46** (2), 187–215.

— and Muhtadi, B. (2019). The mobilisation of intolerance and its trajectories: Indonesian muslims' views of religious minorities and ethnic chinese. In *Contentious Belonging: The Place of Minorities in Indonesia*, ISEAS Yusof Ishak Institute.

Mogilski, J. K., Wade, T. J. and Welling, L. L. M. (2014). Prioritization of potential mates' history of sexual fidelity during a conjoint ranking task. *Personality and Social Psychology Bulletin*, **40** (7), 884–897.

Moskowitz, H. R. and Silcher, M. (2006). The applications of conjoint analysis and their possible uses in sensometrics. *Food Quality and Preference*, **17** (3-4), 145–165.

Mutz, D. C. (1998). *Impersonal influence: How perceptions of mass collectives affect political attitudes*. Cambridge University Press.

Neuman, E. and Neuman, S. (2008). Reference-dependent preferences and loss aversion: A discrete choice experiment in the health-care sector. *Judgment and Decision Making*, **3** (2), 162–173.

Nord, M., Altman, D., Angiolillo, F., Fernandes, T., Good God, A. and Lindberg, S. I. (2025). *Democracy Report 2025: 25 Years of Autocratization – Democracy Trumped?* Tech. rep., V-Dem Institute, University of Gothenburg, Gothenburg.

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, **17** (11), 776–783.

Pande, R. (2011). Can informed voters enforce better governance? experiments in low-income democracies. *Annu. Rev. Econ.*, **3** (1), 215–237.

Pepinsky, T. (2019). Religion, ethnicity, and indonesia's 2019 presidential election. *New Mandala*.

Pepinsky, T. B., Liddle, R. W. and Mujani, S. (2012). Testing islam's political advantage: Evidence from indonesia. *American Journal of Political Science*, **56** (3), 584–600.

Platas, M. R. and Raffler, P. J. (2021). Closing the gap: Information and mass support in a dominant party regime. *The Journal of Politics*, **83** (4), 1619–1634.

Przeworski, A., Stokes, S. C. and Manin, B. (1999). *Democracy, accountability, and representation*, vol. 2. Cambridge University Press.

Robinson, G. (2018). *The Killing Season: A History of the Indonesian Massacres, 1965-66*. Princeton University Press.

Robinson, J. A. and Verdier, T. (2013). The political economy of clientelism. *The Scandinavian Journal of Economics*, **115** (2), 260–291.

Rothschild, D. and Malhotra, N. (2014). Are public opinion polls self-fulfilling prophecies? *Research & Politics*, **1** (2), 2053168014547667.

Shiller, R. J. (2017). Narrative economics. *American Economic Review*, **107** (4), 967–1004.

Shofia, N. and Pepinsky, T. (2019). Measuring the 'nu effect' in indonesia's election. *New Mandala*.

Soedirgo, J. (2018). Informal networks and religious intolerance: how clientelism incentivizes the discrimination of the ahmadiyah in indonesia. *Citizenship Studies*, **22** (2), 191–207.

Stokes, S. C. (2005). Perverse accountability: A formal model of machine politics with evidence from argentina. *American political science review*, **99** (3), 315–325.

— (2011). Political clientelism. In *The Oxford Handbook of Political Science*, Oxford University Press.

Suryanarayan, P. (2019). When do the poor vote for the right wing and why: Status hierarchy and vote choice in the indian states. *Comparative Political Studies*, **52** (2), 209–245.

Svolik, M. W. (2019). Polarization versus democracy. *Journal of democracy*, **30** (3), 20–32.

Toha, R. J., Gueorguiev, D. D. and Sinpeng, A. (2021). The normalization of intolerance: The 2019 presidential election in indonesia. *Electoral Studies*, **74**, 102391.

Vaishnav, M. (2017). *When crime pays: Money and muscle in Indian politics*. Yale University Press.

Vicente, P. C. (2014). Is vote buying effective? evidence from a field experiment in west africa. *The Economic Journal*, **124** (574), F356–F387.

— and Wantchekon, L. (2008). Clientelism and vote buying: Lessons from field experiments in west africa. *Oxford Review of Economic Policy*.

Wade, T. J. and Mogilski, J. K. (2018). Emotional accessibility is more important than sexual accessibility in evaluating romantic relationships – especially for women: A conjoint analysis. *Frontiers in Psychology*, **9**, 632.

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

Warburton, E., Muhtadi, B., Aspinall, E. and Fossati, D. (2021). When does class matter? unequal representation in indonesian legislatures. *Third World Quarterly*, **42** (6), 1252–1275.

Wiswall, M. and Zafar, B. (2018). Preference for the workplace, investment in human capital, and gender. *The Quarterly Journal of Economics*, **133** (1), 457–507.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, **13** (1), 75–98.

A Auxiliary Tables and Figures

Table A.1: Respondent Descriptives Across Waves

Variable	Baseline	Pre	Post	Main Sample
Age category				
17 - 19	11.7	8.2	7.3	8.3
20 - 24	24.9	23.2	23.4	23.6
25 - 29	21.9	22.2	22.0	21.2
30 - 34	16.3	17.0	17.2	17.6
35 - 39	11.2	13.0	13.3	13.4
40 - 44	7.3	8.7	9.1	8.3
45 - 49	6.7	7.6	7.9	7.7
Gender category				
Female	58.3	54.7	54.0	55.0
Male	41.7	45.3	46.0	45.0
Ethnic category				
Chinese Non-Titular	3.1	3.4	3.6	3.9
Non-Titular Ethnic	34.7	33.6	33.1	29.8
Titular Ethnic	62.3	63.0	63.4	66.3
Religious category				
Non-muslim	13.2	13.1	13.3	13.2
Muslim	86.8	86.9	86.7	86.8
Education category				
HS	57.2	56.0	55.4	55.8
University	37.2	39.1	39.5	39.9
Less than HS	5.2	4.6	4.8	4.0
Socio-economic status category				
Lower 1	10.7	10.9	10.8	10.6
Lower 2	0.6	0.5	0.5	0.4
Middle 1	38.9	38.8	38.0	38.5
Middle 2	21.1	20.7	20.9	20.5
Upper 1	7.7	8.2	8.4	8.3
Upper 2	21.0	20.9	21.5	21.7
Geographic location				
DKI Jakarta	9.0	9.3	9.1	9.7
Jawa Barat	20.7	21.0	20.7	20.7
Jawa Tengah	12.3	13.2	13.0	14.6
Jawa Timur	14.8	14.9	15.5	15.8
Sumatera Utara	4.3	4.8	5.1	5.4
Other	38.8	36.8	36.7	33.8
Num. of Obs	8664	5385	4840	3155

Note: The table presents the fraction of respondents' characteristics in different waves; the last column presents the statistics of respondents in the main analysis (candidate-choice module).

Table A.2: Summary Statistics

Variable	Obs	Mean	Std. dev.	Min	Max
A. Random utility model - Aggregated traits					
$Voted_1 = 1$	52,136	0.495	0.500	0	1
Ideology	52,136	-0.003	0.471	-1	1
Valence	52,136	-0.004	0.706	-1	1
Identity	52,136	0.004	0.628	-1	1
Treated = 1	52,136	0.498	0.500	0	1
B. Random utility model - Disaggregated traits					
Diff in ethnicity	52,136	0.004	0.692	-1	1
Diff in religion	52,136	0.001	0.604	-1	1
Diff in ideology: Army role	52,136	0.001	0.202	-1	1
Diff in ideology: Morality	52,136	-0.002	0.357	-1	1
Diff in ideology: Freedom to organise	52,136	-0.000	0.229	-1	1
Diff in ideology: Free speech	52,136	-0.002	0.201	-1	1
Diff in ideology: Immigration	52,136	-0.001	0.208	-1	1
Diff in valence: Corruption	52,136	-0.001	0.313	-1	1
Diff in valence: Economy	52,136	-0.001	0.317	-1	1
Diff in valence: Health	52,136	-0.001	0.317	-1	1
Diff in valence: Women & children	52,136	-0.000	0.316	-1	1
Diff in valence: Environment	52,136	-0.001	0.316	-1	1
C. Vote analysis: Deferential voters					
Actual vote: Front-runner	1,242	0.604	0.489	0	1
Actual vote: Underdog	1,242	0.108	0.310	0	1
Actual vote: Contender	1,242	0.232	0.422	0	1
Actual vote: No vote	1,242	0.056	0.231	0	1
Treated = 1	1,242	0.525	0.500	0	1
D. Response time analysis					
Log(Response time)	52,134	9.436	0.908	6.279	17.587
Low factual knowledge	52,134	0.580	0.494	0	1
Deferential	52,134	0.480	0.500	0	1
Treated = 1	52,134	0.498	0.500	0	1
E. Persistence analysis					
$Score_{POST}$	664	0.147	0.665	-1	1
$Score_{PRE}$	664	-0.051	0.632	-1	1
Treated = 1	664	0.523	0.500	0	1
Factual knowledge index	664	0.102	0.638	-1.569	0.752

Continued on next page

Variable	Obs	Mean	Std. dev.	Min	Max
(Continued)					
F. Non-parametric AMCE model - Aggregated traits					
$Voted = 1$	104,272	0.500	0.500	0	1
Share Identity	104,272	0.299	0.458	0	1
High Valence	104,272	0.498	0.500	0	1
Ideology Same Side	104,272	0.454	0.498	0	1
Ideology Neutral	104,272	0.092	0.289	0	1
Treated = 1	104,272	0.498	0.500	0	1
G. Non-parametric AMCE model - Disaggregated traits					
Same religion	104,272	0.643	0.479	0	1
Same ethnicity	104,272	0.455	0.498	0	1
Agree: Army role	104,272	0.063	0.242	0	1
Agree: Morality	104,272	0.198	0.399	0	1
Agree: Freedom to organise	104,272	0.067	0.250	0	1
Agree: Freedom of speech	104,272	0.063	0.243	0	1
Agree: Immigration	104,272	0.063	0.243	0	1
Disagree: Army role	104,272	0.062	0.241	0	1
Disagree: Morality	104,272	0.199	0.400	0	1
Disagree: Freedom to organise	104,272	0.067	0.251	0	1
Disagree: Freedom of speech	104,272	0.061	0.238	0	1
Disagree: Immigration	104,272	0.064	0.245	0	1
Corruption High	104,272	0.098	0.298	0	1
Corruption Low	104,272	0.099	0.298	0	1
Economy High	104,272	0.101	0.301	0	1
Economy Low	104,272	0.100	0.300	0	1
Education High	104,272	0.101	0.301	0	1
Education Low	104,272	0.101	0.301	0	1
Environment High	104,272	0.100	0.300	0	1
Environment Low	104,272	0.100	0.301	0	1
Women & Children High	104,272	0.098	0.298	0	1
Women & Children Low	104,272	0.102	0.302	0	1

Note: The table reports summary statistics for the variables used in the main analysis. Panels B and G disaggregate the main explanatory traits reported in Panels A (random utility model) and F (non-parametric approach), respectively. The unit in Panels A and B is respondent-candidate pairs, while in Panels F and G it is respondent-candidate profiles. Detailed variable construction is provided in Section 4 and Appendix Section C. In Panel D, two observations are dropped due to invalid response time; the correlation between *Low Factual Knowledge* and $Deferential$ is 0.08. Panel E only includes respondents for whom the Agency score is measured in both pre-election and post-election waves.

Table A.3: Evidence on Selective Attrition

	(1) In post wave	(2) In post wave	(3) In vote analysis	(4) In vote analysis	(5) In vote analysis
Treated = 1	0.006	0.069	0.020	0.001	0.148
	(0.018)	(0.067)	(0.032)	(0.035)	(0.096)
Deferential = 1	-0.009			-0.002	
	(0.019)			(0.021)	
Treated× Deferential	0.025			0.038	
	(0.026)			(0.030)	
intent: Underdog			0.054	0.008	
			(0.038)	(0.038)	
intent: Front-runner			-0.021	-0.070**	
			(0.026)	(0.028)	
intent: no vote			0.074	0.074	
			(0.073)	(0.065)	
intent: Don't know			0.005	-0.033	
			(0.032)	(0.034)	
Treated× intent: Underdog			-0.034	0.009	
			(0.054)	(0.053)	
Treated× intent: Front-runner			0.003	0.009	
			(0.038)	(0.039)	
Treated× intent: no vote			-0.015	-0.081	
			(0.102)	(0.094)	
Treated× intent: Don't know			-0.017	0.005	
			(0.045)	(0.047)	
SES = Lower 2		-0.053			-0.025
		(0.131)			(0.144)
SES = Middle 1		-0.002			0.044
		(0.026)			(0.034)
SES = Middle 2		0.021			0.059
		(0.028)			(0.037)
SES = Upper 1		0.011			0.093*
		(0.036)			(0.048)
SES = Upper 2		0.005			0.078**
		(0.029)			(0.038)
Treated× SES = Lower 2		0.205			0.003
		(0.134)			(0.221)
Treated× SES = Middle 1		-0.045			-0.066
		(0.036)			(0.049)
Treated× SES = Middle 2		-0.053			-0.062
		(0.039)			(0.053)
Treated× SES = Upper 1		-0.036			-0.114*
		(0.049)			(0.068)
Treated× SES = Upper 2		-0.024			-0.044
		(0.040)			(0.054)
Male		0.048***			0.025
		(0.016)			(0.021)
Treated × Male		-0.008			-0.008
		(0.022)			(0.030)
Muslim = 1		-0.012			0.016
		(0.026)			(0.035)
Treated× Muslim		0.029			0.067
		(0.036)			(0.049)
Ethnic = Non-Titular Ethnic		-0.025			-0.076
		(0.044)			(0.063)
Ethnic = Titular Ethnic		-0.018			-0.055
		(0.044)			(0.062)
Treated× Non-Titular Ethnic		-0.038			-0.098
		(0.059)			(0.088)
Treated× Titular Ethnic		-0.040			-0.096
		(0.058)			(0.087)
Edu. Cat. = University and above		-0.001			0.003
		(0.017)			(0.022)
Edu. Cat. = Other/Missing		-0.018			-0.039
		(0.053)			(0.069)
Treated× Edu. Cat. = University and above		0.021			-0.027
		(0.023)			(0.030)
Treated× Edu. Cat. = Other/Missing		-0.068			0.096
		(0.083)			(0.101)

Continued on next page

Table A.3: Evidence on Selective Attrition (Continued)

	In post wave	In post wave	In vote analysis	In vote analysis	In vote analysis
Continued					
Age Category = 17		-0.161***			-0.061
		(0.038)			(0.043)
Age Category = 20		-0.021			0.004
		(0.025)			(0.033)
Age Category = 25		-0.019			-0.017
		(0.025)			(0.032)
Age Category = 35		0.034			0.042
		(0.026)			(0.037)
Age Category = 40		0.053*			0.035
		(0.027)			(0.041)
Age Category = 45		0.031			0.054
		(0.031)			(0.045)
Treated× Age Category = 17		0.044			-0.036
		(0.052)			(0.061)
Treated× Age Category = 20		-0.003			-0.033
		(0.034)			(0.046)
Treated× Age Category = 25		-0.016			-0.013
		(0.034)			(0.046)
Treated× Age Category = 35		-0.012			-0.024
		(0.036)			(0.052)
Treated× Age Category = 40		-0.021			-0.086
		(0.039)			(0.060)
Treated× Age Category = 45		-0.017			-0.055
		(0.043)			(0.062)
Constant	0.825***	0.837***	0.506***	0.780***	0.482***
	(0.013)	(0.050)	(0.022)	(0.024)	(0.068)
Observations	3,374	4,954	4,966	3,373	4,954
R-squared	0.001	0.025	0.002	0.007	0.011

Note: The first two columns report regressions of the probability of being included in the post-election wave, conditional on participation in the pre-election wave. The last three columns report regressions of the probability of being included in the vote-outcome analysis, again conditional on participation in the pre-election wave. The samples in models 1 and 4 are restricted to respondents who received one of the experimental treatments and completed the *drag & drop* task, as these variables are required for both analyses. The dependent variables are indicators for inclusion in the post-election wave and in the vote-outcome analysis, respectively. The baseline for the predictor categories are: Untreated, Agency, Intention to vote for the contender, Lowest SES category, Female, Non-Muslim, Chinese Non-titular ethnicity, No University, and Age category 30. *** p<0.01, ** p<0.05, * p<0.1.

Table A.4: Balance Test

	$Treated = 1$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Male = 1	-0.010						-0.009
	(0.023)						(0.024)
Respondent age category							
17 - 19		-0.029					-0.026
		(0.046)					(0.047)
20 - 24		-0.028					-0.026
		(0.036)					(0.037)
25 - 29		-0.059					-0.060
		(0.038)					(0.038)
35 - 39		-0.013					-0.015
		(0.042)					(0.042)
40 - 44		-0.097					-0.099
		(0.050)					(0.050)
45 and above		0.004					-0.003
		(0.052)					(0.053)
Respondent education category							
Other/Missing			-0.185				-0.190
			(0.173)				(0.172)
University			-0.011				-0.020
			(0.020)				(0.021)
Less than HS			-0.046				-0.042
			(0.048)				(0.048)
Respondent Socioeconomic Status Category							
Lower 2				-0.064			-0.069
				(0.158)			(0.159)
Middle 1				0.017			0.013
				(0.037)			(0.038)
Middle 2				0.007			0.005
				(0.040)			(0.041)
Upper 1				0.073			0.073
				(0.058)			(0.061)
Upper 2				0.050			0.046
				(0.042)			(0.044)
Respondent Ethnicity							
Native Non-Titular					-0.022		-0.008
					(0.074)		(0.081)
Titular Ethnic					-0.032		-0.014
					(0.072)		(0.080)
Muslim = 1						-0.003	0.012
						(0.038)	(0.043)
Constant	0.502***	0.528***	0.502***	0.477***	0.525***	0.499***	0.522***
	(0.016)	(0.028)	(0.013)	(0.032)	(0.070)	(0.036)	(0.084)
Observations	3,155	3,155	3,155	3,155	3,155	3,155	3,155
R-squared	0.000	0.003	0.001	0.002	0.000	0.000	0.006

Note: The table presents the estimates of an OLS regression of the treatment status on dummies for the values of baseline characteristics of respondents. The base for each category (not shown) is: Female, Age group 30-34, High-school educated, Lowest SES category, Chinese non-titular ethnicity, and Non-Muslim. *** p<0.01, ** p<0.05, * p<0.1.

Table A.5: Predictors of Agency Score

	Score			
	(1)	(2)	(3)	(4)
Hierarchical	-0.229***			
	(0.055)			
Tolerant of selfish		-0.376***		
		(0.067)		
Factual knowledge				0.074**
				(0.030)
Education: Middle school			-0.090	-0.103
			(0.126)	(0.129)
Education: High school			0.132	0.094
			(0.106)	(0.111)
Education: Diploma			0.246**	0.208*
			(0.115)	(0.121)
Education: Bachelor			0.378***	0.328***
			(0.108)	(0.113)
Education: Post graduate			0.565***	0.530***
			(0.136)	(0.141)
Education: Others			-0.225	-0.243*
			(0.142)	(0.147)
Constant	-0.109***	-0.120***	-0.286***	-0.237*
	(0.020)	(0.019)	(0.104)	(0.124)
Other controls	✓	✓	✓	✓
Observations	1,668	1,670	1,752	1,752
R-squared	0.012	0.009	0.029	0.047

Notes: The table reports OLS estimates of the score obtained in the *drag & drop* categorization game on various predictors using control arms only. *Hierarchical* is defined as those respondents who respond "No, because society must obey political leaders" and "No, because political leaders have higher power than society" to the question of whether (and why) voters can or cannot punish politicians with their vote. *Tolerant of selfish* is defined as those respondents who think that politicians can pursue their own self-interest. The base category for education is "Primary." *Factual knowledge* refers to the index of factual knowledge in politics. Other demographic controls not shown are age-group dummies and gender, all of which are not significantly different from zero. All predictors are measured at baseline. Survey weights are applied to ensure representativeness of Indonesia's 2024 vote-eligible population. *** p<0.01, ** p<0.05, * p<0.1.

Table A.6: Benchmark vs. Realized Distribution of Profile Characteristics

Variable	Benchmark	Realized distribution		
		Freq	Percent	Cum
Ethnic group category				
Chinese Non-Titular	10	11,170	10.71	10.71
Native Non-Titular	40	41,219	39.53	50.24
Titular	50	51,883	49.76	100.00
Religion Category				
Non-muslim	30	28,466	27.30	27.30
Muslim	70	75,806	72.70	100.00
Gender Category				
Male	60	62,663	60.10	60.10
Female	40	41,609	39.90	100.00
Age Category				
23 - 55	80	84,365	80.91	80.91
56 - 69	15	14,557	13.96	94.87
70 - 75	5	5,350	5.13	100.00
Education Category				
Primary school/Junior high school	5	5,082	4.87	4.87
Senior high school	50	52,041	49.91	54.78
D1/D2/D3	5	5,208	4.99	59.78
D4/S1	30	31,372	30.09	89.86
S2/S3	10	10,569	10.14	100.00
Marital status				
Not Married	20	20,874	20.02	20.02
Married	75	78,251	75.05	95.06
Divorced/widowed	5	5,147	4.94	100.00
Job group Category				
Stay home	10	10,713	10.27	10.27
Enterpreneur	10	10,440	10.01	20.29
Politician	5	5,203	4.99	25.28
Retiree	5	5,176	4.96	30.24
Student	5	5,135	4.92	35.16
Teacher	5	5,181	4.97	40.13
Worker	50	51,912	49.79	89.92
Others	10	10,512	10.08	100.00
Total Profile		104,272	100.00	100.00

Note: The table compares the distribution of candidate profile characteristics in the benchmark design specified in the pre-analysis plan with the empirical distribution actually shown to respondents. The benchmark distribution was determined ex ante, based on the observed characteristics of legislative candidates who contested in Indonesia's 2019 general election. It serves as a reference for assessing how closely the realized profiles align with the intended design.

Table A.7: Random Utility Model: Weight Assigned to Traits–Disaggregated

	$P(voted_1) = 1$		
	Combined	Deferential	Agency
Diff in ethnicity	0.051***	0.051**	0.048**
	(0.016)	(0.022)	(0.023)
	[0.017]	[0.018]	[0.016]
Diff in religion	0.287***	0.326***	0.243***
	(0.022)	(0.031)	(0.031)
	[0.098]	[0.113]	[0.081]
Diff in ideology: Army role	-0.116**	-0.140*	-0.089
	(0.058)	(0.082)	(0.082)
	[-0.039]	[-0.049]	[-0.029]
Diff in ideology: Morality	-0.102***	-0.023	-0.197***
	(0.034)	(0.044)	(0.053)
	[-0.035]	[-0.008]	[-0.065]
Diff in ideology: Freedom to organise	-0.120**	-0.092	-0.160**
	(0.047)	(0.066)	(0.068)
	[-0.041]	[-0.032]	[-0.053]
Diff in ideology: Free speech	-0.061	0.046	-0.187**
	(0.058)	(0.076)	(0.086)
	[-0.021]	[0.016]	[-0.062]
Diff in ideology: Immigration	-0.046	0.062	-0.162**
	(0.052)	(0.071)	(0.076)
	[-0.016]	[0.022]	[-0.054]
Diff in valence: Corruption	1.128***	0.994***	1.290***
	(0.046)	(0.060)	(0.070)
	[0.385]	[0.346]	[0.427]
Diff in valence: Economy	0.876***	0.824***	0.934***
	(0.044)	(0.061)	(0.064)
	[0.299]	[0.286]	[0.309]
Diff in valence: Educ & Health	0.776***	0.706***	0.861***
	(0.041)	(0.057)	(0.057)
	[0.265]	[0.245]	[0.285]
Diff in valence: Women & Children	0.544***	0.545***	0.547***
	(0.037)	(0.050)	(0.056)
	[0.185]	[0.189]	[0.181]
Diff in valence: Environment	0.550***	0.483***	0.627***
	(0.036)	(0.051)	(0.051)
	[0.187]	[0.168]	[0.208]
Constant	-0.020	-0.027	-0.012
	(0.013)	(0.018)	(0.019)
Observations	26,378	12,248	14,130

Notes: The table reports Probit estimates of the weight assigned to candidates' full-dimensional traits corresponding to the model in Equation 2 using the control arm only. The outcome is an indicator of choosing Candidate 1 (left profile in the pair). *Deferential* includes voters with scores < 0 in the *drag & drop* game; *Agency* includes those with scores ≥ 0; *Combined* pools both types. Numbers in round brackets signify the standard errors of the Probit estimates clustered at the individual respondent level, while numbers in square brackets signify the marginal effect of the Probit estimates. Survey weights are applied to ensure representativeness of Indonesia's 2024 vote-eligible population. *** p<0.01, ** p<0.05, * p<0.1.

Table A.8: Treatment Effects (Unweighted)

	$P(voted_1 = 1)$			
Traits & Treatment	Combined	Deferential	Agency	
Identity × Treated	-0.008	-0.018	0.003	
	(0.021)	(0.030)	(0.030)	
Valence × Treated	0.050*	0.091**	0.012	
	(0.027)	(0.038)	(0.038)	
Ideology × Treated	-0.050*	-0.093**	-0.014	
	(0.026)	(0.038)	(0.037)	
				Deferential (treated)
Identity	0.165***	0.173***	0.158***	0.155***
	(0.015)	(0.022)	(0.020)	(0.0209)
Valence	0.806***	0.745***	0.863***	0.836***
	(0.019)	(0.027)	(0.027)	(0.027)
Ideology	-0.098***	-0.039	-0.151***	-0.131***
	(0.018)	(0.026)	(0.025)	(0.0269)
Treated = 1	-0.002	0.005	-0.009	
	(0.014)	(0.020)	(0.019)	
Constant	-0.010	-0.009	-0.011	
	(0.009)	(0.014)	(0.013)	
Observations	52,136	25,020	27,116	

Notes: The table reports treatment effects on the relative importance (weights) assigned to candidate traits. *Deferential* includes voters with Agency scores < 0; *Agency* includes voters with scores ≥ 0; *Combined* pools both types. The last column reports the "total effect" of each trait for *Deferential* voters under treatment. Robust standard errors clustered at the respondent level are in parentheses. *Identity* aggregates religion and ethnicity; *Valence* captures the candidate's track record on a key governance issue; *Ideology* is the respondent–candidate distance on a key ideological dimension. Survey weights are not applied to regressions in this table. *** p<0.01, ** p<0.05, * p<0.1.

Table A.9: Treatment Effects: The "Horse Race" Model–Non-binary

	$P(voted_1 = 1)$				
Alternative Source of Heterogeneity	Unconditional	Education	Knowldg(std)	Interest(std)	SES(std)
Identity	0.157***	0.155***	0.158***	0.156***	0.164***
	(0.019)	(0.022)	(0.019)	(0.019)	(0.019)
Treated	0.007	0.003	0.008	0.013	0.006
	(0.019)	(0.022)	(0.019)	(0.019)	(0.019)
Treated×Identity	-0.003	-0.005	-0.004	-0.006	-0.010
	(0.027)	(0.031)	(0.027)	(0.028)	(0.028)
Agency Score	0.004	0.001	0.004	0.008	0.003
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Identity×Agency Score	-0.010	-0.013	-0.008	-0.011	-0.011
	(0.019)	(0.019)	(0.019)	(0.019)	(0.019)
Treated×Agency Score	0.010	0.013	0.008	0.005	0.010
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Treated×Identity×Agency Score	0.015	0.016	0.012	0.021	0.016
	(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Valence	0.776***	0.781***	0.776***	0.788***	0.782***
	(0.025)	(0.028)	(0.025)	(0.025)	(0.025)
Treated×Valence	0.050	0.038	0.056	0.049	0.042
	(0.035)	(0.040)	(0.035)	(0.036)	(0.036)
Valence×Agency Score	0.114***	0.108***	0.109***	0.108***	0.113***
	(0.025)	(0.025)	(0.024)	(0.025)	(0.025)
Treated×Valence×Agency Score	-0.073**	-0.072**	-0.086**	-0.076**	-0.073**
	(0.035)	(0.035)	(0.034)	(0.035)	(0.035)
Ideology	-0.107***	-0.103***	-0.106***	-0.114***	-0.110***
	(0.023)	(0.026)	(0.023)	(0.024)	(0.023)
Treated×Ideology	-0.088***	-0.110***	-0.088***	-0.096***	-0.081**
	(0.034)	(0.038)	(0.033)	(0.035)	(0.034)
Ideology×Agency Score	-0.084***	-0.085***	-0.084***	-0.086***	-0.082***
	(0.023)	(0.024)	(0.023)	(0.023)	(0.023)
Treated×Ideology×Agency Score	0.025	0.028	0.029	0.021	0.023
	(0.033)	(0.034)	(0.033)	(0.034)	(0.033)
2.Edu Cat.		-0.318			
		(0.225)			
3.Edu Cat.		0.017			
		(0.022)			
4.Edu Cat.		-0.044			
		(0.057)			
2.Edu Cat.×Identity		-0.130			
		(0.213)			
3.Edu Cat.×Identity		0.039			
		(0.034)			
4.Edu Cat.×Identity		-0.019			
		(0.073)			
Treated×2.Edu Cat.		0.217			
		(0.271)			
Treated×3.Edu Cat.		-0.013			
		(0.032)			
Treated×4.Edu Cat.		0.063			
		(0.083)			
Treated×2.Edu Cat.×Identity		0.160			
		(0.321)			
Treated×3.Edu Cat.×Identity		0.013			
		(0.048)			
Treated×4.Edu Cat.×Identity		0.001			
		(0.111)			
2.Edu Cat.×Valence		-0.250			
		(0.364)			
3.Edu Cat.×Valence		0.062			
		(0.043)			
4.Edu Cat.×Valence		-0.136			
		(0.098)			
Treated×2.Edu Cat.×Valence		-0.335			
		(0.473)			
Treated×3.Edu Cat.×Valence		0.037			
		(0.062)			
Treated×4.Edu Cat.×Valence		0.139			
		(0.128)			

Continued on next page

Table A.9: Treatment Effects: The "Horse Race" Model–Non-binary (Continued)

	$P(voted_1 = 1)$				
Alternative heterogeneity	Unconditionl	Education	Knowldg(std)	Interest(std)	SES(std)
Continued					
2.Edu Cat.×Ideology		-0.077			
		(0.255)			
3.Edu Cat.×Ideology		0.002			
		(0.041)			
4.Edu Cat.×Ideology		-0.058			
		(0.103)			
Treated×2.Edu Cat.×Ideology		0.071			
		(0.394)			
Treated×3.Edu Cat.×Ideology		0.098*			
		(0.059)			
Treated×4.Edu Cat.×Ideology		0.229			
		(0.152)			
Alternative hetero. dimension			-0.008	0.002	0.016
			(0.015)	(0.013)	(0.013)
Identity×hetero			-0.007	0.025	0.028
			(0.022)	(0.019)	(0.018)
Treated×hetero			0.021	-0.013	0.004
			(0.024)	(0.020)	(0.019)
Treated×Identity×hetero			0.012	-0.054*	-0.028
			(0.032)	(0.028)	(0.027)
Valence×hetero			0.106***	-0.028	0.025
			(0.027)	(0.024)	(0.024)
Treated×Valence×hetero			0.035	-0.016	-0.056
			(0.040)	(0.034)	(0.036)
Ideology×hetero			0.021	0.048**	-0.014
			(0.026)	(0.021)	(0.023)
Treated×Ideology×hetero			-0.045	0.003	0.042
			(0.036)	(0.032)	(0.033)
Constant	-0.020	-0.017	-0.021	-0.026**	-0.017
	(0.013)	(0.015)	(0.013)	(0.013)	(0.013)
Observations	52,136	52,136	52,136	50,971	52,136

Notes: The table reports Probit estimates of the weights assigned to candidate traits interacted with various alternative dimensions of heterogeneity. The main dimension of heterogeneity is the Agency score (leftmost model), while the alternative dimensions of heterogeneity listed are: educational attainment, factual political knowledge, interest in politics, and socioeconomic status. All heterogeneity variables are collected prior to treatment. Survey weights are applied to ensure representativeness of Indonesia's 2024 vote-eligible population. *** p<0.01, ** p<0.05, * p<0.1.

Table A.10: Heterogeneity Among Controls: The "Horse Race" Model

	$P(voted_1 = 1)$				
Alternative Source of Heterogeneity	Unconditional	Education	Knowledge	Interest	SES
Identity	0.144***	0.178***	0.139***	0.174***	0.154***
	(0.028)	(0.029)	(0.035)	(0.036)	(0.033)
Deferential = 1	-0.010	-0.009	-0.011	-0.010	-0.010
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
Deferential × Identity	0.031	0.032	0.027	0.031	0.032
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Valence	0.833***	0.918***	0.956***	0.800***	0.865***
	(0.037)	(0.038)	(0.047)	(0.048)	(0.048)
Deferential × Valence	-0.141***	-0.136***	-0.135***	-0.140***	-0.140***
	(0.050)	(0.050)	(0.049)	(0.050)	(0.050)
Ideology	-0.172***	-0.178***	-0.118***	-0.126***	-0.208***
	(0.034)	(0.036)	(0.044)	(0.041)	(0.040)
Deferential = 1 × Ideology	0.149***	0.148***	0.151***	0.151***	0.148***
	(0.047)	(0.047)	(0.047)	(0.046)	(0.046)
No College = 1		-0.024			
		(0.022)			
No College × Identity		-0.038			
		(0.033)			
No College × Valence		-0.096**			
		(0.042)			
No College × Ideology		0.006			
		(0.040)			
Low Politics Knowldg = 1			0.015		
			(0.027)		
Low Politics Knowldg × Identity			0.011		
			(0.038)		
Low Politics Knowldg × Valence			-0.197***		
			(0.051)		
Low Politics Knowldg × Ideology			-0.083*		
			(0.048)		
Low Politics Interest = 1				0.007	
				(0.027)	
Low Politics Interest × Identity				-0.053	
				(0.038)	
Low Politics Interest × Valence				0.057	
				(0.051)	
Low Politics Interest × Ideology				-0.089*	
				(0.046)	
Low SES = 1					-0.050*
					(0.026)
Low SES × Identity					-0.020
					(0.038)
Low SES × Valence					-0.057
					(0.050)
Low SES × Ideology					0.064
					(0.047)
Constant	-0.016	0.005	-0.025	-0.020	0.012
	(0.018)	(0.019)	(0.025)	(0.025)	(0.023)
Observations	26,378	26,378	26,378	26,378	26,378

Notes: The table reports Probit estimates of the weights assigned to candidate traits interacted with alternative dimensions of heterogeneity. *Deferential, No College, Low Political Knowledge, Low Political Interest, and Low SES* are dummy variables indicating respondents at the lower end of the respective distributions (below median for all variables except *No College*). These dimensions correspond to the Agency score, educational attainment, factual political knowledge, interest in politics, and socioeconomic status. All heterogeneity variables are collected prior to treatment. Survey weights are applied to ensure representativeness of Indonesia's 2024 vote-eligible population. *** p<0.01, ** p<0.05, * p<0.1.

Table A.11: Heterogeneity of Treatment: The "Horse Race" Model

	$P(voted_1 = 1)$				
Alternative Source of Heterogeneity	Unconditional	Education	Knowledge	Interest	SES
Identity	0.143***	0.178***	0.140***	0.173***	0.155***
	(0.028)	(0.030)	(0.035)	(0.036)	(0.033)
Treated = 1	-0.005	-0.016	0.012	-0.007	-0.008
	(0.028)	(0.028)	(0.036)	(0.038)	(0.034)
Treated×Identity	0.012	0.025	0.023	-0.038	-0.001
	(0.040)	(0.043)	(0.050)	(0.051)	(0.050)
Deferential = 1	-0.015	-0.014	-0.016	-0.015	-0.015
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
Deferential×Identity	0.029	0.030	0.025	0.029	0.031
	(0.038)	(0.038)	(0.038)	(0.038)	(0.038)
Treated×Deferential	0.019	0.019	0.023	0.019	0.018
	(0.038)	(0.039)	(0.038)	(0.038)	(0.038)
Treated×Deferential×Identity	-0.032	-0.032	-0.026	-0.032	-0.034
	(0.054)	(0.055)	(0.055)	(0.054)	(0.054)
Valence	0.836***	0.920***	0.958***	0.802***	0.869***
	(0.037)	(0.038)	(0.047)	(0.049)	(0.048)
Treated×Valence	-0.023	-0.006	-0.006	-0.029	-0.090
	(0.052)	(0.054)	(0.067)	(0.068)	(0.067)
Deferential×Valence	-0.143***	-0.138***	-0.137***	-0.142***	-0.142***
	(0.050)	(0.050)	(0.050)	(0.050)	(0.050)
Treated×Deferential×Valence	0.156**	0.158**	0.174**	0.158**	0.156**
	(0.071)	(0.071)	(0.070)	(0.070)	(0.071)
Ideology	-0.171***	-0.176***	-0.115***	-0.128***	-0.209***
	(0.034)	(0.036)	(0.044)	(0.041)	(0.040)
Treated×Ideology	-0.070	0.013	-0.141**	-0.025	0.013
	(0.049)	(0.053)	(0.061)	(0.060)	(0.062)
Deferential×Ideology	0.142***	0.142***	0.144***	0.145***	0.141***
	(0.047)	(0.047)	(0.047)	(0.047)	(0.047)
Treated×Deferential×Ideology	-0.042	-0.037	-0.049	-0.047	-0.043
	(0.067)	(0.068)	(0.067)	(0.067)	(0.067)
No College = 1		-0.022			
		(0.022)			
No College×Identity		-0.040			
		(0.033)			
Treated×No College		0.013			
		(0.031)			
Treated×No College×Identity		-0.014			
		(0.048)			
No College×Valence		-0.096**			
		(0.042)			
Treated×No College×Valence		-0.019			
		(0.061)			
No College×Ideology		0.006			
		(0.040)			
Treated×No College×Ideology		-0.094			
		(0.058)			
Low Politics Knowldg = 1			0.016		
			(0.027)		
Low Politics Knowldg×Identity			0.009		
			(0.038)		
Treated×Low Politics Knowldg			-0.030		
			(0.039)		
Treated×Low Politics Knowldg×Identity			-0.024		
			(0.055)		
Low Politics Knowldg×Valence			-0.195***		
			(0.051)		
Treated×Low Politics Knowldg×Valence			-0.039		
			(0.073)		
Low Politics Knowldg×Ideology			-0.086*		
			(0.048)		
Treated×Low Politics Knowldg×Ideology			0.115*		
			(0.068)		

Continued on next page

Table A.11: Heterogeneity of Treatment: The "Horse Race" Model (Continued)

	$P(voted_1 = 1)$				
Alternative Source of Heterogeneity	Unconditional	Education	Knowledge	Interest	SES
Continued					
Low Politics Interest = 1				0.007	
				(0.027)	
Low Politics Interest×Identity				-0.052	
				(0.038)	
Treated×Low Politics Interest				0.003	
				(0.040)	
Treated×Low Politics Interest×Identity				0.087	
				(0.054)	
Low Politics Interest×Valence				0.058	
				(0.051)	
Treated×Low Politics Interest×Valence				0.008	
				(0.072)	
Low Politics Interest×Ideology				-0.083*	
				(0.047)	
Treated×Low Politics Interest×Ideology				-0.072	
				(0.067)	
Low SES = 1					-0.053**
					(0.026)
Low SES×Identity					-0.023
					(0.038)
Treated×Low SES					0.004
					(0.038)
Treated×Low SES×Identity					0.025
					(0.054)
Low SES×Valence					-0.059
					(0.050)
Treated×Low SES×Valence					0.124*
					(0.071)
Low SES×Ideology					0.067
					(0.047)
Treated×Low SES×Ideology					-0.151**
					(0.068)
Constant	-0.012	0.007	-0.022	-0.017	0.017
	(0.018)	(0.019)	(0.025)	(0.025)	(0.023)
Observations	52,136	52,136	52,136	52,136	52,136

Notes: The table reports Probit estimates of the weights assigned to candidate traits interacted with alternative dimensions of heterogeneity. *Deferential, No College, Low Political Knowledge, Low Political Interest, and Low SES* are dummy variables indicating respondents at the lower end of the respective distributions (below median for all variables except *No College*). These dimensions correspond to the Agency score, educational attainment, factual political knowledge, interest in politics, and socioeconomic status. All heterogeneity variables are collected prior to treatment. Survey weights are applied to ensure representativeness of Indonesia's 2024 vote-eligible population. *** p<0.01, ** p<0.05, * p<0.1.

Table A.12: Regression of Actual Vote on Vote Intention

	Actual votes (self-declared)			
	Front-runner	Underdog	Contender	No vote
intent: Underdog × Deferential	0.053	**-0.168****	0.064	0.051
	(0.082)	(0.083)	(0.064)	(0.039)
intent: Front-runner × Deferential	-0.010	-0.017	0.045	-0.018
	(0.038)	(0.020)	(0.045)	(0.032)
intent: No vote × Deferential	-0.092	-0.021	0.087	0.026
	(0.163)	(0.015)	(0.146)	(0.186)
intent: Don't know × Deferential	**0.147****	**-0.085***	-0.035	-0.027
	(0.072)	(0.047)	(0.068)	(0.049)
Deferential = 1	0.007	0.021	-0.030	0.002
	(0.031)	(0.015)	(0.043)	(0.029)
intent: Underdog	0.204***	0.667***	-0.821***	-0.050***
	(0.056)	(0.057)	(0.040)	(0.017)
intent: Front-runner	0.868***	0.023***	-0.874***	-0.018
	(0.024)	(0.009)	(0.026)	(0.020)
intent: No vote	0.229*	0.000	-0.750***	0.521***
	(0.123)	(.)	(0.097)	(0.134)
intent: Don't know	0.415***	0.171***	-0.649***	0.064*
	(0.049)	(0.034)	(0.046)	(0.034)
base intent: Contender	0.057***	-0.000	0.893***	0.050***
	(0.018)	(.)	(0.025)	(0.017)
Observations	1,262	1,262	1,262	1,262
R-squared	0.514	0.347	0.584	0.138

Notes: The table presents OLS regressions of respondents' actual presidential vote on their stated pre-election vote intention and the voter-type indicator. The analysis is restricted to respondents assigned to the control arm. Actual votes are self-reported in the post-election survey, while vote intentions are recorded in the pre-election wave. *** p<0.01, ** p<0.05, * p<0.1.

Table A.13: Relative Weight of Candidate Traits: Response Time

	Log(time spent on pairs)				
	All	Male	Female	Islamist	Nationalist
same religion	0.055***	0.036**	0.075***	0.051***	0.058***
	(0.011)	(0.015)	(0.016)	(0.015)	(0.016)
same ethnicity	0.006	-0.016	0.030**	0.024*	-0.016
	(0.011)	(0.015)	(0.015)	(0.014)	(0.016)
same valence	0.149***	0.188***	0.112***	0.137***	0.164***
	(0.018)	(0.028)	(0.024)	(0.025)	(0.027)
same ideological stance	0.059***	0.048	0.071**	0.073**	0.042
	(0.022)	(0.033)	(0.031)	(0.029)	(0.035)
same gender	0.018	0.010	0.025	0.020	0.014
	(0.011)	(0.015)	(0.015)	(0.014)	(0.017)
same education	0.006	0.018	-0.007	0.007	0.004
	(0.011)	(0.017)	(0.015)	(0.015)	(0.018)
same marital status	0.004	-0.015	0.022	0.011	-0.006
	(0.011)	(0.015)	(0.017)	(0.015)	(0.017)
same age category	0.012	0.002	0.024	0.007	0.018
	(0.013)	(0.018)	(0.019)	(0.018)	(0.018)
same job	-0.001	-0.023	0.022	-0.001	0.000
	(0.011)	(0.016)	(0.016)	(0.015)	(0.017)
Constant	9.329***	9.421***	9.234***	9.303***	9.363***
	(0.014)	(0.017)	(0.023)	(0.020)	(0.020)
Observations	26,376	14,591	11,785	15,149	11,227
R-squared	0.498	0.468	0.524	0.500	0.495

Notes: The table reports OLS estimates of log time spent on candidate *pairs* regressed on indicators that the two candidates share the same trait value. The estimation sample is the same as Column 1 of Table 1, except for two missing response times. Robust standard errors clustered at the respondent level are in parentheses. *Islamist* denotes respondents who reported an Islamist worldview at baseline; *Nationalist* denotes those reporting a Nationalist–Pancasila worldview. Survey weights are applied to ensure representativeness of Indonesia's 2024 vote-eligible population. p<0.01, ** p<0.05, * p<0.1.

Table A.14: Placebo Video Effect – Control Arms

	$P(voted_1 = 1)$		
	Combined	Deferential	Agency
Identity × Placebo Video	-0.017	0.028	-0.061
	(0.038)	(0.053)	(0.055)
Valence × Placebo Video	-0.066	-0.051	-0.091
	(0.050)	(0.067)	(0.073)
Ideology × Placebo Video	-0.045	-0.036	-0.042
	(0.047)	(0.065)	(0.067)
Identity	0.168***	0.161***	0.174***
	(0.026)	(0.035)	(0.038)
Valence	0.789***	0.716***	0.879***
	(0.035)	(0.047)	(0.051)
Ideology	-0.070**	-0.006	-0.150***
	(0.033)	(0.046)	(0.047)
Placebo Video = 1	0.023	0.008	0.041
	(0.026)	(0.036)	(0.036)
Constant	-0.032*	-0.030	-0.037
	(0.019)	(0.026)	(0.026)
Observations	26,378	12,248	14,130

Notes: The table reports the effects of the placebo video on the relative importance (weights) assigned to candidate traits under the random utility model specification restricted to the control arms. The experiment design includes two different control conditions: Placebo Video and No Video. *Deferential* includes voters with scores < 0 in the *drag & drop* game; *Agency* includes voters with scores ≥ 0; *Combined* pools both types. Robust standard errors clustered at the respondent level are in parentheses. *Identity* aggregates religion and ethnicity; *Valence* captures the candidate's track record on a key governance issue; *Ideology* is the respondent–candidate distance on a key ideological dimension. Survey weights are applied to ensure representativeness of Indonesia's 2024 vote-eligible population. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure A.1: Candidate Pairs in Discrete Choice Exercise

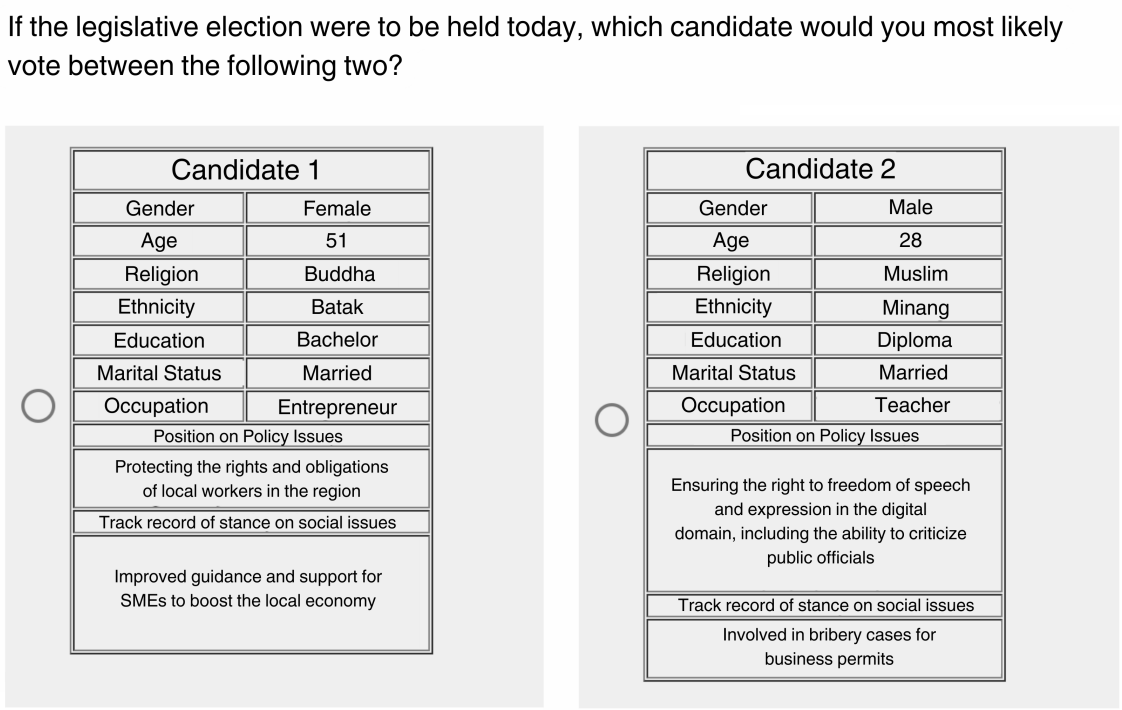
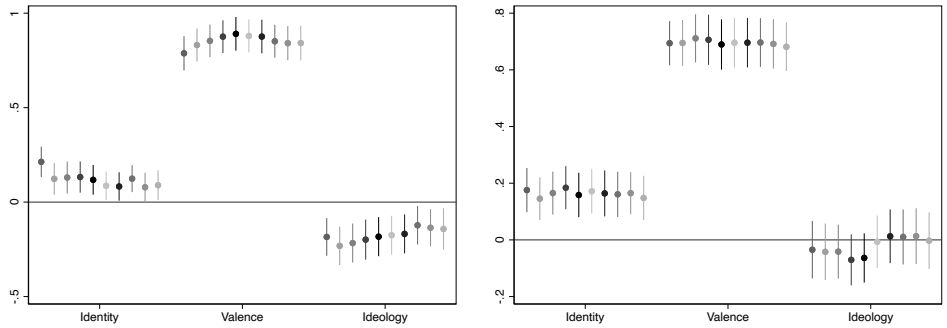
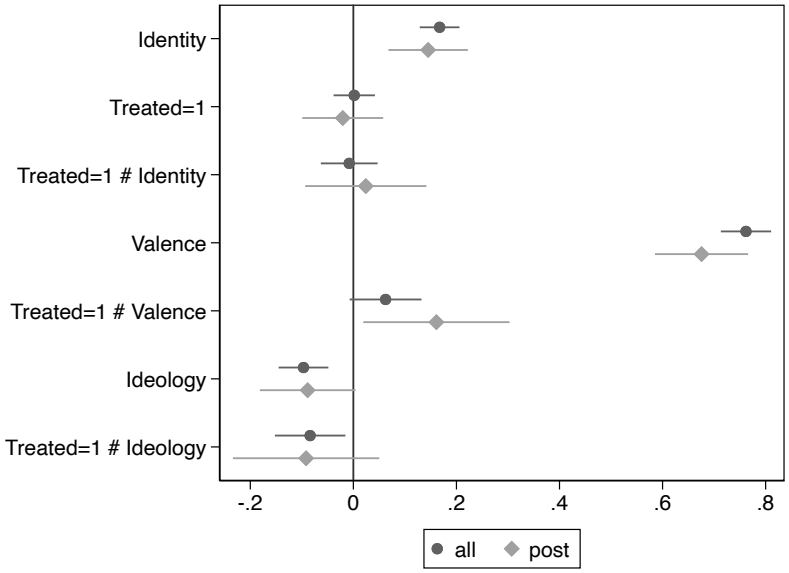


Figure A.2: Robustness: Five-period Rolling Windows Analysis



Note: The left panel reports re-estimates of the weights assigned by Agency voters to candidate characteristics using five-period rolling windows based on the order in which candidate pairs are presented to respondents, while the right panel presents the corresponding estimates for Deferential voters. Vertical bars denote 95% confidence intervals constructed from standard errors clustered at the respondent level. All regressions apply survey weights to ensure representativeness of Indonesia's 2024 electorate.

Figure A.3: Persistence: Treatment Effects Post-election



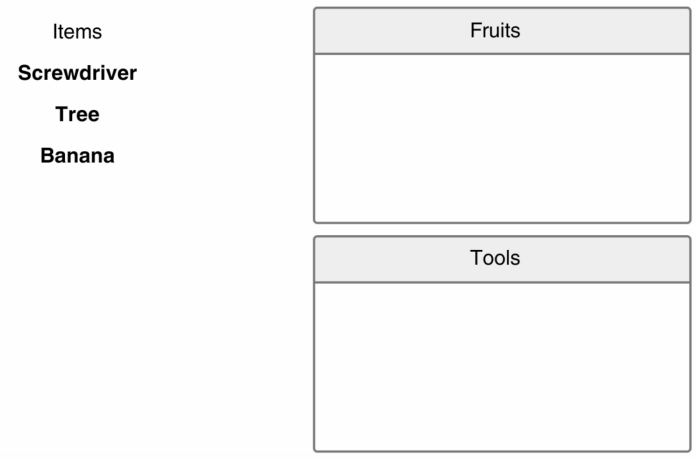
Note: The figure reports estimates of the relative weights assigned to aggregate candidate traits and the corresponding treatment effects from the random–utility specification in Equation 2. Dots and diamonds plot Probit coefficients for two samples: (i) all respondents and (ii) the post–election–only subsample. To preserve power, we adopt a broader sample definition that retains respondents with missing *drag & drop* categorization game scores. Horizontal bars denote 95% confidence intervals based on standard errors clustered at the respondent level. All regressions apply survey weights to be representative of Indonesia's 2024 electorate.

B Measurement of Democratic Narratives

This appendix provides full details on the measurement and scoring of respondents' democratic narratives using a *drag & drop* categorization task. Respondents completed three categorization tasks in which they were asked to match objects displayed in a left-hand box to category labels displayed in two boxes on the right. The first question is a practice question (categorizing fruits and tools), while the two subsequent questions are substantive and enter the calculation of the Agency score.

Figure B.4: *Drag & Drop* Categorization Game: Practice Question

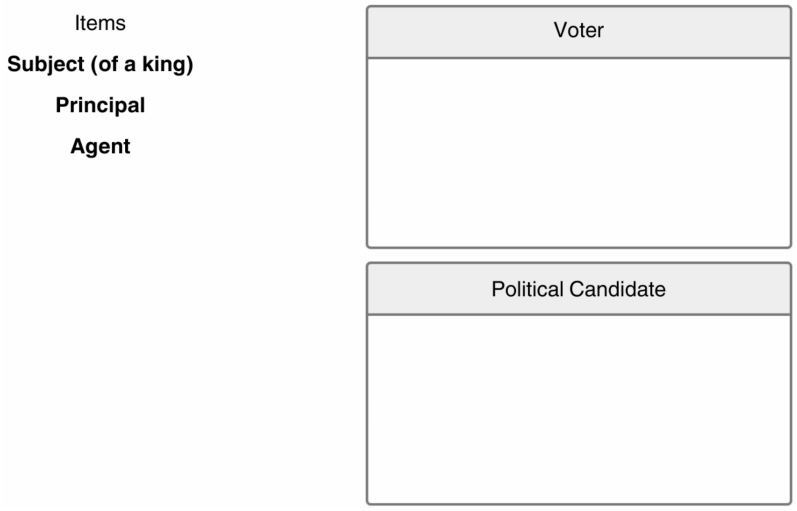
"From the three objects on the left, there is one object that is similar to the object in each box on the right. Drag and drop one object from the left box into each corresponding box on the right. (Each box on the right has only one matching object.)"



In the first substantive question, respondents were asked to categorize three objects into two boxes labeled *"Voter"* and *"Political candidate."* Although the terms "principal" and "agent" may sound very technical in English, Bahasa Indonesia has widely used and intuitive equivalents: *pemberi amanat* (giver of mandate) for principal and *penerima amanat* (receiver of mandate) for agent.

Figure B.5: *Drag & Drop* Categorization Game: Substantive Question 1

"Same as before, we kindly ask for your assistance in matching one object in the left box with the category in the right box that you find most suitable."



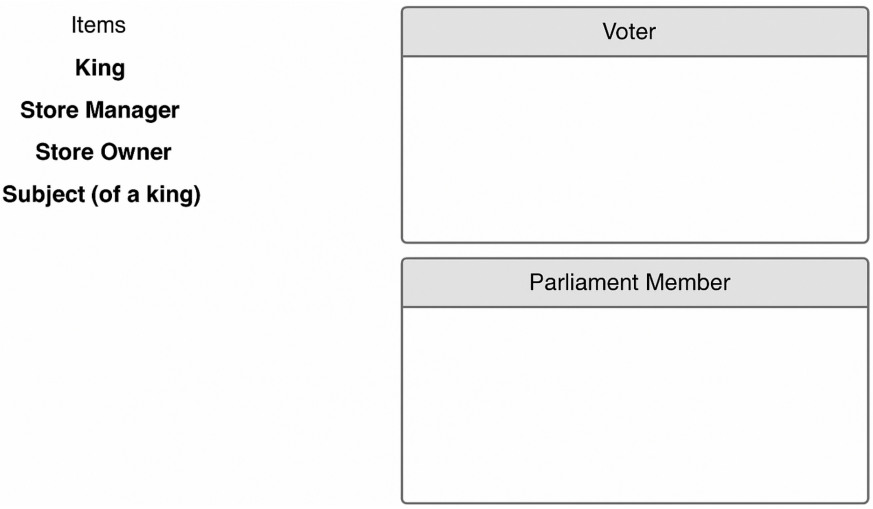
Responses were scored as follows:

- Score = 1: Voter matched with *Principal* and political candidate matched with *Agent*.
- Score = −1: Voter matched with *Agent* and political candidate matched with *Principal*.
- Score = 0: All other assignments.

In the second substantive question, respondents categorized four objects into boxes labeled *"Voter"* and *"Parliament member."*

Figure B.6: *Drag & Drop* Categorization Game: Substantive Question 2

"Same as before, we kindly ask for your assistance in matching one object in the left box with the category in the right box that you find most suitable."



Responses were scored as follows:

- Score = 1: Voter matched with *Shop owner* and parliament member matched with *Shop employee*.
- Score = 0.5: Voter matched with *King* and parliament member matched with *Subjects (of a king)*. While this mapping is less precise—since subjects are not accountable to kings in the same way employees are to employers—it nonetheless reflects voter sovereignty and a non-hierarchical conception of democratic authority.
- Score = −1: Parliament member matched with *King* or *Shop owner*, and voter matched with *Subjects (of a king)* or *Shop employee*, reflecting a hierarchical view in which political authority flows from elected officials to voters.
- Score = 0: All other combinations.

We construct a respondent-level democratic narrative index as a weighted average of the two substantive questions, where the weight to the second question is twice that of the first. The second substantive question is cognitively more demanding than the first: It involves four objects rather than three and requires respondents to first identify a principal–agent structure before mapping it onto a metaphor.

All task design, scoring rules, and weighting procedures were fully pre-specified in the pre-analysis plan, available at the Social Science Registry (https://www.socialscienceregistry.org/trials/12872).

C Average Marginal Component Effect (AMCE)

Despite its advantage of providing a more structural way of modeling voter utility with clearly laid out assumptions and hence permitting the interpretation of the estimated coefficients as the weights assigned to traits (Abramson *et al.*, 2022), the random utility framework is somewhat inflexible, since one cannot really compare the estimates across different specifications, e.g., with interaction terms or different sets of regressors. In this section, we use a non-parametric approach. These estimates provide complementary perspectives on how respondents choose, which may not be obtained using the random utility model.

AMCE analysis following Hainmueller *et al.* (2014) has been the most widely used technique to analyze full-factorial choice data in political science. Using this method, one can obtain a non-parametric estimate of the effect of a given value of a trait, compared to the base category. The effect of moving from base category t_0 to the category of interest t_1 for a given trait is defined as

$$\hat{\pi}(t_1, t_0) = \mathbb{E}(Y_{ij}|T_{ij} = t_1) - \mathbb{E}(Y_{ij}|T_{ij} = t_0) \tag{4}$$

where Y_{ij} is an indicator for whether individual i chose candidate j. Notice that in practice, this involves pooling the data from fifteen choices made by each respondent, for a total of thirty profiles. Given that the profiles are randomly and independently drawn, and under the assumption that *when* a profile is viewed is inconsequential for the choice (no profile-order effects), one can pool all the profiles seen by a respondent (of which fifteen were chosen and fifteen were rejected) and calculate the effect as the difference in mean approval for profiles that have value t_1 on the trait (averaging over all other characteristics) and the mean approval for profiles that have value t_0. One valid interpretation of $\hat{\pi}(t_1, t_0)$ in the AMCE framework, that carries substantive meaning, is the expected change in vote share of a candidate due to a change in their attributes from t_0 to t_1, regardless of the remainder of the traits (Abramson *et al.*, 2022).

AMCE estimation is somewhat more flexible than the random utility model, and does not rely on a fully specified utility model for voters. All inferences for the estimation strategy in this section are based on standard errors clustered at the respondent level, as per Hainmueller *et al.* (2014). The results of the basic non-parametric estimation are reported in Table C.1. Reassuringly, the results are substantially equivalent to those of the random utility estimation.[30]

Valence is the trait with the largest effect, followed by identity and then ideology. The magnitude is also comparable with the marginal effects resulting from the random utility model. Using this model, we find that Agency voters value valence and ideological position

[30]Identity and valence are parametrized with one dummy variable each (respectively an indicator for congruent identity—vs. discordant ethnic or religious identity—and an indicator for profiles with high valence—vs. low valence). As explained in Subsection 3.4, the effect of ideology in the AMCE model is estimated by including two dummies: one dummy equal to one when the respondent and the candidate in the profile are on the same side of an ideological issues, and one dummy for when respondents are neutral about the issue. The base category refers to when the respondent and candidate are on opposite sides, e.g., the candidate is progressive on traditional morality issues and the respondent is conservative on these. We also run an alternative analysis where respondents who do not have a stance on the issue are not included in the estimation sample, and the results are substantially equivalent.

Table C.1: Importance of Traits (Aggregate) and Treatment Effects: AMCE Analysis

	$\boldsymbol{P(voted = 1)}$		
VARIABLES	Combined	Deferential	Agency
Share Identity × Treated	0.001	-0.002	0.004
	(0.009)	(0.013)	(0.014)
High Valence × Treated	0.015	0.038**	-0.011
	(0.011)	(0.015)	(0.016)
Ideology Neutral × Treated	0.002	0.010	-0.005
	(0.011)	(0.016)	(0.016)
Ideology Same × Treated	0.026***	0.027**	0.029**
	(0.009)	(0.012)	(0.012)
Share Identity = 1	0.050***	0.055***	0.044***
	(0.007)	(0.009)	(0.009)
High Valence = 1	0.273***	0.252***	0.296***
	(0.008)	(0.011)	(0.011)
Ideology side = Neutral	0.019**	0.006	0.034***
	(0.008)	(0.011)	(0.012)
Ideology side = Same	0.026***	0.008	0.046***
	(0.006)	(0.008)	(0.008)
Treated = 1	-0.019**	-0.031***	-0.006
	(0.007)	(0.010)	(0.011)
Constant	0.333***	0.353***	0.312***
	(0.005)	(0.007)	(0.008)
Observations	104,272	50,040	54,232
R-squared	0.082	0.077	0.090

Notes: The table reports Average Marginal Component Effects (AMCE) estimates, the counterpart of the random utility model estimated in the main tables. The first column (*Combined*) includes all respondents who completed the *drag & drop* categorization game. *Deferential* includes voters with Agency scores < 0; *Agency* includes voters with scores ≥ 0. *Share Identity* is an indicator for sharing religion *and* ethnicity with the candidate (base category = does not share identity). *High Valence* is an indicator for a high track record on any one of the five governance policies (base = low track record/poor performance). *Ideology side* equals one if the respondent is neutral or on the same ideological side as the candidate on any one of the five ideological dimensions (base = opposing side). Standard errors (in parentheses) are clustered at the respondent level. Survey weights are applied to ensure representativeness of Indonesia's 2024 vote-eligible population. *** p<0.01, ** p<0.05, * p<0.1.

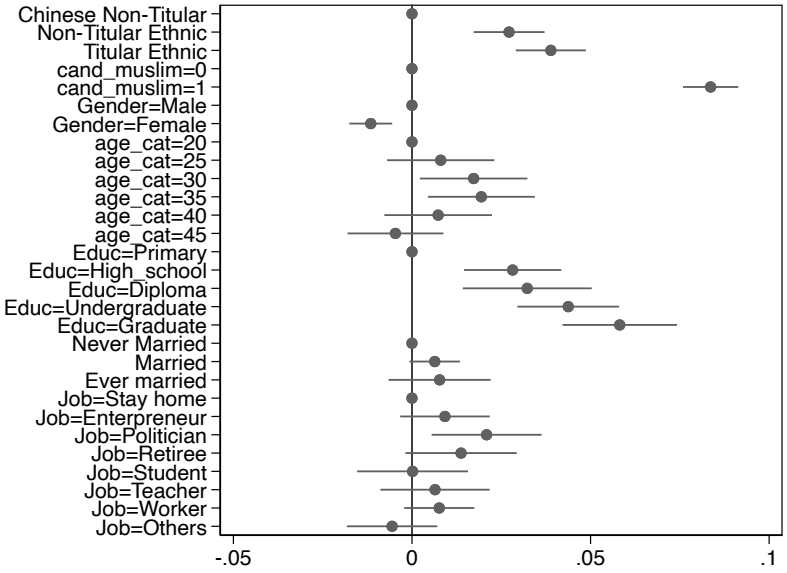
more than Deferential voters. Importantly, by and large, similar conclusions can be drawn for the treatment effects: while no treatment effect is detected on the role assigned to identity, the treatment significantly increases the effect of valence and the ideological position of the candidate. Importantly, these effects are, just like in the main model, mostly driven by voters who, pre-treatment, do not conceptualize democratic representation as an agency relationship.

The pattern of relative importance of traits documented in the random utility model is also observed from the analysis of the disaggregated traits.[31] In particular, Figure C.2 shows that religious congruence is much more important than ethnic congruence. Among the valence traits, a bad record in terms of corruption has the highest effect, followed by economic, health, and education policy, while the environment, and women and child protection, rank last. Some additional results are of substantive interest. In particular, we can estimate separate effects for being on the same side or the opposite side of the ideological divide, testing whether the bonus from ideological proximity is symmetric to the malus from ideological distance. According to the results in Figure C.2, the bonus for sharing a position with the respondent in a given dimension is not necessarily symmetric to the malus from holding an opposing position.[32]

[31] In Table C.1 we use the "aggregated" traits, which means that a profile is classified as having high valence whenever the profile lists a positive track record, regardless of the policy area mentioned; thus, we only have one coefficient. Meanwhile, in the disaggregated analysis, the profile is classified as having high or low valence according to the random assignment to one single policy area; the comparison is to profiles where the policy area is not mentioned.

[32] We find a small anomaly in the effect of the position on the role of the army in politics, where both having the same position as the respondent or an opposing position is found to give a disadvantage, compared to profiles where this issue is not mentioned; hence, profile mentions of *any* stance on the role of the army in national politics harm candidates.

Figure C.1: Weight Assigned to Raw Traits of Candidates

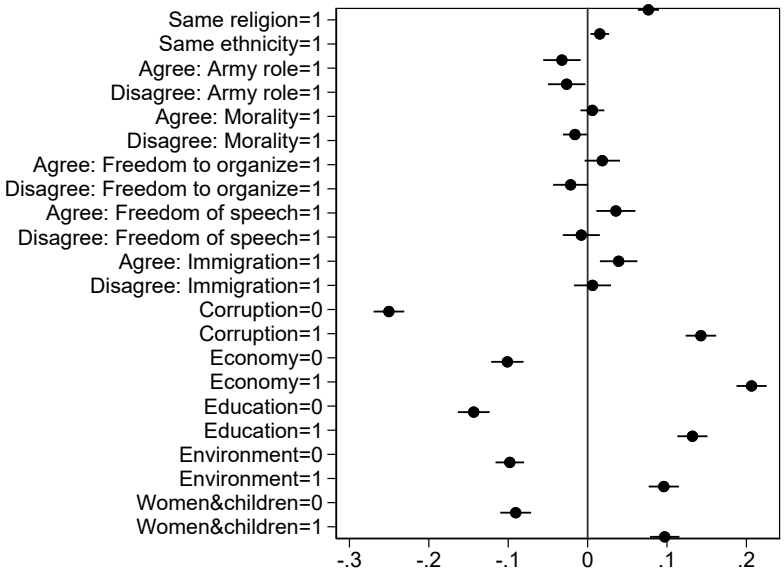


Note: The graph presents the weight assigned to the raw traits of candidates using the Average Marginal Component Effect (AMCE) method. Each dot represents estimates of a linear probability model of being chosen on the raw traits of candidate presented in the candidate choice experiment relative to the base traits, where the base-trait estimates are represented by dots along the zero line. The bars represent 95% confidence intervals. The standard errors are clustered at the individual respondent level. All regressions are weighted to be representative of voter's population in Indonesia's 2024 general election.

We can also look at asymmetric effects of high vs. low valence, disaggregating by policy area. In this way, we can back out separate effects of having a positive vs. a negative track record in each of the five policy areas. The effects are approximately symmetric for education, the environment, and the rights of women and children. On corruption, the malus of having a bad record is larger in absolute value than the bonus of having a good record; conversely, for the economy, the bonus from a good track record is larger in absolute value than the malus from a poor track record.

Another advantage of the non-parametric approach is that we can also estimate the component effects of the raw original value traits instead of the congruence terms. This is informative about which features, averaging across the entire electorate, might make a candidate electorally more successful. The analysis is presented in Appendix Figure C.1. We learn that compared to the non-parametric analysis—where the traits are computed in terms of voter-candidate congruence—the effect of the raw traits is not as large. Note that in this framework we can directly compare the magnitudes in Figure C.2 and Figure C.1: the effect of valence is much larger than the effect of education, age, gender, or previous occupation of the candidate. This provides further justification for the choice of the main traits in the random utility model, as these are essentially the candidate characteristics that matter the most. Crucially, we can conclude that the profile of a successful candidate in the eyes of the Indonesian electorate is that of a Muslim, non-Chinese, male, of prime age (30-40), master's-level educated, married, with prior experience as a politician.

Figure C.2: Weight Assigned to Disaggregated Traits: AMCE Analysis



Notes: The graph shows AMCE estimates for disaggregated candidate traits. Each dot corresponds to the estimate from Eq. (4), obtained separately for each ideological issue and policy area. The reference category—corresponding to the zero line—is the average of all profiles without that trait. Bars indicate 95% confidence intervals. For valence dimensions, performance is coded as one for high performance and zero for low performance. Standard errors are clustered at the individual respondent level. Survey weights are applied to ensure representativeness of Indonesia's 2024 vote-eligible population.

C.1 Interaction Effects

Another advantage of this framework is that one can straightforwardly estimate the interaction of two or more traits. This allows us to answer questions about, for instance, whether being corrupt matters more for candidates that share the ethno-religious identity of the respondent, or more for out-groups.[33] Significant interactions between candidate traits constitute deviations from the additive separability assumption adopted in the utility modeling. In substantive terms, deviations from additive separability mean that the effect of one trait depends on the value that another trait takes: for instance, if voters value valence more for candidates of their own ethno-religious group than for out-groups, or if they value competence more when facing an ideologically close candidate, while they value it less (or even dislike it) when facing an ideologically distant candidate.

[33]Notice that the inclusion of multiplicative interactions is clearly possible also in a Probit setting. Yet, as discussed earlier, in order to meaningfully interpret the estimates of the coefficients as utility weights, the variance of the error term and hence the specification needs to be fixed once and for all; comparing across specifications does violence to the "structural" interpretation of the random utility formulation.

Table C.2: Interaction Effects between Valence, Ideology, and Identity

	$P(voted = 1)$		
	Identity	Religion	Ethnicity
High Valence = 1	0.268***	0.269***	0.267***
	(0.009)	(0.012)	(0.009)
Shared Identity = 1	0.039***	0.071***	0.010
	(0.011)	(0.011)	(0.010)
Ideology side = Neutral	0.010	0.016	0.015
	(0.011)	(0.016)	(0.013)
Ideology side = Same	0.024***	0.017*	0.026***
	(0.007)	(0.010)	(0.008)
High Valence × Shared Identity	0.014	0.005	0.010
	(0.013)	(0.013)	(0.011)
Ideology Neutral × Shared Identity	0.031	0.007	0.008
	(0.023)	(0.021)	(0.021)
Ideology Same × Shared Identity	0.004	0.012	-0.000
	(0.012)	(0.012)	(0.011)
Constant	0.338***	0.303***	0.345***
	(0.006)	(0.008)	(0.007)
Observations	52,756	52,756	52,756
R-squared	0.077	0.080	0.075

Notes: The table reports interaction effects among *Shared Identity*, *High Valence*, and *Ideology side*, estimated on the control arm only. *Shared Identity* is a dummy indicating whether the respondent shares identity with the candidate (base = does not share identity); its definition varies by column: Column (1) uses an aggregated measure (religion *and* ethnicity), Column (2) uses religion only, and Column (3) uses ethnicity only. *High Valence* is a dummy for a high track record on any of the five governance policy areas (base = low track record / poor performance). *Ideology side* is an identifier for the respondent's ideological position relative to the candidate in any one of the five ideological dimensions (base category = opposing side). Numbers in parentheses are standard errors clustered at the respondent level. Survey weights are applied to ensure representativeness of Indonesia's 2024 vote-eligible population. *** p<0.01, ** p<0.05, * p<0.1.

We estimate the effect for the binary identity-congruence variable (equal to one if *both* ethnicity and religion match between respondent and candidate, zero otherwise), as well as for valence and ideological congruence in any of the five dimensions shown for the candidate. Results are reported in Table C.2. The results shows that valence carries the same consequences for candidates that are ethno-religious in-groups and out-groups (Column 1). The same pattern holds if we use religion and ethnicity separately as the moderator (Column 2 and Column 3, respectively). This means that sharing identity with a candidate does not affect voter valuation of the importance of valence. This is reassuring, as these patterns give additional grounds for making a claim that we cautiously made already in the random

utility section: that between a corrupt in-group and a clean out-group, voters would choose the non-corrupt candidate in spite of religious differences. Similarly, ideological congruence also matters in a similar fashion regardless of one's identity congruence. In an additional analysis, we also check that the interaction between valence and ideology is not statistically different from zero. This implies that the assumption of additive separability in this case holds. None of the coefficients for the interaction variables are statistically significant.

D Video Scripts

Duration: 75 seconds

D.1 Video 1 - treatment

EN:

Imagine you inherit a convenient store from your grandparents. However, you have your own job and do not have the expertise to manage it.

Therefore, you run a search for the store manager who will act on your behalf to manage it, so you need to find someone that you can trust to act for your best interest.

Due to the power you give him, he may be tempted to do things that benefit himself but are costly for you. Therefore, you need to check on him from time to time, making sure that he manages your store with competence, dignity, and integrity. If you think he does not fulfill your expectations, or is betraying your trust, you have every right to fire him and elect someone else to do the job.

This is a metaphor of an election (PEMILU) in democracy. In an election, we choose a political candidate to run the country on our behalf.

Just like the case of the store manager, political leaders will be tempted to abuse their power to benefit themselves at our cost.

Therefore, in an election, we can "punish" them by electing someone else when they are lying, cheating, or under-performing.

So, are you ready to vote in the coming election?

ID:

Bayangkan kamu mewarisi sebuah toko serba ada dari kakek kamu. Namun, kamu memiliki pekerjaan sendiri dan tidak memiliki keahlian untuk mengelolanya.

Karena itu, kamu mencari seorang manajer toko yang bertindak atas nama kamu untuk mengelolanya. Orang ini harus bisa kamu percayai akan bertindak sebaik-baiknya demi kepentinganmu.

Karena kepercayaan yang diberikan kepadanya, dia mungkin akan tergoda untuk melakukan hal-hal yang menguntungkan dirinya sendiri tetapi merugikan kamu; sehingga kamu perlu mengawasinya dari waktu ke waktu, memastikan dia mengelola tokomu dengan kompetensi, martabat, dan integritas. Jika dia tidak memenuhi harapanmu, atau mengkhianati kepercayaanmu, kamu punya segala hak untuk memecatnya dan memilih orang lain untuk melakukan pekerjaan itu

Ini adalah perumpamaan PEMILU dalam demokrasi. Dalam pemilu, kita memilih CALEG untuk menjalankan negara atas nama kita. Seperti kasus manajer toko di atas, pemimpin politik sering tergoda untuk menyalahgunakan kekuasaan demi keuntungan sendiri; merugikan kita tanpa kita sadari. Dengan PEMILU, kita bisa "menghukum" mereka dengan memilih orang lain jika kita dapati mereka berbohong, curang, atau berkinerja buruk.

Jadi, apakah kamu siap untuk memilih pada pemilu mendatang?

D.2 Video 2 – control

EN:

Election (PEMILU) is around the corner, but what do we know about elections? PEMILU stands for Pemilihan Umum. We held our first democratic election in 1995 and have experienced twelve general elections since. An interesting fact about this year's election is that we will hold the national as well as the regional level elections all at once.

Who is eligible to vote? If you are 17 years old or older, or you have had been married before, then you are eligible to vote. Of course, you need to have a valid national ID (KTP) to certify that you are indeed a citizen of the Republic of Indonesia. Moreover, a person who is eligible to vote has to be mentally sound, and they cannot be a member of the national police force (POLRI) or a member of the national army (TNI).

We should check whether our names are registered as valid voters (Daftar Pemilih Tetap - DPT) on the KPU website. When you find yourself registered, you should also check if your information details are accurate. Make sure that your residence is in line with the location where you will vote (Dapil). If you detect any discrepancy, you should report it to KPU and request a data change. You will then be added to the additional voter list (Daftar Pemilih Tambahan).

So, are you ready to vote in the coming election?

ID:

PEMILU sudah dekat, tapi apa yang kita ketahui tentang pemilu? PEMILU adalah singkatan dari Pemilihan Umum. Kita menyelenggarakan pemilu demokratis pertama kali pada tahun 1995 dan telah menjalani dua belas kali semenjak itu. Fakta menarik pada pemilu tahun ini adalah kita akan menyelenggarakan pemilu tingkat nasional dan tingkat daerah secara serentak.

Siapa yang berhak memilih? Jika kamu sudah berusia 17 tahun, atau pernah menikah sebelumnya, maka kamu berhak memilih. Tentunya kamu harus memiliki KTP yang masih berlaku untuk membuktikan bahwa kamu memang warga negara Republik Indonesia. Selain itu, orang yang berhak memilih harus sehat secara mental, dan anggota POLRI atau TNI tidak boleh ikut PEMILU.

Kita bisa mengecek apakah nama kita terdaftar sebagai pemilih sah dalam Daftar Pemilih Tetap (DPT) di website KPU. Ketika terdaftar, kamu juga sebaiknya memeriksa apakah rincian informasi kamu akurat. Pastikan tempat tinggal kamu sesuai dengan tempat kamu akan memilih (Dapil). Jika kamu menemukan kesalahan, kamu bisa melaporkannya ke KPU dan meminta perubahan data. Kemudian, kamu akan ditambahkan ke daftar pemilih tambahan (DPT).

Jadi, apakah kamu siap untuk memilih pada pemilu mendatang?

E List of Statements Associated with Candidates

Table E.1: Statements on Track-Record/Performance Associated with the Candidate

Dimension	Performance	Statement alternative
Corruption	High	• No track record of corruption. • Successfully reduced corruption in local government.
	Low	• Previously involved in corruption cases. • Involved in bribery cases for business permits. • Former convict.
Education and health	High	• Improved the quality of schools in the region. • Improved service quality and infrastructure of health facilities. • Established free schools. • Increased school participation rates.
	Low	• Failed to reduce maternal and infant mortality and stunting. • Unsuccessful in curbing the dropout rate.
Economy	High	• Reduced poverty, inequality, and unemployment. • Strengthened guidance and support for SMEs to boost the local economy.
	Low	• Failed to stabilize prices of basic commodities. • Poverty, inequality, and unemployment increased during the term.
Environment	High	• Reduced regional carbon emissions. • Addressed emergency waste problems.
	Low	• Failed to address air and river pollution from factory waste. • Failed to resolve the regional waste emergency. • Granted permits for opening mining companies.
Rights of women and children	High	• Proposed new programs for women's empowerment. • Improved reporting services for domestic violence.
	Low	• Did not reduce domestic violence cases. • Failed to curb child labor.

Table E.2: Statements on Ideological Positions Associated with the Candidates

Issue	Stance	Statements
Issue 1: Roles of the military	Conservative	Strengthen the role of the armed forces (TNI) within state institutions for national stability and public security.
	Progressive	Prioritize a democratic system that guarantees the neutrality of the armed forces.
Issue 2: Traditional morality	Conservative	Strengthen law enforcement to protect children even from the womb (criminalizing abortion).
	Progressive	Safeguard women's abortion rights in cases of rape or medical emergencies.
	Conservative	Uphold moral and religious values in society, such as criminalizing prostitution.
	Progressive	Protect the labor rights of women working in the sex industry.
	Conservative	Enact local regulations consistent with sharia, e.g., regulating adultery and sexual deviance.
	Progressive	Uphold human rights and reject laws that discriminate against specific groups.
Issue 3: Freedom to organize	Conservative	Oppose radicalism, extremism, and intolerance that threaten regional security.
	Progressive	Guarantee freedom of association and organization.
Issue 4: Freedom of speech	Conservative	Strengthen law enforcement against the spread of hate speech toward public officials online.
	Progressive	Guarantee freedom of expression and the right to criticize public officials in digital spaces.
Issue 5: Immigration	Conservative	Promote foreign investment to expand employment and support regional economic growth.
	Progressive	Protect the rights and obligations of local workers.

F The weighting method

We detect some discrepancies between the marginal distributions of some traits in our sample and those in official census data. We instructed the survey firm to collect data based on the marginal distribution of age, province, and gender, and the conditional distribution of education by age. Yet, our sample is more educated than the general population and there are some discrepancies in terms of geography, in particular an overrepresentation of the city of Jakarta. In addition, there is some imbalance in terms of gender, with an overrepresentation of females. Hence, we estimate raking weights based on census data and individual education (treated as a binary indicator for university-educated respondents), gender, and a six-category variable for residence in one of the five largest provinces in terms of population [34], aggregating the remainder of the provinces in one residual category. We use the `anesrake` package in the R environment to estimate weights to match the sample marginals to the census. We estimate three different vector of weights, to account for attrition. The first vector of weights rebalances the full sample interviewed at baseline with census marginals. The second vector of weights rebalances participants in the pre-election wave with census marginals, and the third vector rebalances participants in the post-election wave with census marginals. Ultimately, given the small number of variables used in the raking exercise, we are able to rebalance the sample to census marginals with "gentle" weighting: the maximum weight in the baseline sample is 2.9; it is 2.6 for the pre-election wave, and 2.4 in the post-election wave. In any case, we are able to exactly match the census targets for these three variables.

[34] Jawa Barat, Jawa Tengah, Jawa Timur, Sumatera Utara, DKI Jakarta.