

Unbundling Autocratic Capture: Bundled Reform, Selective Emphasis, and Democratic Accountability

Alper Şükrü Gençer

August 28, 2026

Abstract

Why do voters support reforms that weaken democratic oversight? I argue that incumbents can make harmful provisions easier to accept by packaging them with attractive or technical components and emphasizing the attractive parts, making the trade-offs embedded in the package harder to identify. Support for such reforms can therefore look like voters are trading democracy for other benefits, even when the democratic downside was never clear enough to shape their choices. Bundling turns the voters' decision into a judgment about a package; complexity makes the package harder to parse; selective emphasis highlights benefits while obscuring democratic risks. This process is especially consequential in captured media environments, where government-aligned outlets amplify promised benefits and make democratic risks harder to encounter. I test this argument in Türkiye, a hard case because voters have extensive experience with omnibus reform politics. Across randomized abstract choice tasks and political evaluation vignettes, bundled presentation worsens choice quality, selective emphasis can distort evaluation when the task is demanding, and oversight-weakening proposals gain support while recognition of their democratic costs weakens. The results show that democratic accountability can fail even before voters make a transparent trade-off, because whether they see what is at stake depends on how reforms are assembled, communicated, and contested.

1 Introduction

Why do voters sometimes support reforms that weaken separation of powers and democratic oversight? Take the case of Türkiye’s 2010 constitutional referendum, for example. The government presented the package through broadly appealing claims about democratization, civilianization, and institutional reform, yet it also changed the composition and appointment rules of key judicial bodies, including the Constitutional Court and the High Council of Judges and Prosecutors (Somer, 2019; Esen, 2021). The referendum was one constitutional example of a broader pattern. Much of Türkiye’s institutional change under the Justice and Development Party (AKP) moved through *torba yasalar*¹ and other reform packages, where oversight-weakening provisions were placed alongside technical, administrative, or publicly attractive measures (Kadioğlu, 2021; Özcan and Kimya, 2024). Existing accounts explain support for such democracy-weakening reforms through partisan loyalty, polarization, charismatic attachment, weak democratic commitments, or strategic trade-offs between democracy and other outcomes (Svolik, 2018; Graham and Svolic, 2020; Hollyer, Klačnja and Titiunik, 2023; Grossman et al., 2022; Krishnarajan, 2023; Luo and Przeworski, 2019; Grillo and Prato, 2023; Gratton and Lee, 2024). These explanations remain important, but they often begin after a demanding prior step: voters must recognize that an otherwise attractive reform package weakens separation of powers or democratic oversight before the democratic cost can shape their judgment. This problem is not confined to Türkiye. Hungary’s post-2010 constitutional transformation, India’s 2020 farm-law episode, and Israel’s 2023 judicial overhaul show how consequential institutional or policy changes can reach citizens through multi-component packages whose public defense emphasizes reform, efficiency, public benefit, or democratic restoration (Bánkuti, Halmai and Scheppele, 2012; Behl, 2022; Navot, 2023).

I argue that this failure can arise from the way reforms are bundled and presented, before voters decide whether they are willing to trade democratic oversight for other benefits. The packaging at the center of this argument is not unique to Türkiye: legislative scholars have long shown how omnibus bills and multi-issue reform packages combine otherwise separate provisions into a single proposal (Krutz, 2001; Sinclair, 2012). When a reform reaches voters as a package, its democratic implications must be understood from the combined effect of several provisions rather than from a single institutional change; as the proposal includes more components, policy domains, or technical details, that judgment becomes harder (Gabaix and Graeber, 2024; Oprea, 2020; Kendall and Oprea, 2024). This evaluation

¹ *Torba yasa* literally means “bag law” in Turkish. I use the term to refer to omnibus legislation that combines amendments across multiple statutes or policy domains into a single legislative package (Sav, 2014).

burden makes presentation more consequential because selective emphasis—the highlighting of promised benefits such as efficiency, security, service delivery, or democratization—can shape which parts of the package voters notice and which risks to oversight they miss. This emphasis should matter more in captured media environments, where government-aligned outlets repeat favorable claims and make warnings about weakened oversight harder to encounter, because voters then face a package whose benefits are easier to see than its democratic downside (Tversky and Kahneman, 1981; Bordalo, Gennaioli and Shleifer, 2012, 2013; Gabaix, 2014; Matějka and Tabellini, 2021). I develop this argument as a formal theory of bundle evaluation.

I test this argument in Türkiye using four randomized studies that include abstract tasks with known payoffs as well as realistic candidate-choice and municipal reform vignettes. Across the studies, package presentation increases suboptimal choice in abstract tasks compared with simpler presentation, while selective emphasis can further distort evaluation when the task is demanding. In the politically realistic choice tasks, executive-aggrandizing platforms and oversight-weakening reform packages gain more support while recognition of their democratic costs weakens. The central implication is that democratic resistance can fail even when citizens would reject a clearly identified threat to oversight, because reforms may be assembled and communicated in ways that make their democratic downside harder to understand.

This misevaluation gap between the reform citizens are formally asked to judge and the reform they actually understand at the moment of choice does not require voters to be indifferent to democracy, hostile to constraints, or blindly loyal to the actor proposing reform. The same package can produce several kinds of failure. Some voters may miss the oversight-weakening implication altogether. Others may understand that something changes, but fail to connect the provisions to separation of powers or democratic oversight. Still others may recognize the risk, yet give it less weight once it appears alongside promised benefits such as efficiency, security, or service delivery. In each case, democratic resistance weakens because citizens evaluate the reform without a clear understanding of what it would do to oversight and institutional constraint.

Türkiye is a useful case because it is not an idiosyncratic outlier: it belongs to a broader family of cases in which democratic erosion proceeds through formally legal, reformist, and multi-component institutional change. The ingredients required by the argument are widespread: incumbents in many backsliding settings weaken oversight through constitutional amendments, ordinary legislation, administrative reforms, or emergency measures that are defended through claims about efficiency, stability, security, democratization, or public benefit (Sebők, Kiss and Kovács, 2023; Druckman, 2024; Grillo et al., 2024). Within this

broader family, Türkiye is a demanding case because bundled reform is familiar rather than novel. The repeated use of *torba yasalar*, constitutional packages, emergency decrees, and executive-centered administrative reforms means that Turkish voters have repeatedly encountered institutional change in packaged form (Kadioğlu, 2021; Özcan and Kimya, 2024). This prior exposure matters: if packaging still shapes evaluation in this setting, the result is harder to dismiss as confusion caused by an unfamiliar experimental format. Türkiye also helps connect selective emphasis to real politics. As media capture deepened, citizens were more likely to encounter favorable descriptions of reform through government-aligned mainstream outlets than warnings about weakened oversight (Esen, 2021). The case therefore helps study a broader problem in contemporary backsliding, where citizens encounter formally legal reform packages in captured media environments that make promised benefits easier to encounter than their democratic implications (Shirikov, 2024).

The empirical design begins by isolating the evaluative problem and then asks whether the same pattern appears in political choice. The abstract tasks do this by stripping away partisan attachment, policy agreement, candidate appeal, and other political motivations that could explain support for a harmful reform. Respondents choose between options with known payoffs, so accepting the worse option can be interpreted as suboptimal choice rather than ideological agreement, partisan loyalty, or a conscious democratic trade-off. I then move to candidate-choice and municipal reform modules, where the same presentational features are embedded in more realistic political settings: complexity is varied through shorter versus longer platforms or packages, and selective emphasis is varied through neutral versus favorable presentation of attractive components. These modules preserve a fixed oversight-weakening contrast while bringing back what makes real reform choices difficult: voters may support a proposal because they miss the democratic cost, understand it only partially, recognize it but give it less weight inside a broader package, or perceive other benefits from the proposal.

Study 1 establishes the benchmark result: package presentation alone increases suboptimal choice. Respondents are 17.0 percentage points more likely to choose the objectively worse option when its consequences must be inferred from a bundle of three changes rather than directly as a simple option (95% CI: 12.2, 21.8), and they are less likely to identify the objectively better option when its value must be reconstructed from those changes. Study 2 then varies complexity and selective emphasis while holding payoff-relevant consequences fixed. The full-sample interactions are positive but imprecisely estimated. The clearest pattern appears among lower-capacity respondents² evaluating harmful bundles. Relative

²I use “lower-capacity” to refer to the preregistered cognitive-type classification described in the research design section. Among retained respondents, higher-capacity respondents answered both simple CRT items

to the low-complexity, neutral-presentation condition (three policy changes without selective emphasis), the high-complexity, selective-emphasis condition increases harmful-bundle acceptance by 11.6 percentage points (95% CI: 3.4, 19.8). Together, the abstract choice tasks show that presentation can impair evaluation even when the stakes are objective and experimentally controlled.

The third and fourth studies move to politically realistic choice tasks. Study 3 examines complexity and selective emphasis in a candidate-choice setting. Relative to the low-complexity, neutral-presentation condition, the high-complexity, selective-emphasis condition increases support for the candidate associated with an executive-aggrandizing commitment by 17.0 percentage points (95% CI: 11.1, 22.8), while recognition of the harmful implication falls by 7.9 percentage points in the same condition (95% CI: -14.1, -1.6). Study 4 then moves to a multidimensional municipal reform package. Relative to the two-policy, neutral-presentation condition, the four-policy, selective-emphasis condition increases package acceptance by 8.1 percentage points in the full sample (95% CI: 1.8, 14.3), while consequence recognition declines by an estimated 4.1 percentage points (95% CI: -8.8, 0.5). The political studies show that package presentation can increase support and weaken recognition, while also leaving room for more than one evaluative path: some respondents may miss the democratic cost, others may understand it only partially, and others may recognize it but give it less weight inside a broader package.

The paper makes three main contributions. First, it develops a formal theory of bundled policy evaluation that connects work on complexity, rational inattention, sparse attention, and salience. Existing models of inattention explain why decision makers economize on costly information processing, while salience models explain why prominent attributes receive disproportionate weight (Sims, 2003; Gabaix, 2014, 2019; Maćkowiak, Matějka and Wiederholt, 2023; Bordalo, Gennaioli and Shleifer, 2012, 2022). I bring these ideas together by distinguishing the quantity of attention from its composition: voters may devote more or less attention to a reform package, but presentation shapes which provisions that attention reaches. Complexity raises the burden of understanding the package, while selective emphasis changes which consequences voters are most likely to notice and use. The main text develops the theory’s logic and empirical implications, while Appendix A presents the full model and proofs. Although the empirical application centers on democratic backsliding, the theory applies more broadly to political choices that arrive as packages: campaign platforms, omnibus legislation, policy reforms, and proposals that combine popular benefits with more

correctly and at least one harder reasoning item correctly; lower-capacity respondents are those who do not meet that threshold. The label is used only to examine heterogeneity in the evaluative burden of reconstructing bundles, not to rank respondents more broadly.

extreme or harder-to-evaluate components. In these settings, political communication can shape not only how voters feel about a proposal, but also what they understand the proposal to contain.

Second, the paper clarifies a recognition problem at the center of democratic accountability. Existing work shows that citizens may tolerate democratic violations because of partisan loyalty, polarization, weak democratic commitments, or trade-offs between democracy and other valued outcomes (Svolik, 2018; Graham and Svolik, 2020; Grossman et al., 2022; Grillo and Prato, 2023; Braley et al., 2023). Recent work on democratic backsliding also emphasizes that erosion depends on the weakening of constraints on elected executives, the actors who can restrain them, and the psychological processes through which citizens respond to democratic threats (Druckman, 2024; Grillo et al., 2024). This paper asks whether voters recognize the oversight-weakening or executive-aggrandizing provisions inside omnibus reforms clearly enough for accountability or democratic trade-off theories to operate. This shifts attention to a recognition condition that trade-off accounts often leave implicit: voters can only knowingly tolerate democratic costs when those costs have become visible enough to enter their judgment.

Third, the paper gives a demand-side account of why legislative packaging matters for democratic backsliding. Research on omnibus legislation and policy bundling emphasizes how elites assemble coalitions, protect controversial provisions, or move otherwise separate measures through a single proposal (Krutz, 2001; Sinclair, 2012; Lee, 2022). Work on legislative backsliding shows that democratic erosion can also proceed through ordinary law, including changes in policy content, legal-constitutional form, legislative procedure, and institutional stability (Sebők, Kiss and Kovács, 2023). I connect these perspectives by showing that packaging can also matter after a proposal reaches citizens. By changing how reform is encountered, bundling, complexity, and selective emphasis can alter whether voters recognize a proposal’s oversight-weakening consequences and how those consequences enter political judgment.

The remainder of the paper proceeds as follows. The next section develops the theoretical framework, with the full formal model and proofs in Appendix A. I then situate the argument in Türkiye’s reform politics and explain why Türkiye provides a demanding setting in which to study the proposed mechanism. The research design section explains the experimental sequence. The empirical sections move from the abstract bundling benchmark to the core complexity-by-emphasis test, then to candidate choice and municipal reform evaluation. The final section discusses what the results imply for democratic backsliding and reform evaluation.

2 Theoretical Framework

Citizens can punish democratic backsliding only when they identify what a proposal actually does to oversight and institutional constraint. Existing trade-off accounts usually begin after that step: voters see a democratic cost and decide whether to accept it for partisan, policy, or expressive reasons (Svolik, 2018; Graham and Svolik, 2020; Grossman et al., 2022; Grillo and Prato, 2023; Braley et al., 2023). This paper moves one step earlier. When a reform reaches citizens as a package, the democratic cost may have to be reconstructed from a larger set of provisions before it can shape what they learn and whether they support or oppose the reform. I use reconstruction to mean the evaluative step in which citizens infer the relevant consequences of a package from its components before forming a judgment. A voter who misses that cost, identifies it only partly, or sees it less clearly than the attractive parts of the package is not making the same choice as a voter who knowingly accepts democratic harm.

Bundling makes the prior step harder because it changes what voters are asked to judge. A stand-alone provision can be considered on its own. A package asks citizens to form a judgment over several provisions at once, and the burden grows as the package includes more issues, policy domains, and consequences. Legislative scholars show that elites use omnibus packages to build coalitions, protect controversial measures, and move provisions that might struggle on their own (Krutz, 2001; Sinclair, 2012). The demand-side implication is that packaging can make the consequences of a proposal harder to identify. A voter may understand some parts of the package while missing, misclassifying, or only partly identifying others, including provisions that alter oversight or executive discretion.

H1. *Relative to direct presentation of the same final consequences, bundled presentation increases suboptimal choice because citizens must reconstruct the package before evaluating it.*

Complexity captures how demanding the package is once reform is presented in bundled form. Some packages contain only a few clear provisions; others are longer, more technical, or spread consequences across several policy domains. Work on complexity, sparse attention, and rational inattention shows that more demanding choice environments can make otherwise engaged evaluation less complete or less reliable (Gabaix, 2014; Oprea, 2020; Kendall and Oprea, 2024; Gabaix and Graeber, 2024). In the model here, this happens because additional issues and domains give the same effort more to cover. Citizens may spend more time or seek more information, but that effort can still leave them with less clarity about each component.

H2. *Higher complexity reduces clarity about the components of a package. This can increase*

evaluation error when citizens lose sight of decision-relevant consequences, even if they spend more effort evaluating the package.

The quantity-composition distinction matters because greater effort is not the same as more complete understanding. A harder package may induce more time, more clicks, or greater effort, yet still leave citizens with a less complete understanding of what the reform does. As the package becomes longer, more technical, or more multidimensional, the same amount of effort buys less scrutiny per unit of burden. Under those conditions, presentation becomes more consequential: when attractive components are emphasized while institutional costs remain less salient, citizens may understand the attractive parts of the reform more clearly than its costs. The result is not necessarily disengagement. Citizens can work at evaluating a reform and still come away with a partial, presentation-shaped version of its consequences.

Selective emphasis changes which parts of a package citizens are most likely to identify. Governing actors, campaigns, and government-aligned media rarely need to deny that a reform contains many provisions. They can instead describe the package through the parts that are easiest to defend: service delivery, administrative repair, security, rights, efficiency, or material benefits. This logic is related to framing and shortcut-based accounts of opinion formation, but it operates one step earlier. Framing research shows that presentation can change which considerations citizens use when forming opinions (Tversky and Kahneman, 1981; Kahneman and Tversky, 1984; Druckman, 2001; Chong and Druckman, 2007), and work on information shortcuts shows how accessible signals can substitute for direct evaluation (Zaller, 1992; Lupia, 1994). In bundled reform, however, the prior problem is identifying what the package contains. Selective emphasis can shape which components enter judgment clearly enough to be weighed in the first place. If citizens identify the attractive components more clearly than the oversight cost, support can rise because the package they judge is incomplete.

H3. *Selective emphasis increases support for a harmful package when attractive components are made easier to identify than costly institutional components.*

The model places an important restriction on this prediction. Selective emphasis should have little room to matter when the package is simple enough for citizens to identify both the attractive components and the institutional cost. It should also weaken when the package is so difficult that citizens identify little meaningful content at all. The effect should be most visible in the engaged range: the package is difficult enough to produce uneven identification, but still clear enough for emphasized components to enter judgment.

Cognitive capacity also matters because evaluating a package requires effort. Some re-

spondents can inspect several provisions with less strain; others face greater difficulty when the same package contains multiple components, technical language, or consequences spread across domains. While citizens remain engaged, this difference should affect what they identify inside the package. Lower-capacity respondents should identify fewer components and rely more on the components made easiest to see. When governing actors, campaigns, or aligned media emphasize attractive parts of the package, those parts should therefore play a larger role in lower-capacity respondents' judgments.

H4. *Lower-capacity citizens are more vulnerable to selective emphasis in the engaged range, because the same package leaves them with less clarity about its components.*

The same logic also cautions against treating lower capacity as a fixed ranking of vulnerability. Under severe complexity, some lower-capacity respondents may disengage and stop identifying meaningful content, including the emphasized components. Higher-capacity respondents may continue trying to evaluate the package, but their understanding can still become uneven: they identify some components, especially the emphasized ones, while missing or only partly identifying others. For this reason, the subgroup analyses should be read together with engagement and recognition outcomes rather than as a simple test of which group is always more susceptible.

Appendix A formalizes this theory. A reform package has full consequences, denoted by V , but citizens act on the version of the package they identify at the moment of judgment, denoted by $\hat{V}$. Bundling introduces the reconstruction problem by requiring citizens to infer consequences from components; additional complexity raises the burden of that reconstruction; selective emphasis changes which components are easiest to identify; and capacity affects how costly evaluation is for different respondents. Support depends on $\hat{V}$, while recognition depends on whether the harmful component enters judgment clearly enough to be reported. The hypotheses above therefore describe when $\hat{V}$ should depart from V , and when that departure should make harmful packages more acceptable.

The empirical sequence tests these expectations in complementary ways. Study 1 isolates the benchmark implication of H1 by comparing bundled presentation with direct presentation of objectively ranked consequences. Study 2 provides the cleanest controlled tests of H2–H4 by varying complexity and selective emphasis while holding payoff-relevant consequences fixed and examining heterogeneity by cognitive capacity. Studies 3 and 4 then move these presentational features into political settings, where support and recognition can diverge and where respondents may recognize a cost only partially or give a recognized cost less weight in their final judgment. Table 1 summarizes the relationship between the theoretical expectations and the four studies.

Table 1: Theoretical expectations and empirical tests

Hypothesis	Expectation	Primary test	Political extension
H1	Bundled presentation impairs evaluation relative to direct presentation of the same consequences.	Study 1	Studies 3–4 examine the same evaluative problem once consequences are embedded in political platforms and reform packages.
H2	Greater complexity reduces clarity about the components of a package.	Study 2	Studies 3–4 examine whether more complex political presentations alter support and recognition.
H3	Selective emphasis makes harmful packages more acceptable when attractive components are easier to identify than costly institutional components.	Study 2	Studies 3–4 examine whether favorable emphasis changes support and recognition in political choices.
H4	Lower-capacity citizens are more vulnerable to selective emphasis while they remain engaged.	Study 2	Capacity comparisons in Studies 3–4 provide secondary evidence about heterogeneity in political evaluation.

The theory makes a specific claim about democratic accountability: citizens may spend effort evaluating a reform and still judge an incomplete version of it. Complexity affects how clearly they identify the components of a package; selective emphasis affects which components become easiest to identify; and capacity affects how strongly those burdens are felt. Support for a harmful reform can rise without requiring citizens to prefer institutional harm, because the package they evaluate may differ from the package’s full consequences.

3 Türkiye as a Scope-Condition Case

Türkiye is useful for this paper because, during the AKP period, governing actors often advanced changes to judicial oversight, executive discretion, emergency authority, policing, and media regulation through legal reform packages rather than openly anti-democratic measures. They packaged oversight-weakening provisions with technical, administrative, distributive, or rights-oriented changes, while government officials and government-aligned media described these packages through favorable language: modernization, security, efficiency, stability, democratization, or administrative necessity. Türkiye therefore belongs to a broader family of legalistic backsliding cases in which incumbents weaken oversight through constitutional amendments, omnibus laws, decree-laws, or administrative reforms. It is also a demanding case for the experiments because packaged reform is familiar rather than novel. The studies ask whether bundled presentation and selective emphasis still shape evaluation in a setting where citizens already know this kind of politics.

The key legislative form is the *torba yasa*, or omnibus law. The term refers to laws that gather amendments across multiple statutes or policy domains into a single legislative package. Legal scholarship describes this practice as a de facto lawmaking form that became especially important after 2001 and helped shift legislative initiative toward the executive (Sav, 2014; Haspolat, 2016). A provision that weakens oversight or expands discretion can appear inside a broader reform that also includes technical fixes, administrative changes, distributive benefits, or rights-oriented measures. Citizens then evaluate the package as a whole, while the institutional consequences of one provision must be separated from the surrounding measures before they can shape judgment.

Torba yasa was a routine feature of Turkish lawmaking rather than an occasional tactic. Drawing on information from the Turkish parliament, BBC Türkçe reported that only two omnibus laws were passed during the AKP’s first governing period from 2002 to 2007, while twenty-three were passed between November 2015 and November 2017; by late 2017, almost one in every two laws passed in that legislative period was an omnibus law (Erem, 2017). Critics objected to the practice because the package form made separation difficult: controversial provisions could be placed beside less controversial ones, and legislators could be forced to vote up or down on a package containing both provisions they supported and provisions they opposed. The same problem matters for citizens. When reform arrives as a package, recognizing its democratic cost requires connecting particular provisions to oversight, executive discretion, or institutional constraint.

The 2010 constitutional referendum gives a clear example. The ballot placed rights-oriented reforms and judicial-governance changes inside one constitutional package. Rights-

oriented components included individual application to the Constitutional Court; oversight-relevant components changed the composition and appointment rules of key judicial bodies, including the Constitutional Court and the HSYK (Somer, 2019; Esen, 2021). Citizens therefore evaluated judicial reform as an integrated package. Support for the rights-oriented parts could coexist with weaker attention to, uncertainty about, or disagreement with the judicial-governance changes. The democratic cost entered the vote as one part of a broader reform package rather than as a separable proposal.

Other episodes show how routine legalistic reform politics became. Law No. 6524 changed the governance of the HSYK in technical legal-administrative terms. The 2014–2015 internet-law packages expanded blocking powers through privacy and public-order language. The 2015 Internal Security Package expanded police authority through a broader security reform frame. Emergency decree-laws after 2016, including Decree-Law No. 674, made lasting institutional changes under emergency authority, while the 2017 constitutional amendment package consolidated executive power through a multi-article constitutional proposal defended partly through efficiency and stability. Together, these episodes show how changes to courts, police powers, internet regulation, emergency governance, and executive authority often reached citizens through reform packages, technical legal language, emergency authority, and favorable public-facing claims (Kadioğlu, 2021; Özcan and Kimya, 2024; Esen, Gumuscu and Yavuzylmaz, 2023).

The media environment made this evaluative task harder. During the AKP period, major newspapers and television channels increasingly came under the ownership or influence of business groups close to the government. Scholars describe this transformation as media capture, involving ownership changes, regulatory pressure, legal intimidation, and state-linked advertising used to reward supportive outlets and punish critical ones (Yeşil, 2016; Coşkun, 2020; Yanatma, 2021; Akser and Baybars, 2023). The 2018 sale of Doğan Media Group to Demirören Holding was especially important because it transferred major outlets such as *Hürriyet*, CNN Türk, Kanal D, and DHA to a conglomerate widely described as government-aligned (Reporters Without Borders, 2018). As media capture deepened, citizens were more likely to encounter favorable descriptions of reform through government-aligned outlets: efficiency, security, democratization, service delivery, administrative order, or crisis response. Detailed criticism of what a package did to courts, oversight bodies, local authority, or executive discretion became harder to encounter and more costly to verify (Yanatma, 2021; Çarkoğlu and Yıldırım, 2024). In this setting, selective emphasis matters because citizens first have to identify which parts of a broader package change accountability, and those parts may be less visible than the benefits repeatedly attached to the reform.

The experiments do not try to reproduce Turkish politics as a whole. They isolate the

presentation problem that recurs in this setting: governing actors place oversight-weakening provisions inside broader reform packages, and citizens encounter those packages through language that makes some components easier to notice than others. By varying presentation while holding the relevant consequences fixed where the design permits, the studies ask whether harmful options become more acceptable when citizens encounter them as reform packages rather than as isolated institutional costs. The Turkish case makes this a demanding test because packaged reform is politically familiar. If bundled presentation and selective emphasis shape evaluation in this setting, the results are harder to attribute to novelty or confusion about the experimental format. The case also clarifies where similar dynamics should matter: legalistic backsliding settings where oversight-weakening changes are embedded in broader reform packages and communicated through publicly defensible claims.

4 Research Design Overview

The first two studies deliberately begin with abstract, incentivized lottery choices. This abstraction makes it possible to isolate evaluation from political motivations that could otherwise explain support for a harmful option, including partisan attachment, policy agreement, candidate appeal, or a willingness to accept institutional costs for other benefits. Because the lottery payoffs are known and objectively ranked, choosing a strictly worse option can be interpreted as an evaluation error rather than as a difference in political preferences. Study 1 establishes the bundling benchmark by comparing direct presentation of the final consequences with a low-complexity bundle in which those same consequences must be inferred from three changes. Study 2 then holds bundling constant and increases complexity, comparing three-change bundles with five-change bundles while also varying selective emphasis. This sequence separates the effect of requiring respondents to reconstruct the final consequences from the additional effect of increasing the complexity of that task, while allowing Study 2 to test whether higher complexity becomes more consequential for evaluation under selective emphasis.

Studies 3 and 4 then reintroduce political content while preserving the same basic presentational logic. Study 3 embeds an oversight-weakening commitment in a candidate platform and varies platform complexity and selective emphasis. Study 4 moves to a multidimensional municipal reform package, again varying complexity and selective emphasis. These political tasks trade the objective error benchmark of the lottery studies for greater political realism. Support for a candidate or reform package can reflect failed recognition, partial recognition, a willingness to accept a recognized institutional cost, or some combination of these paths. I therefore examine support and recognition together rather than treating every instance of support for the harmful political option as an evaluation error.

Across the four studies, I measure final judgments alongside what respondents identify from the task. The main outcomes are suboptimal lottery choice, harmful-bundle acceptance, support for the candidate tied to an oversight-weakening commitment, acceptance of the municipal reform package, and recognition of the harmful implications in the political tasks. I also measure perceived difficulty, confidence, time, and clicks to characterize how respondents experience the evaluative burden created by the treatments. I use these measures as supporting evidence about task engagement and difficulty rather than as separate tests of the mechanism.

Table 2: Study sequence and inferential role

Study	Setting		Main comparison	Main judgment	Inferential role
Study 1	Incentivized choice	lottery	Bundled presentation vs. direct stand-alone presentation	Choice of the lower-value option	Tests whether bundling-induced reconstruction impairs evaluation.
Study 2	Incentivized choice	lottery	Higher vs. lower complexity; selective emphasis vs. neutral presentation	Choice of the worse option and acceptance of the harmful bundle	Tests whether complexity and favorable emphasis jointly change evaluation under fixed payoffs.
Study 3	Candidate choice		Longer vs. shorter candidate platforms; favorable vs. generic summary headline for the direct-sign-off candidate	Support for the candidate tied to a harmful institutional commitment; recognition of that commitment	Tests whether platform length and headline emphasis change electoral support and recognition.
Study 4	Municipal package	reform	Two- vs. four-policy package; neutral vs. favorable policy headlines	Acceptance of the reform package; recognition of harmful consequences	Tests the logic in a multidimensional reform task closer to real policy evaluation.

The studies were fielded as modules within the same survey. Respondents therefore come from the same recruitment pool and share the same pre-treatment measures as the tasks move from abstract lotteries to political reform evaluation. The studies should be read as an experimental sequence rather than as independent replications.

4.1 Sample, Recruitment, and Data Collection

The studies use original online data from adult Turkish citizens who could complete the survey in Turkish. Respondents were recruited through paid Meta advertisements and directed to a Qualtrics survey, where they provided consent before completing the randomized experimental modules and the pretreatment measures used for heterogeneity analyses. This recruitment strategy was designed to estimate treatment effects within a recruited online sample, not to estimate population prevalence of support for any particular reform, candidate, or institutional view.

The raw data export contains 9,904 records. Of these, 2,150 respondents provided permission for data use and completed the survey. The main analytic specification excludes

responses completed in less than 450 seconds, responses longer than 6,000 seconds, and respondents who failed the pretreatment CRT-based quality threshold, producing a restricted analytic sample of 1,871 respondents. The 450-second lower threshold and CRT screen were specified in the preregistration; the 6,000-second upper threshold was added to exclude sessions likely to contain extended interruptions or unattended browser time. Because total survey duration is measured after treatment exposure, Appendix D.5 reports parallel estimates under alternative inclusion rules, including a specification that removes the upper-duration cutoff. The substantive conclusions are unchanged across these specifications.

The final restricted analytic sample contains 1,871 respondents. Study 1 has a smaller sample of 1,642 respondents because the module was introduced after the pilot wave and therefore was not administered to respondents in that wave. Studies 2–4 share the restricted analytic sample of 1,871 respondents. Appendix D.5 reports the corresponding sample sizes under the alternative inclusion rules.

The randomized modules were fielded in a fixed sequence within the same survey. To reduce confounding between task order and complexity exposure, the design counterbalanced complexity across the abstract and political modules: respondents assigned to higher-complexity presentation in the lottery-bundle module were assigned to lower-complexity presentation in the candidate and municipal modules, and respondents assigned to lower-complexity lottery bundles were assigned to higher-complexity political modules. Selective-emphasis assignment is analyzed using the study-specific treatment indicators corresponding to the presentation shown in each module. The estimates therefore use the randomized presentation variation within each study, while the appendix reports balance and assignment diagnostics separately for each article-labeled study.

The main design and analysis plan were preregistered on OSF on March 28, 2026, under the registration titled *Unbundling Autocratic Capture* (OSF registry code `py2xc`; project link: osf.io/q4su5). The study was reviewed under NYU IRB protocol IRB-FY2025-10304. Rather than listing every preregistered design element here, I report the relevant departures from the registration directly. One sample-construction difference is the treatment of post-treatment duration restrictions: the main restricted sample remains the main specification, while Appendix D.5 reports estimates under alternative inclusion rules to assess whether the duration rule affects the results. Appendix D.13 summarizes how the analyses reported in the article map onto the preregistration, identifies departures from the registered plan, and notes preregistered components that are reported only in the appendix or fall outside the scope of the present article.

Pilot waves were used to inform instrument revisions. The main analyses pool only waves and modules whose assignment structure, treatment conditions, and outcome measures are

comparable.

The estimand is the average treatment effect within the recruited survey sample: the average difference between the potential outcome under one presentation and the potential outcome under the comparison presentation. Random assignment identifies this estimand by making the assigned presentation independent of respondents’ pretreatment characteristics. I therefore interpret the estimates as internally valid treatment effects for this sample, while treating external representativeness as a scope condition rather than a design assumption.

The survey recorded both final judgments and process measures. The final judgments are whether respondents choose the worse lottery option, accept a harmful lottery bundle, support a candidate tied to a harmful institutional commitment, accept the municipal reform package, and recognize the harmful implication of the proposal. The process measures are time spent on task, clicks for additional information, perceived difficulty, and confidence. I use these measures as supporting evidence about task engagement and difficulty. Full process-outcome estimates are reported in Appendix D.10.

4.2 Cognitive Types

The capacity analyses use a pretreatment cognitive-performance measure specified in the preregistration. The five Cognitive Reflection Test (CRT) items were administered before treatment assignment and before respondents entered the randomized experimental modules. They therefore cannot be affected by the complexity or selective-emphasis treatments. Two items are used as basic pretreatment screen items, and three items are harder reasoning items. Appendix D.12 reports the exact item wording, correct answers, and coding rule.

The preregistered CRT-based inclusion rule excludes respondents who missed both basic pretreatment screen items. This exclusion changes the respondent population to which the restricted-sample estimates apply, but because the items were measured before treatment assignment, it does not create treatment-induced selection. Among retained respondents, I classify respondents as higher-capacity if they answered both basic items correctly and at least one of the three harder reasoning items correctly. Respondents who passed the basic-CRT-item inclusion rule but do not meet this higher-capacity threshold are classified as lower-capacity.

I use this distinction only to examine heterogeneity in treatment response. It is not meant to rank respondents more broadly. The measure captures a pretreatment difference in performance on reasoning items relevant to the theory: how costly it is to identify the consequences of a package when those consequences are spread across several components. Because the preregistration treated the lower-cognition joint pattern as less certain in sign,

I interpret subgroup estimates as evidence about where the evaluative burden binds most clearly rather than as the paper's main confirmatory evidence.

5 Study 1: Bundling Benchmark

5.1 Design

Study 1 begins with the simplest version of the problem. Before introducing political content, selective emphasis, or variation in bundle size, it asks whether bundling itself creates enough evaluative complexity to reduce respondents’ ability to identify the better option. In the direct-comparison baseline, respondents see each lottery’s final composition. In the bundled conditions, one option must instead be reconstructed from a starting bowl and three changes (see Table 3). In this lottery task, reconstruction means calculating the final bowl implied by the starting bowl and the displayed changes before comparing it with the directly shown alternative. The treatment therefore captures the complexity introduced by bundling: the relevant consequences are no longer directly visible, even though the bundle is short and its size is held fixed. Study 2 then asks whether this problem deepens when the bundle itself becomes larger and respondents must process more changes.

Respondents choose between two bowls of colored balls. After they choose, one ball is drawn at random, and the color of the drawn ball determines their experimental credits: Purple pays 60 credits, Blue pays 30 credits, and Gray pays zero. Each option ultimately produces a final bowl with 20 balls. One final bowl is strictly better: it contains 4 Purple, 5 Blue, and 11 Gray balls, for an expected value of 19.5 credits. The other is strictly worse: it contains 3 Purple, 2 Blue, and 15 Gray balls, for an expected value of 12 credits. The lower-value bowl is not a safer alternative or a trade-off against risk; it has fewer high-paying balls, fewer intermediate-paying balls, and more zero-paying balls. A respondent therefore makes a suboptimal choice when they select the lower expected-value bowl.

The three treatment arms vary whether those same final lotteries are directly visible or must be understood from components. In the direct-comparison baseline, both final bowls are shown directly. In the lower-value bundled condition, the better lottery is shown directly, while the worse lottery is shown as a starting bowl plus three changes that produce the worse final bowl. In the higher-value bundled condition, the worse lottery is shown directly, while the better lottery is shown as a starting bowl plus three changes that produce the better final bowl. Table 3 summarizes the comparison. Full condition details and payoff calculations are reported in Appendix Tables 14 and 15; the full survey screen is reported in Appendix C.

This design distinguishes directional reactions to bundled presentation from an evaluation problem created by reconstruction. If respondents simply favor bundled options, they should become more likely to choose the bundled alternative regardless of whether it is better or worse. If they avoid bundled options because they dislike complexity, they should become

less likely to choose the bundled alternative regardless of its value. By contrast, if bundling makes evaluation harder by forcing respondents to reconstruct the final bowl, it should reduce choice quality in both directions: respondents should become more likely to accept a worse bundled option and less likely to identify a better bundled option.

Table 3: Study 1 treatment arms and treatment presentations

Treatment arm	Strictly better option	Strictly worse option
Direct comparison	Direct final bowl: 4P, 5B, 11G. Expected value = 19.5.	Direct final bowl: 3P, 2B, 15G. Expected value = 12.
Lower-value bundled	Direct final bowl: 4P, 5B, 11G. Expected value = 19.5.	Bundled presentation. <i>Start:</i> 4P, 5B, 11G. <i>Changes:</i> -2G, +3B, -2P; +4P, +1G, -4B; -2B, -3P, +5G. <i>Implied final:</i> 3P, 2B, 15G. Expected value = 12.
Higher-value bundled	Bundled presentation. <i>Start:</i> 3P, 2B, 15G. <i>Changes:</i> +2B, +3P, -5G; -4P, -1G, +4B; +2G, -3B, +2P. <i>Implied final:</i> 4P, 5B, 11G. Expected value = 19.5.	Direct final bowl: 3P, 2B, 15G. Expected value = 12.

Note: P = Purple, B = Blue, and G = Gray. The column labels identify the payoff ranking for the reader; respondents saw generic option labels rather than “better” and “worse.” Option placement was randomized. In the bundled conditions, respondents saw a starting bowl and a sequence of three changes, then had to infer the final bowl before comparing it with the directly shown alternative. The lower expected-value option is strictly worse, not a safer alternative: it contains fewer high-payoff balls, fewer medium-payoff balls, and more zero-payoff balls.

The main outcome is suboptimal choice. I estimate linear probability models in which the dependent variable equals one when the respondent chooses the lower expected-value lottery. The main specification compares the bundled arms to the direct-comparison baseline and includes the preregistered risk-aversion measure as a covariate. Because each respondent contributes one Study 1 decision, I use heteroskedasticity-robust standard errors and report effects in percentage points. I also examine perceived difficulty, confidence, time, and clicks as secondary process measures that characterize how respondents experienced the task.³

³Process outcomes are estimated with the same treatment specification. Treatment effects are reported in standardized outcome units; the Baseline column reports difficulty and confidence on their original 1–7 scales, time in the original recorded units, and clicks as raw counts. Full process-outcome estimates are reported in Appendix D.10. Because these measures are secondary and may reflect several aspects of task experience, I use them to characterize engagement and difficulty rather than as separate tests of the mechanism.

5.2 Results

Study 1 shows that bundling alone is enough to impair evaluation even when the underlying choice contains no payoff trade-off. In the direct-comparison baseline, where both options are visible as simple lotteries, 26.4 percent of respondents choose the objectively inferior option. When one of the same final lotteries must instead be inferred from a three-change bundle, suboptimal choice increases by 17.0 percentage points, with a 95 percent confidence interval from 12.2 to 21.8 points.

Figure 1 shows that when the lower-value option is bundled, respondents become more likely to accept it: suboptimal choice rises by 14.0 percentage points (95% CI: 8.3, 19.6). When the higher-value option is bundled, respondents become less likely to identify it as the better option: suboptimal choice rises by 20.0 percentage points (95% CI: 14.4, 25.7). Bundling therefore does more than make respondents favorable toward whichever option is bundled. It makes them worse at distinguishing better from worse options when the payoff-relevant final bowl must be inferred rather than inspected directly.

The capacity pattern is especially informative. Higher-capacity respondents perform much better in the direct-comparison baseline, but their advantage shrinks sharply once an option is bundled. For this group, bundling the lower-value option increases suboptimal choice by 22.1 percentage points (95% CI: 13.3, 30.8), and bundling the higher-value option increases suboptimal choice by 24.2 percentage points (95% CI: 15.4, 33.1). Among lower-capacity respondents, the corresponding estimates are 8.8 percentage points (95% CI: 1.4, 16.2) and 17.2 percentage points (95% CI: 9.7, 24.6). The smaller lower-capacity estimates should be read against a much higher baseline error rate. Higher-capacity respondents have more room to fall because they begin from a more accurate baseline.

This benchmark establishes the paper’s first claim: when citizens must infer the consequences of a package, the format of presentation can change the quality of evaluation before any political content, selective emphasis, or partisan signal enters the task. Bundling affects more than respondents who are already prone to error. It can also erode the advantage of respondents who evaluate simple alternatives accurately.

The process outcomes point in the same direction. Bundled options lead respondents to spend more time, click more, report greater difficulty, and express lower confidence. These patterns support an evaluative-burden interpretation: respondents are not merely ignoring the task, but the additional effort does not fully offset the difficulty of aggregating the option mentally. Study 1 supplies the baseline for the later experiments by showing that bundling alone creates an evaluative vulnerability. Study 2 asks whether that vulnerability deepens when the bundle becomes more complex and attractive components receive selective

emphasis.

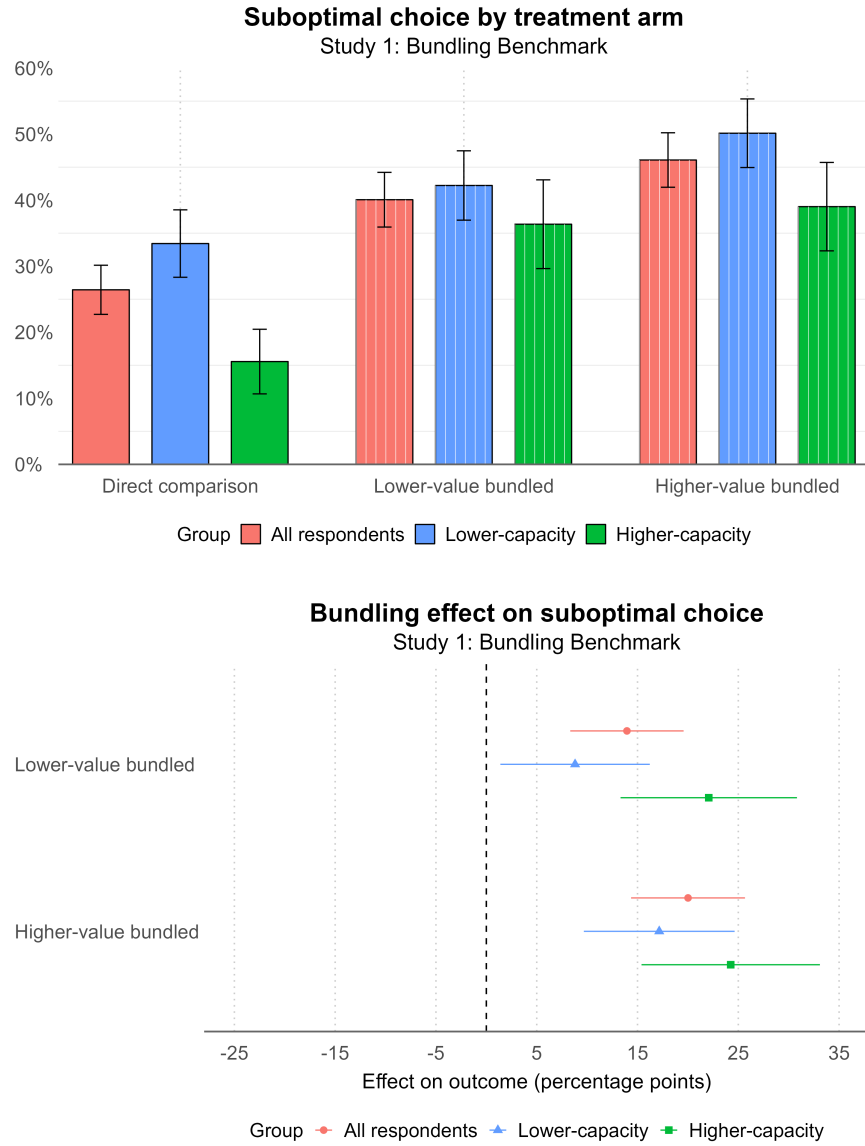


Figure 1: Study 1

Note: The analytic sample contains 1,642 respondents. The top panel reports treatment-arm means for suboptimal choice, coded as one when the respondent chooses the lower expected-value lottery. The bottom panel reports linear probability model estimates relative to the direct-comparison baseline. The specification includes the preregistered risk-aversion measure as a covariate; standard errors are heteroskedasticity-robust. Baseline suboptimal-choice rates in the direct-comparison condition are 26.4 percent among all respondents, 15.6 percent among higher-capacity respondents, and 33.4 percent among lower-capacity respondents.

6 Study 2: Complexity and Selective Emphasis

6.1 Design

Study 2 keeps the choice task abstract but varies how difficult it is to infer the value of a bundled option. Study 1 showed that bundling can impair evaluation because one option must be inferred from a starting point *and* a sequence of changes, raising the complexity and cost of evaluation. Study 2 asks whether that problem worsens when the inference task becomes more demanding and attractive changes are selectively emphasized.

Respondents again make incentivized lottery choices. In each task, they compare a directly shown status quo bowl with one bundled alternative. The bundled alternative is shown as a starting bowl plus a sequence of rule changes that respondents must use to infer the final bowl. The bundle can be harmful, meaning that its final outcome is strictly worse than the directly shown alternative, or beneficial, meaning that its final outcome is strictly better. This structure keeps the payoff ranking objective: there is no induced payoff trade-off between the two options.

The design varies two features of the bundled alternative. The first is complexity. Study 2 manipulates complexity by changing the number of rule lines respondents must process. Low-complexity bundles contain three rule lines; high-complexity bundles contain five. In both cases, respondents must infer the final payoff-relevant bowl from all rule changes before comparing it with the directly shown alternative.

The second feature is selective emphasis. In the neutral condition, all rule changes are displayed without visual emphasis. In the selective-emphasis condition, attractive changes are highlighted with bold, colored font while the full rule list remains available. This manipulation does not change the final payoff-relevant bowl. It changes which parts of the bundle are easiest to notice while respondents infer the outcome.

Table 4 summarizes the treatment structure. In every task, respondents compare a directly shown status-quo bowl with one bundled alternative. The treatment varies whether that bundled alternative is described with three or five rule lines and whether attractive changes are visually emphasized. In the harmful-bundle task, the question is whether a net-worse bundle becomes more acceptable when the inference task is harder and favorable changes are highlighted. In the beneficial-bundle task, the question is whether a strictly better bundle becomes harder to identify under the same presentation conditions.

Table 4: Study 2 treatment cells and rule changes

Treatment cell	Directly shown alternative	Rule changes shown for bundled alternative	Implied final bundle
Low complexity / neutral / harmful bundle	Task A: 2P, 6B, 12G. Task B: 2P, 4B, 14G.	1. +2B, -1P, +3G. 2. +2P, +3G, -5B. 3. -4G, +2B, -2P.	Task A: 1P, 5B, 14G. EV = 10.5. Task B: 1P, 3B, 16G. EV = 7.5.
Low complexity / selective emphasis / harmful bundle	Task A: 2P, 6B, 12G. Task B: 2P, 4B, 14G.	1. +2B , -1P, +3G. 2. +2P , +3G, -5B. 3. -4G , +2B , -2P.	Task A: 1P, 5B, 14G. EV = 10.5. Task B: 1P, 3B, 16G. EV = 7.5.
High complexity / neutral / harmful bundle	Task A: 2P, 6B, 12G. Task B: 2P, 4B, 14G.	1. +2B, -1P, +3G. 2. +3P, -1B, +2G. 3. +2P, +3G, -5B. 4. +1B, -2G, -3P. 5. -4G, +2B, -2P.	Task A: 1P, 5B, 14G. EV = 10.5. Task B: 1P, 3B, 16G. EV = 7.5.
High complexity / selective emphasis / harmful bundle	Task A: 2P, 6B, 12G. Task B: 2P, 4B, 14G.	1. +2B , -1P, +3G. 2. +3P , -1B, +2G. 3. +2P , +3G, -5B. 4. +1B , -2G , -3P. 5. -4G , +2B , -2P.	Task A: 1P, 5B, 14G. EV = 10.5. Task B: 1P, 3B, 16G. EV = 7.5.
Low complexity / neutral / beneficial bundle	Task A: 2P, 6B, 12G. Task B: 2P, 4B, 14G.	1. +3B, -6G, -1P. 2. +1P, -3B, +1G. 3. +2G, +1P, +2B.	Task A: 3P, 8B, 9G. EV = 21.0. Task B: 3P, 6B, 11G. EV = 18.0.
Low complexity / selective emphasis / beneficial bundle	Task A: 2P, 6B, 12G. Task B: 2P, 4B, 14G.	1. +3B , -6G , -1P. 2. +1P , -3B, +1G. 3. +2G, +1P , +2B .	Task A: 3P, 8B, 9G. EV = 21.0. Task B: 3P, 6B, 11G. EV = 18.0.
High complexity / neutral / beneficial bundle	Task A: 2P, 6B, 12G. Task B: 2P, 4B, 14G.	1. +3B, -6G, -1P. 2. +1P, -1B, +2G. 3. +1P, -3B, +1G. 4. -2G, -1P, +1B. 5. +2G, +1P, +2B.	Task A: 3P, 8B, 9G. EV = 21.0. Task B: 3P, 6B, 11G. EV = 18.0.
High complexity / selective emphasis / beneficial bundle	Task A: 2P, 6B, 12G. Task B: 2P, 4B, 14G.	1. +3B , -6G , -1P. 2. +1P , -1B, +2G. 3. +1P , -3B, +1G. 4. -2G , -1P, +1B . 5. +2G, +1P , +2B .	Task A: 3P, 8B, 9G. EV = 21.0. Task B: 3P, 6B, 11G. EV = 18.0.

Note: P = Purple, B = Blue, and G = Gray. Purple pays 60 credits, Blue pays 30 credits, and Gray pays zero. The directly shown alternative is the status-quo bowl. The bundled alternative starts from the same bowl and then applies the listed rule changes. Neutral and selective-emphasis versions use the same rule lists and imply the same final bowls. In the selective-emphasis condition, favorable changes were shown in bold colored font; all other changes were shown in regular black font.

The main outcomes are suboptimal choice and harmful bundle acceptance. Suboptimal

choice equals one when the respondent chooses the lower expected-value option. Harmful bundle acceptance equals one when the respondent accepts a bundled alternative that is worse than the directly shown status quo. I estimate linear probability models with indicators for high complexity, selective emphasis, and their interaction, using the low-complexity neutral condition as the baseline. Because respondents contribute repeated lottery decisions in Study 2, standard errors are clustered by respondent. Effects are reported in percentage points. I also examine perceived difficulty, confidence, time, and clicks as secondary process measures.⁴

6.2 Results

Study 2 shows that the complexity-by-emphasis pattern is concentrated among lower-capacity respondents. In the full sample, the interaction estimates are small because the lower- and higher-capacity patterns move in different directions. The aggregate estimate therefore does not summarize the main pattern well. I report the full-sample contrasts first, then turn to the preregistered capacity groups.

In the low-complexity, neutral-emphasis baseline, 41.5 percent of respondents in the full sample choose the suboptimal option and 42.0 percent accept the harmful bundle (see Figure 2). High complexity alone produces little change in either outcome in the full sample. The sharper pattern emerges in how complexity combines with selective emphasis, particularly among lower-capacity respondents. In the full sample, the complexity-by-selective-emphasis interaction is 3.7 percentage points for suboptimal choice (95% CI: -2.4, 9.8) and 4.8 percentage points for harmful bundle acceptance (95% CI: -4.1, 13.8). These interaction terms are difference-in-differences estimates: they ask whether the effect of high complexity is larger under selective emphasis than under neutral presentation.

The lower-capacity subgroup shows the clearest pattern. Among lower-capacity respondents, the interaction on suboptimal choice is 9.7 percentage points (95% CI: 2.1, 17.4). This means that high complexity increases suboptimal choice more under selective emphasis than under neutral presentation. The harmful-bundle outcome shows the same pattern. The interaction on harmful bundle acceptance is 12.6 percentage points (95% CI: 1.2, 24.1). In substantive terms, added complexity makes lower-capacity respondents more likely to accept a harmful bundle when attractive changes are highlighted.

⁴Process outcomes are estimated with the same treatment specification. Treatment effects are reported in standardized outcome units; the Baseline column reports difficulty and confidence on their original 1–7 scales, time in the original recorded units, and clicks as raw counts. Full process-outcome estimates are reported in Appendix D.10. Because these measures are secondary and may reflect several aspects of task experience, I use them to characterize engagement and difficulty rather than as separate tests of the mechanism.

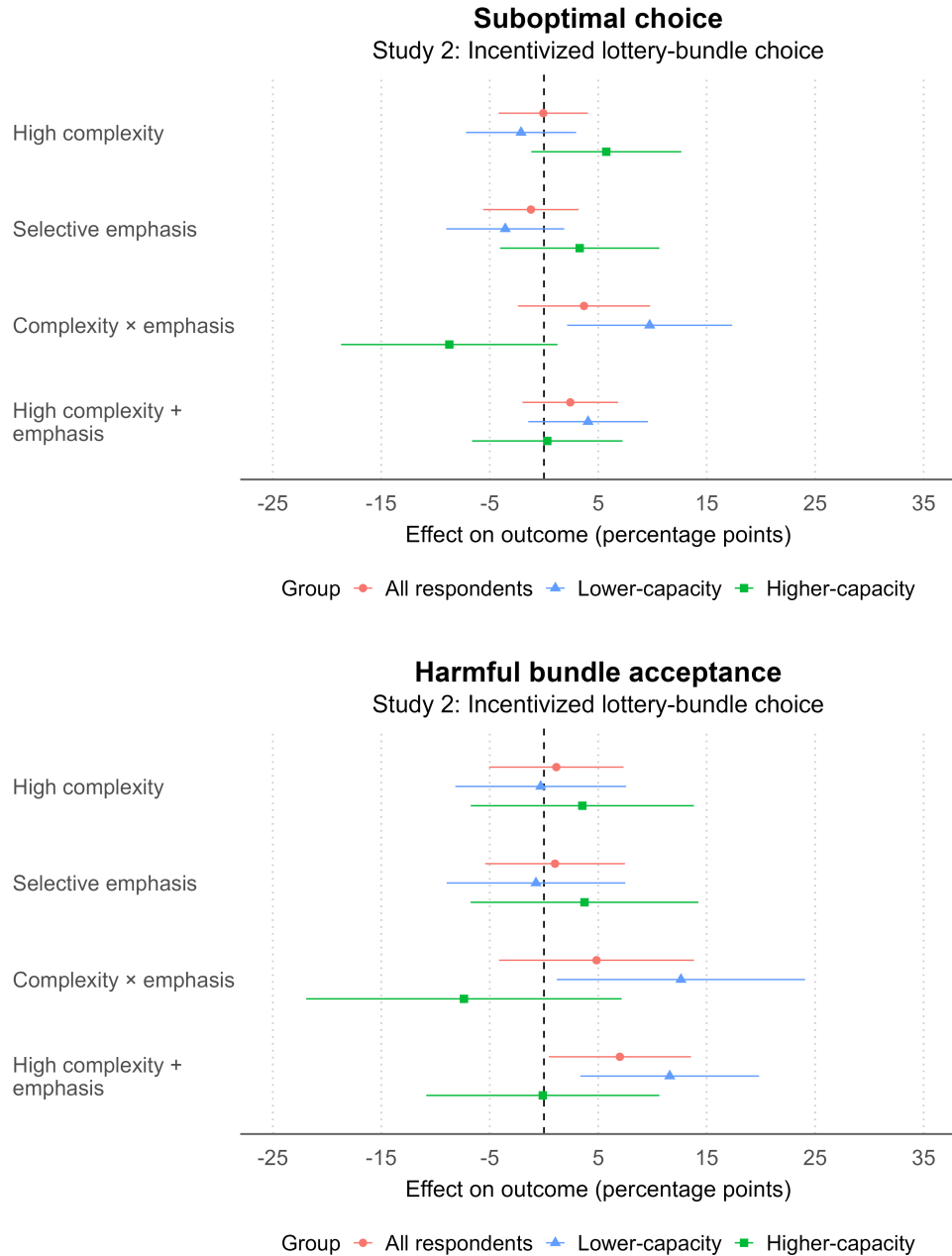


Figure 2: Study 2

Note: The analytic sample contains 1,871 respondents. Outcomes are coded as binary indicators and effects are reported in percentage points. The top panel reports treatment effects on suboptimal choice, coded as one when the respondent chooses the lower expected-value option. The bottom panel reports treatment effects on harmful bundle acceptance, coded as one when the respondent accepts a bundled alternative that is worse than the directly shown status quo. Linear probability models include indicators for high complexity, selective emphasis, and their interaction, with the low-complexity neutral condition as the baseline. Standard errors are clustered by respondent. Baseline rates are from the low-complexity, neutral-emphasis, no-advisor condition, matching the reported no-advisor simple-effect estimand. For suboptimal choice, baseline rates are 41.5 percent among all respondents, 32.8 percent among higher-capacity respondents, and 46.4 percent among lower-capacity respondents. For harmful bundle acceptance, baseline rates are 42.0, 36.7, and 45.1 percent, respectively.

The combined treatment contrast points in the same direction. Relative to the low-complexity neutral condition, the high-complexity selective-emphasis treatment increases harmful bundle acceptance among lower-capacity respondents by 11.6 percentage points (95% CI: 3.4, 19.8). This comparison is easier to read than the interaction: lower-capacity respondents are more likely to accept the harmful bundle when they face both a more complex inference task and highlighted attractive changes.

The higher-capacity pattern is different. These respondents are less moved by the interaction, and in some specifications the interaction moves in the opposite direction. That divergence helps explain why the full-sample interaction is small. The result therefore suggests a capacity-specific heterogeneity pattern: the combination of added complexity and selective emphasis matters most among lower-capacity respondents.

The process outcomes are informative but should be interpreted cautiously. I do not find clear evidence that added complexity produced a consistent increase in time, clicks, or perceived difficulty in Study 2. Clicks rise modestly with complexity in the suboptimal-choice task, but the pattern is not strong enough to support a separate search-effort account. Confidence provides the most suggestive process evidence: it declines most clearly in the high-complexity selective-emphasis condition for the suboptimal-choice task, especially among lower-capacity respondents. I therefore treat the process outcomes as supporting context rather than as a separate test of the mechanism.

Study 2 supports the argument on a specific margin. In a controlled lottery setting with objective payoffs, added complexity and selective emphasis make harmful bundles more acceptable among lower-capacity respondents. The study does not show that all bundle evaluation deteriorates in the same way, and it does not show a large full-sample interaction. Its contribution is narrower but important: when a harmful bundle must be inferred from several changes, highlighting attractive parts can make that bundle more acceptable among respondents for whom the task is most demanding. The next studies ask whether the same presentational logic appears in political evaluation.

7 Study 3: Candidate Choice

7.1 Design

Study 3 asks whether the same bundle-evaluation problem appears in electoral choice. Respondents compare two candidates running on proposals about small-business permits and inspections. The political studies deliberately use local administrative settings rather than nationally salient reforms. A national-level vignette in Türkiye would be more likely to activate prior knowledge, established partisan attachments, and existing views about prominent reforms, making it harder to separate responses to the experimental presentation from responses to familiar political actors or issues. The local setting preserves the institutional stakes central to the argument—executive discretion, procedural review, and public accountability—while reducing these preexisting national cues. Permitting, inspections, and discretionary approval also provide a setting in which speed and simplification can be attractive while weaker review and less public documentation increase executive discretion.

The candidates operate in the same policy domain, but one carries the oversight-weakening commitment. That candidate proposes faster permit approval through mayoral direct sign-off without committee review. The comparison candidate preserves committee approval and published minutes. The design therefore makes the institutional contrast concrete: one candidate offers speed and administrative flexibility while weakening review, while the other preserves a more procedural approval process.

The experiment varies two features of presentation. The first is complexity. In the low-complexity condition, respondents see two short three-item candidate platforms. In the high-complexity condition, both candidates' platforms expand to eight items. Five administrative items are added to each platform using closely matched language adapted to the candidate's proposal, while the original three core items remain unchanged. The institutional contrast therefore remains fixed across complexity conditions: one candidate proposes permit approval through the mayor's direct sign-off without committee review, while the comparison candidate preserves committee approval with published minutes. Complexity changes the amount of surrounding information respondents must process without changing the core institutional difference between the candidates.

Table 5: Study 3 treatment arms and full text of candidate platforms

Treatment arm	Candidate with direct mayoral sign-off	Comparison candidate
Low complexity / neutral	At a glance: Municipal Permit and Inspection Plan Set a clear timeline for permit decisions. Use an expedited lane for some applications by making exceptions. Approve permits through the mayor’s direct sign-off, without a committee.	At a glance: Municipal Permit and Inspection Plan Establish a fixed schedule for permit processing. Offer a fast option for applications that meet preset standards. Approve permits through a committee with published minutes.
Low complexity / selective emphasis	At a glance: Clear deadlines for permit decisions and Faster processing for routine applications Set a clear timeline for permit decisions. Use an expedited lane for some applications by making exceptions. Approve permits through the mayor’s direct sign-off, without a committee.	At a glance: Municipal Permit and Inspection Plan Establish a fixed schedule for permit processing. Offer a fast option for applications that meet preset standards. Approve permits through a committee with published minutes.
High complexity / neutral	At a glance: Municipal Permit and Inspection Plan Use a single standard application form for permits. Use a checklist to reduce missing documents. Assign a case owner for each application. Set internal time limits for department feedback. Use a shared tracking system to follow applications. Set a clear timeline for permit decisions. Use an expedited lane for some applications by making exceptions. Approve permits through the mayor’s direct sign-off, without a committee.	At a glance: Municipal Permit and Inspection Plan Use a standard application form for permits. Use a checklist to reduce missing documents. Assign a point person for each application. Set internal time limits for departments to respond. Use a shared record system to track applications. Establish a fixed schedule for permit processing. Offer a fast option for applications that meet preset standards. Approve permits through a committee with published minutes.
High complexity / selective emphasis	At a glance: Clear deadlines for permit decisions and Faster processing for routine applications Use a single standard application form for permits. Use a checklist to reduce missing documents. Assign a case owner for each application. Set internal time limits for department feedback. Use a shared tracking system to follow applications. Set a clear timeline for permit decisions. Use an expedited lane for some applications by making exceptions. Approve permits through the mayor’s direct sign-off, without a committee.	At a glance: Municipal Permit and Inspection Plan Use a standard application form for permits. Use a checklist to reduce missing documents. Assign a point person for each application. Set internal time limits for departments to respond. Use a shared record system to track applications. Establish a fixed schedule for permit processing. Offer a fast option for applications that meet preset standards. Approve permits through a committee with published minutes.

Note: Candidate order was randomized. The column labels identify the institutional contrast for the reader; respondents saw generic candidate labels. Complexity varies the length of both candidate platforms: both contain three items in the low-complexity condition and eight items in the high-complexity condition. The five additional items in the high-complexity condition describe closely matched administrative procedures, with minor wording differences adapted to each candidate’s proposal. Within each complexity condition, the underlying platform items are held fixed across neutral and selective-emphasis arms. In the neutral condition, both candidates receive the generic “At a glance: Municipal Permit and Inspection Plan” headline. In the selective-emphasis condition, the comparison candidate retains that headline, while the direct-sign-off candidate instead receives an “At a glance” headline emphasizing clear deadlines and faster processing for routine applications. In Turkish, “At a glance” was presented as “Kısaca.” The institutional contrast between mayoral direct sign-off without committee review and committee approval with published minutes remains unchanged across treatment arms.

The second feature is selective emphasis. Each candidate is presented with an “At a glance” headline. In the neutral condition, both candidates receive the generic headline “Municipal Permit and Inspection Plan.” In the selective-emphasis condition, the comparison candidate retains that generic headline, while the candidate proposing mayoral direct sign-off receives a favorable headline emphasizing “Clear deadlines for permit decisions and Faster processing for routine applications.” Within each complexity condition, the underlying platform items are held fixed across neutral and selective-emphasis arms. Selective emphasis therefore changes how the direct-sign-off candidate’s platform is summarized without changing the proposals themselves. In Turkish, “At a glance” was presented as “Kısaca.”

The main outcomes are harmful-option selection and recognition. Harmful-option selection is coded as one when the respondent supports the candidate proposing direct mayoral sign-off without committee review. This outcome is the political analogue of choosing or accepting a harmful option in the lottery studies. Recognition is coded as one when the respondent correctly identifies which candidate proposed approving permits through mayoral direct sign-off without a committee. I also examine perceived difficulty, confidence, time, and clicks as secondary process measures.

I estimate linear probability models with indicators for high complexity, selective emphasis, and their interaction, using the low-complexity neutral condition as the baseline. The models adjust for module order, municipal-module order, and survey date. Because each respondent contributes one Study 3 candidate-choice observation, I use heteroskedasticity-robust standard errors and report effects in percentage points. Recognition is a binary outcome and is analyzed with the same treatment specification.⁵

7.2 Results

When both candidates are shown with longer platforms and the candidate proposing direct mayoral sign-off receives the favorable summary headline, harmful-option selection rises and recognition of the institutional commitment falls. Both outcomes move in the direction predicted by the theory: respondents become more willing to support the candidate tied to weaker oversight, while they become less likely to identify the direct-sign-off commitment.

⁵Process outcomes are estimated with the same treatment specification. Treatment effects are reported in standardized outcome units; the Baseline column reports difficulty and confidence on their original 1–7 scales, time in the original recorded units, and clicks as raw counts. Full process-outcome estimates are reported in Appendix D.10. Because these measures are secondary and may reflect several aspects of task experience, I use them to characterize engagement and difficulty rather than as separate tests of the mechanism.

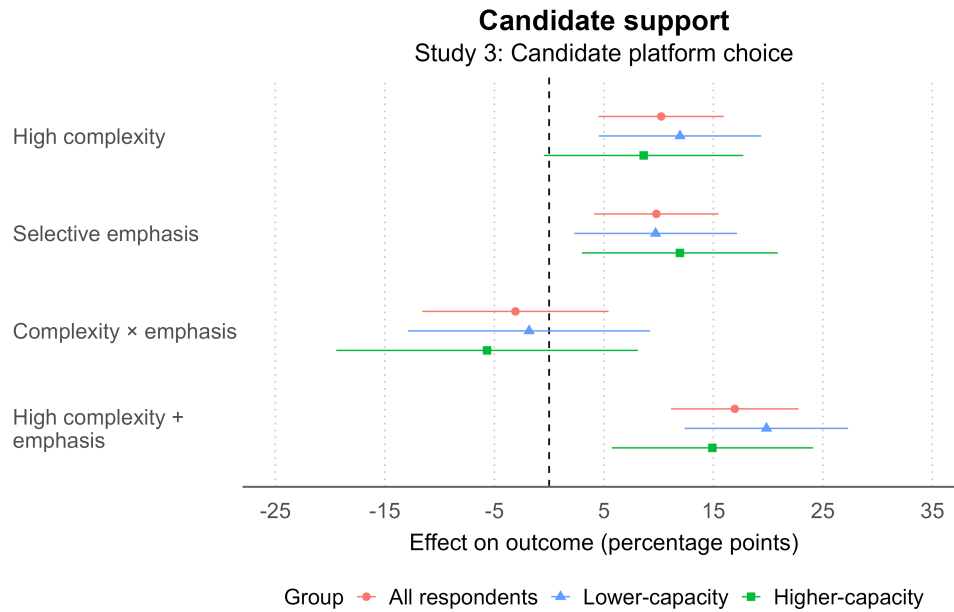
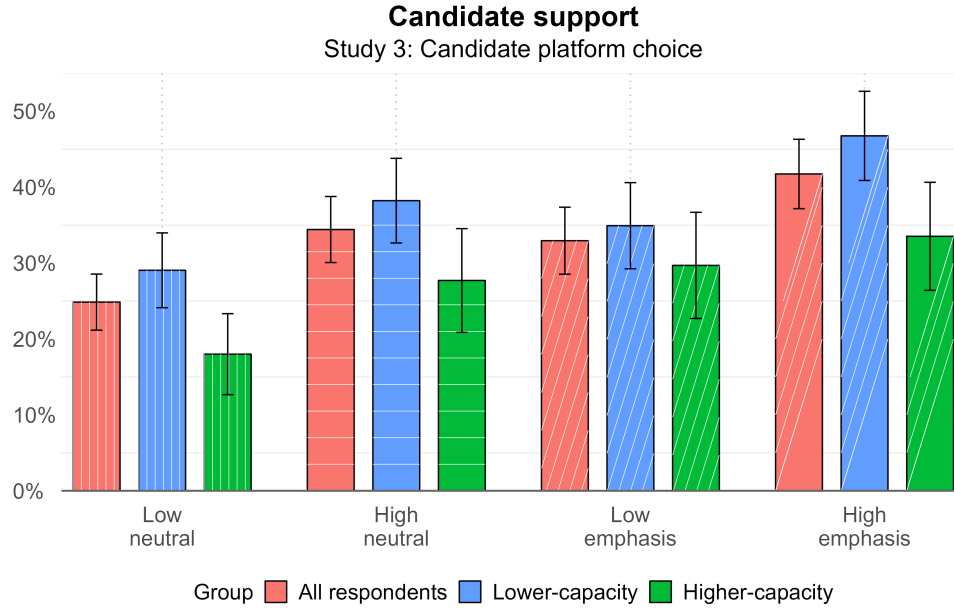


Figure 3: Study 3: Candidate support.

Note: The analytic sample contains 1,871 respondents. The top panel reports treatment-arm means for harmful-option selection, coded as one when the respondent supports the candidate proposing mayoral direct sign-off without committee review. The bottom panel reports linear probability model estimates relative to the low-complexity neutral baseline. The specification includes indicators for high complexity, selective emphasis, and their interaction, and adjusts for module order, municipal-module order, and survey date. Standard errors are heteroskedasticity-robust. Baseline candidate-support rates in the low-complexity, neutral-emphasis condition are 24.9 percent among all respondents, 18.0 percent among higher-capacity respondents, and 29.1 percent among lower-capacity respondents.

Recognition moves in the opposite direction. Relative to the low-complexity neutral baseline, the joint high-complexity and selective-emphasis condition reduces correct recognition of the oversight-weakening commitment by 7.9 percentage points (95% CI: -14.1, -1.6; Figure 4).

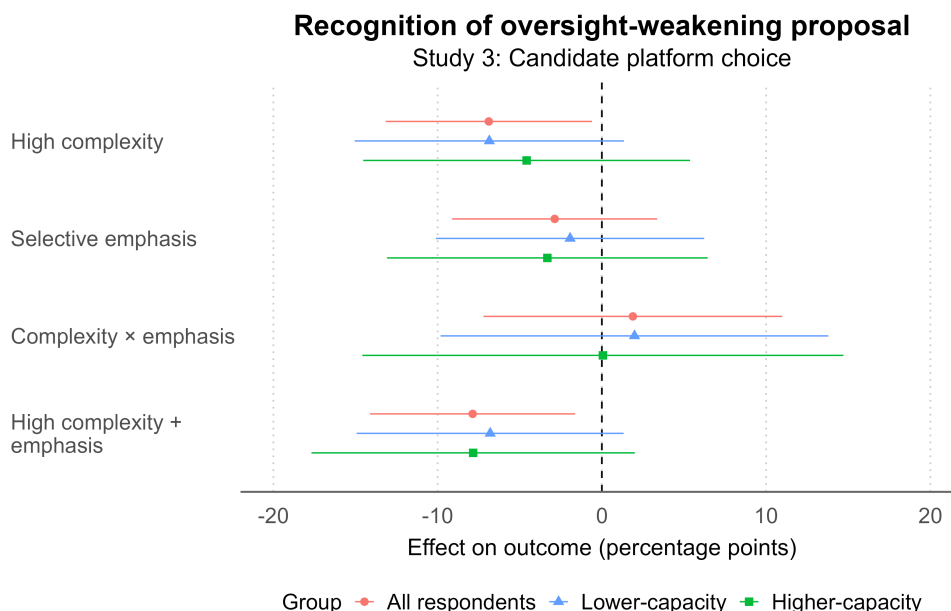


Figure 4: Study 3: Recognition of oversight-weakening proposal.

Note: The analytic sample contains 1,871 respondents. Recognition is coded as one when the respondent correctly identifies which candidate proposed approving permits through mayoral direct sign-off without a committee. The figure reports linear probability model estimates relative to the low-complexity neutral baseline. The specification includes indicators for high complexity, selective emphasis, and their interaction, and adjusts for module order, municipal-module order, and survey date. Standard errors are heteroskedasticity-robust. Baseline recognition rates in the low-complexity, neutral-emphasis condition are 63.4 percent among all respondents, 72.0 percent among higher-capacity respondents, and 58.1 percent among lower-capacity respondents.

In the low-complexity neutral baseline, 24.9 percent of respondents support the candidate proposing direct mayoral sign-off (see Figure 3). Relative to that baseline, high complexity increases support by 10.2 percentage points (95% CI: 4.5, 15.9), while selective emphasis increases support by 9.8 percentage points (95% CI: 4.1, 15.5). When both features are present, support rises by 17.0 percentage points (95% CI: 11.1, 22.8). The joint condition produces the largest shift in support, but the two presentational features do not amplify one another. The complexity-by-selective-emphasis interaction is -3.1 percentage points (95% CI: -11.6, 5.4).

Appendix D.8.9 further unpacks this decline by distinguishing correct identification, identification of the wrong candidate, and uncertainty. The decline in correct recognition is driven primarily by an increase in uncertainty rather than by respondents shifting toward the wrong candidate. These diagnostics are descriptive rather than a causal mediation analysis: they show how recognition changes under treatment, but they cannot establish how much of the increase in candidate support operates through recognition. Descriptively, support for the direct-sign-off candidate also rises among respondents who correctly identify the institutional commitment, indicating that failed recognition alone does not account for the overall support pattern.

The process outcomes help rule out a simple disengagement interpretation. High complexity increases time spent and clicks, and the joint condition also raises perceived difficulty. Confidence falls under complexity in the full sample, especially among lower-capacity respondents, although the joint confidence effect is weaker. Respondents therefore do not appear to ignore the longer platform. They spend more time with it, interact more with it, and experience it as more demanding. Yet harmful-option selection rises and recognition falls under the joint presentation. The candidate platform changes what respondents take away from the choice even when it induces more engagement rather than less.

Study 3 therefore extends the argument from direct bundle evaluation to candidate choice. A candidate tied to weaker oversight becomes more electorally acceptable when the choice involves longer platforms and that candidate receives a favorable summary headline. Harmful-option selection and recognition each move in the predicted direction: support for the direct-sign-off candidate rises, while identification of the direct-sign-off commitment declines.

8 Study 4: Municipal Reform Package

8.1 Design

Study 4 brings the argument back to direct reform evaluation in a setting closer to ordinary local politics. Respondents read a municipal reform package that affects employment, service capacity, and discretionary allocation. As in the lottery tasks, each component changes the consequences of the option, and respondents must combine those component-level changes before deciding whether to accept the package.

The package deliberately mixes attractive administrative aims with costly institutional consequences. Some components promise faster delivery, investment, or revitalization. Others increase administrative discretion, or reduce service capacity or employment opportunities. This structure mirrors the reform bundles that motivate the paper: institutional costs do not appear alone, but alongside provisions that may be useful, publicly defensible, or attractive on their face.

The experiment varies two features of presentation. The first is complexity. In the low-complexity condition, respondents see a shorter two-policy version of the package. In the high-complexity condition, they see the full four-policy version. The longer package asks respondents to track more moving parts and integrate consequences across several domains.

The second feature is selective emphasis. Within each complexity condition, the same measures appear in the neutral and selective-emphasis arms. What changes is how those measures are labeled. The neutral version uses more technocratic titles, while the selective-emphasis version uses favorable policy headlines that emphasize attractive aspects of the package, such as speed, revitalization, and delivery. The treatment therefore changes which components are easiest to notice, not the underlying measures within a given complexity level.

Table 6 summarizes the measures shown in each arm. Low-complexity respondents saw Measures 1 and 3; high-complexity respondents saw all four measures. The added measures in the high-complexity condition increase the amount of information respondents must process, but they do not change the package’s net consequences across the three measured domains. Within each complexity level, selective emphasis changes how the same measures are displayed: favorable policy titles are added, and good-direction numerical consequences are made visually prominent. The full text is reported in Appendix C.

Table 6: Study 4 municipal measures, selective emphasis, and domain consequences

Policy	Shown in	Neutral headline	Selective-emphasis headline	Emphasized consequences	Employment	Discretionary allocation	Service capacity
1	Low and high complexity	Regulation 1	Faster Project Delivery.	+300 jobs and +3,000 residents served.	+300	+200 million TL	+3,000
2	High complexity only	Regulation 2	Stronger Oversight.	+200 jobs and -250 million TL discretionary allocation.	+200	-250 million TL	0
3	Low and high complexity	Regulation 3	Downtown Revitalization Fund.	—	-700	+350 million TL	-6,000
4	High complexity only	Regulation 4	New Transparency Portal.	—	-200	+250 million TL	0
Low-complexity total (Measures 1 and 3)					-400	+550 million TL	-3,000
High-complexity total (Measures 1–4)					-400	+550 million TL	-3,000

Note: The table summarizes the municipal package shown in Study 4. Low-complexity respondents saw Measures 1 and 3; high-complexity respondents saw all four measures. Within each complexity level, neutral and selective-emphasis arms imply the same domain consequences. Selective emphasis changes the display of the same measures: favorable titles are added, and good-direction numerical consequences are bolded. Good-direction consequences are positive employment changes, reductions in discretionary allocation, and positive service-capacity changes. Positive discretionary allocation means more spending under municipal discretion.

The main outcomes are harmful-package acceptance and consequence recognition. Harmful-package acceptance is coded as one when respondents accept the municipal reform package, which is strictly worse than the status-quo in all three domains. Consequence recognition is coded as one when respondents identify that at least two relevant domains worsen under the package: employment, discretionary allocation, or service capacity. Recognition matters because the municipal task can fail in more than one way. Respondents may fail to integrate the domain-level consequences correctly, or they may recognize the costs but still accept the package when judging it as a whole.⁶ I also examine perceived difficulty, confidence, time,

⁶This measure is deliberately more demanding than the Study 3 recognition item. Study 3 asks respondents to recall which candidate proposed a specific oversight-weakening provision. Study 4 asks respondents to integrate consequences across several policy domains and identify whether multiple domains worsen. Consequence recognition in Study 4 therefore captures multidomain integration rather than recall of a single provision.

and clicks as secondary evidence about how respondents experienced the task.

I estimate linear probability models with indicators for high complexity, selective emphasis, and their interaction, using the low-complexity neutral condition as the baseline. The models adjust for module order, candidate-module order, survey date, and the preregistered risk-aversion measure. Because each respondent contributes one Study 4 reform-package decision, I use heteroskedasticity-robust standard errors and report effects in percentage points. Consequence recognition is a binary outcome and is analyzed with the same treatment specification.

8.2 Results

In the low-complexity neutral baseline, 59.1 percent of respondents accept the municipal package (Figure 5). Relative to that baseline, the high-complexity selective-emphasis condition increases harmful-package acceptance by 8.1 percentage points, with a 95 percent confidence interval from 1.8 to 14.3 points. The joint condition produces the largest shift in acceptance: respondents become more willing to accept the package when they see the longer version with favorable policy headlines. I therefore interpret the result as evidence that the combined presentation changes package acceptance, while treating the interaction term itself as secondary. The complexity-by-selective-emphasis interaction is -6.8 percentage points (95% CI: -15.6, 2.0), providing no evidence that the two presentational features amplify one another in this task.

Consequence recognition moves in the predicted direction, but the evidence is weaker than for package acceptance (Figure 6). In the low-complexity neutral baseline, 16.3 percent of respondents correctly identify that at least two relevant domains worsen. That low baseline reflects the difficulty of the recognition task and leaves limited room for further decline. Relative to that baseline, the high-complexity selective-emphasis condition lowers consequence recognition by 4.1 percentage points, with a 95 percent confidence interval from -8.8 to 0.5 points. The estimate is consistent with weaker recognition under the joint presentation, but it is also compatible with a small decline or no change.

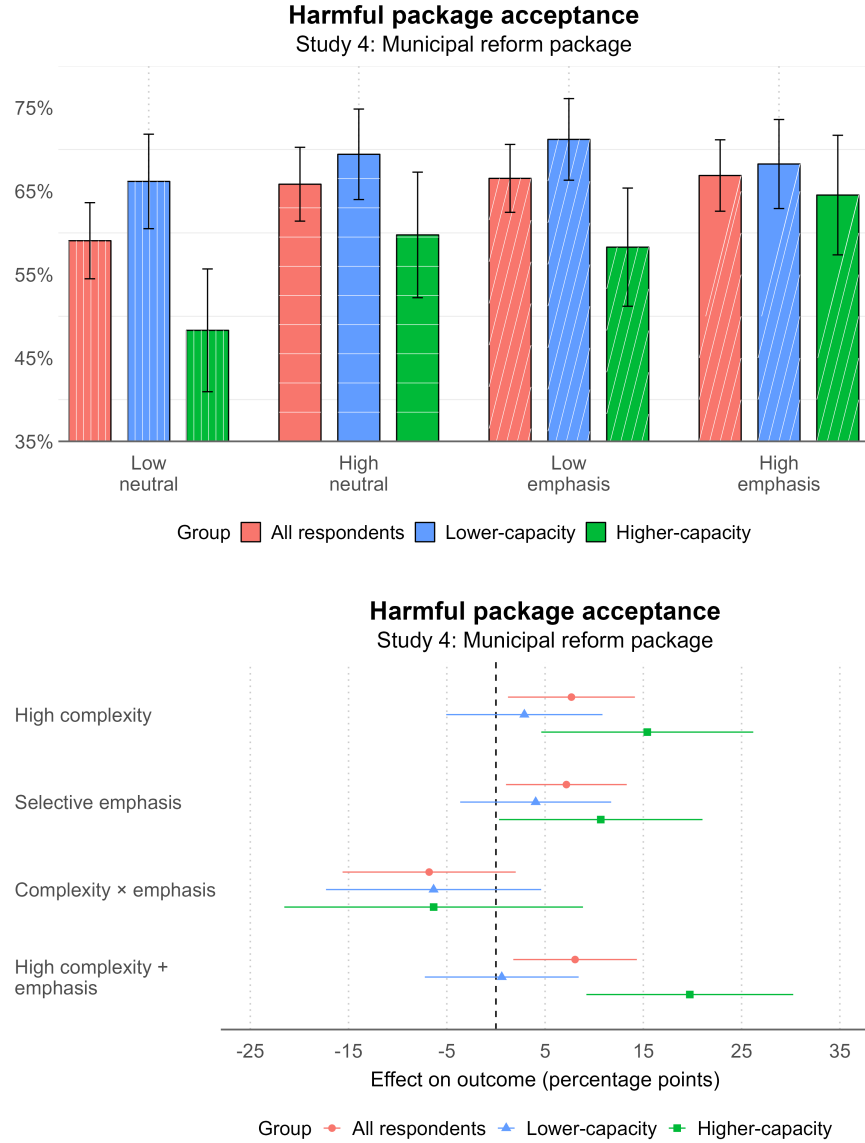


Figure 5: Study 4: Harmful-package acceptance.

Note: The analytic sample contains 1,871 respondents. The top panel reports treatment-arm means for harmful-package acceptance, coded as one when the respondent accepts the municipal reform package. The bottom panel reports linear probability model estimates relative to the low-complexity neutral baseline. The specification includes indicators for high complexity, selective emphasis, and their interaction, and adjusts for module order, candidate-module order, survey date, and the preregistered risk-aversion measure. Standard errors are heteroskedasticity-robust. Baseline municipal package-acceptance rates in the low-complexity, neutral-emphasis condition are 59.1 percent among all respondents, 48.3 percent among higher-capacity respondents, and 66.2 percent among lower-capacity respondents.

The subgroup estimates suggest that acceptance and recognition move on different margins. Acceptance rises most among higher-capacity respondents (Figure 5). For this group,

the joint condition increases package acceptance by 19.7 percentage points (95% CI: 9.2, 30.3). Among lower-capacity respondents, the corresponding estimate is 0.6 percentage points (95% CI: -7.2, 8.4). This difference should be read alongside the baseline rates. Lower-capacity respondents begin from a higher baseline acceptance rate of 66.2 percent, while higher-capacity respondents begin from a lower baseline of 48.3 percent. The combined presentation therefore has more room to increase acceptance among higher-capacity respondents, which may partly account for the larger estimated effect in that subgroup. The baseline difference, however, does not by itself explain the subgroup contrast.

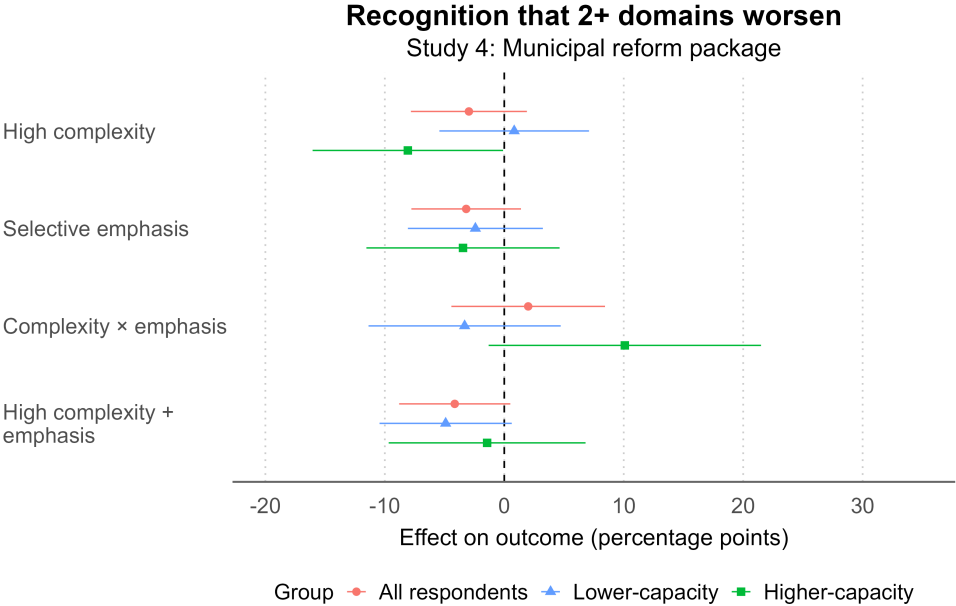


Figure 6: Study 4: Consequence recognition. The figure reports treatment effects on the indicator for correctly identifying that at least two relevant package domains worsen, relative to the low-complexity, neutral-emphasis baseline.

Note: The analytic sample contains 1,871 respondents. Consequence recognition is coded as one when the respondent correctly identifies that at least two relevant domains worsen under the municipal package. The figure reports linear probability model estimates relative to the low-complexity neutral baseline. The specification includes indicators for high complexity, selective emphasis, and their interaction, and adjusts for module order, candidate-module order, survey date, and the preregistered risk-aversion measure. Standard errors are heteroskedasticity-robust. Baseline consequence-recognition rates in the low-complexity, neutral-emphasis condition are 16.3 percent among all respondents, 19.7 percent among higher-capacity respondents, and 14.1 percent among lower-capacity respondents.

Recognition shows a different pattern. Higher-capacity respondents correctly identify

the multidomain consequences of the package only slightly more often than lower-capacity respondents in the baseline condition, 19.7 percent versus 14.1 percent (Figure 6). Among higher-capacity respondents, the joint condition changes consequence recognition by -1.4 percentage points (95% CI: -9.7, 6.8). Among lower-capacity respondents, the corresponding estimate is -4.9 percentage points (95% CI: -10.4, 0.6). Acceptance therefore rises most among higher-capacity respondents, who also begin from a lower baseline acceptance rate, while recognition declines more visibly among lower-capacity respondents.

This subgroup pattern fits the structure of the municipal task. The recognition item requires respondents to combine positive and negative consequences across employment, discretionary allocation, and service capacity. In that setting, presentation can matter in more than one way. Some respondents may fail to identify that multiple domains worsen. Others may recognize some costs but still accept the package because the favorable labels make the package more attractive overall. Study 4 cannot separate those paths, but it shows why acceptance and recognition need not move identically as the task becomes more demanding.

The process outcomes help rule out a pure non-engagement interpretation. Clicks increase under more complex presentation, indicating that respondents interact more with the municipal package rather than simply passing over it. Perceived difficulty moves in the same general direction, with the clearest increase in the high-complexity selective-emphasis condition. These measures do not show that respondents understood the package better; they show that higher acceptance is not simply a result of lower engagement. Respondents engage with a demanding reform package, but that engagement does not prevent higher acceptance or weaker consequence recognition.

Study 4 therefore extends the argument to a harder political setting. The longer package with favorable policy headlines increases harmful-package acceptance in the full sample. Consequence recognition moves downward, though the evidence is weaker. The subgroup results suggest that difficult multidomain packages can affect evaluation through more than one path: weakening recognition for some respondents, while changing how recognized or partially recognized costs enter final judgment for others.

9 Discussion

Support for reforms that weaken democratic oversight is often interpreted as evidence that voters knowingly accept democratic costs in exchange for partisan, policy, or material benefits. The findings in this paper suggest that this inference may sometimes be premature. When an oversight-weakening provision is embedded in a broader package and attractive components receive greater emphasis, citizens may support the proposal without fully understanding its democratic consequences. What appears to be a willingness to trade democracy for other benefits may therefore also reflect how the reform was assembled and presented.

When governments combine several provisions in an omnibus bill or reform package, citizens must assess those provisions together. This judgment becomes more demanding as the package grows longer, spans more policy domains, or relies on technical language whose institutional implications are not immediately clear. Selective emphasis by government officials, campaigns, and government-aligned media can further shape which parts of the package citizens notice and use. Promised benefits such as efficiency, security, or service delivery may become easier to identify than changes to oversight or executive discretion, leaving the democratic downside less clear at the moment of judgment.

The two abstract choice studies isolate this evaluation problem from partisan attachment, policy agreement, and other political motivations. Study 1 shows that bundled presentation increases suboptimal choice by 17.0 percentage points even when respondents choose between options with objectively ranked payoffs. Study 2 then examines whether selective emphasis becomes more consequential as the bundle becomes harder to evaluate. The pattern is concentrated among lower-capacity respondents: relative to the shorter, neutrally presented bundle, the longer bundle with attractive changes emphasized increases acceptance of the harmful option by 11.6 percentage points. Together, these findings show that presentation can impair evaluation even when the consequences are fixed and no partisan or ideological trade-off is present.

The political studies move the same argument into more realistic settings. In Study 3, relative to the low-complexity, neutral-presentation condition, the high-complexity, selective-emphasis condition increases support for the candidate associated with an oversight-weakening commitment by 17.0 percentage points, while recognition of that commitment declines by 7.9 percentage points. Study 4 extends the comparison to a municipal reform package. There, the larger package with favorable headlines increases acceptance by 8.1 percentage points, while consequence recognition moves downward by an estimated 4.1 points. Together, the political studies show that broader and more favorable presentation can increase support for oversight-weakening proposals. The decline in recognition is clearest in Study 3 and more

modest in Study 4.

The design addresses the trade-off interpretation directly, although the studies do so in different ways. In Studies 1 and 2, the harmful option is strictly worse in payoff terms, and in Study 4 the municipal package is worse than the status quo across all three measured domains. Increased acceptance in Studies 1 and 2 therefore cannot be explained by a payoff trade-off built into the alternatives. In Study 4, it cannot be explained by a compensating gain in any of the three experimentally measured domains. Study 3 allows more room for a perceived trade-off because the candidate who weakens committee oversight is also associated with attractive claims about speed and administrative flexibility. Because recognition is measured after respondents make their choice, the political studies cannot determine whether presentation increased support by making the harmful implication harder to identify, by reducing the weight respondents placed on a recognized cost, or by some combination of the two.

The subgroup patterns are substantively stable across the alternative inclusion rules. Most importantly, the Study 2 complexity-by-selective-emphasis pattern among lower-capacity respondents changes little as the sample is broadened. The interaction remains positive for both suboptimal choice and harmful-bundle acceptance under every inclusion rule, and the joint high-complexity and selective-emphasis contrast on harmful-bundle acceptance is similarly stable. Study 1 continues to show bundling effects in both capacity groups, and Study 3 candidate-support effects remain positive in both groups. In Study 4, the acceptance effect remains concentrated among higher-capacity respondents, while the lower-capacity estimate remains near zero. The Study 4 consequence-recognition estimates remain negative but imprecisely estimated and become smaller under broader inclusion rules.

The argument applies most clearly when provisions that weaken oversight are embedded in broader legal or policy packages, must be identified and evaluated alongside several other components, and have not already been made unmistakable through public debate. It should matter less when the harmful provision stands alone, the proposal openly weakens oversight, or opposition parties, independent media, courts, experts, and civil society have clearly explained how the package expands executive discretion or weakens institutional constraints.

Türkiye provides a demanding case because packaged reform is familiar rather than novel: citizens have repeatedly encountered institutional change through *torba yasalar*, constitutional packages, emergency decree-laws, and executive-centered administrative reforms. Finding presentation effects in this setting makes unfamiliarity with reform packages a less plausible explanation. Media capture makes the problem especially consequential by repeatedly emphasizing promised benefits while making criticism harder to encounter, although the experiments do not establish that voters misunderstood any particular Turkish reform.

The same scope conditions can arise beyond Türkiye. Hungary’s post-2010 constitutional transformation, India’s 2020 farm-law episode, and Israel’s 2023 judicial overhaul each placed major institutional or policy changes within multi-component packages (Bánkuti, Halmai and Scheppele, 2012; Behl, 2022; Navot, 2023). India and Israel also show how sustained opposition and mass mobilization can make the disputed costs central to public debate. The evaluation problem studied here should become less consequential once the contents and likely effects of a package are widely understood.

The same evaluation problem can extend beyond democratic backsliding. Governments often place costly or controversial provisions alongside popular benefits in omnibus legislation, ballot measures, campaign platforms, and administrative reforms. In these settings, the relevant harm may concern public spending, privacy, environmental risk, or access to services rather than democratic oversight. The argument is not that policy packages are inherently misleading. It is that when citizens must infer a proposal’s overall consequences from several components, selective emphasis in media coverage can make favorable consequences easier to identify than the costs embedded elsewhere in the package.

The findings also have implications for how researchers study democratic accountability. Experiments that state a democratic violation explicitly are well suited to testing whether citizens will tolerate a cost they clearly recognize. Many real reforms, however, do not present the democratic downside in that form. When institutional changes are embedded in longer platforms or multi-component packages, support alone cannot reveal whether citizens knowingly accepted the cost or failed to identify it clearly. Future research should therefore examine what citizens understood the proposal to contain alongside whether they supported it.

The distinction between conscious trade-offs and incomplete understanding also changes what opposition and civil society need to do. When voters recognize the institutional cost and accept it, opponents must explain why the promised benefits do not justify weaker oversight. When that cost remains unclear, they must first separate the package’s provisions and explain how those provisions alter oversight and executive discretion.

References

- Akser, Murat and Banu Baybars. 2023. “Repressed Media and Illiberal Politics in Turkey: The Persistence of Fear.” *Southeast European and Black Sea Studies* 23(1):159–177.
- Bánkuti, Miklós, Gábor Halmai and Kim Lane Scheppele. 2012. “Hungary’s Illiberal Turn: Disabling the Constitution.” *Journal of Democracy* 23(3):138–146.
- Behl, Natasha. 2022. “India’s Farmers’ Protest: An Inclusive Vision of Indian Democracy.” *American Political Science Review* 116(3):1141–1146.
- Bordalo, Pedro, Nicola Gennaioli and Andrei Shleifer. 2012. “Salience Theory of Choice Under Risk.” *The Quarterly Journal of Economics* 127(3):1243–1285.
- Bordalo, Pedro, Nicola Gennaioli and Andrei Shleifer. 2013. “Salience and Consumer Choice.” *Journal of Political Economy* 121(5):803–843.
- Bordalo, Pedro, Nicola Gennaioli and Andrei Shleifer. 2022. “Salience.” *Annual Review of Economics* 14:521–544.
- Braley, Alia, Gabriel S Lenz, Dhaval Adjodah, Hossein Rahnama and Alex Pentland. 2023. “Why voters who value democracy participate in democratic backsliding.” *Nature human behaviour* 7(8):1282–1293.
- Çarkoğlu, Ali and Kerem Yıldırım. 2024. “Media effects in a polarized political system: The case of Turkey.” *Political Behavior* 46(2):1189–1211.
- Chong, Dennis and James N. Druckman. 2007. “Framing Theory.” *Annual Review of Political Science* 10:103–126.
- Coşkun, Gülizar Betül. 2020. “Media Capture Strategies in New Authoritarian States: The Case of Turkey.” *Publizistik* 65:637–654.
- Druckman, James N. 2001. “The Implications of Framing Effects for Citizen Competence.” *Political Behavior* 23(3):225–256.
- Druckman, James N. 2024. “How to study democratic backsliding.” *Political Psychology* 45:3–42.
- Erem, Onur. 2017. “Son 2 yılda torba yasalardaki rekor artışın arkasında ne var?” BBC Türkçe.

- Esen, Berk. 2021. Competitive authoritarianism in Turkey under the AKP rule. In *The Routledge Handbook on Contemporary Turkey*. Routledge pp. 153–167.
- Esen, Berk, Sebnem Gumuscu and Hakan Yavuzylmaz. 2023. “Competitive yet unfair: May 2023 elections and authoritarian resilience in Turkey.” *South European Society and Politics* 28(3):359–387.
- Gabaix, Xavier. 2014. “A sparsity-based model of bounded rationality.” *The Quarterly Journal of Economics* 129(4):1661–1710.
- Gabaix, Xavier. 2019. Behavioral inattention. In *Handbook of behavioral economics: Applications and foundations 1*. Vol. 2 Elsevier pp. 261–343.
- Gabaix, Xavier and Thomas Graeber. 2024. The Complexity of Economic Decisions. NBER Working Paper 33109 National Bureau of Economic Research.
- Graham, Matthew H and Milan W Svobik. 2020. “Democracy in America? Partisanship, polarization, and the robustness of support for democracy in the United States.” *American Political Science Review* 114(2):392–409.
- Gratton, Gabriele and Barton E Lee. 2024. “Liberty, security, and accountability: The rise and fall of illiberal democracies.” *Review of Economic Studies* 91(1):340–371.
- Grillo, Edoardo and Carlo Prato. 2023. “Reference points and democratic backsliding.” *American Journal of Political Science* 67(1):71–88.
- Grillo, Edoardo, Zhaotian Luo, Monika Nalepa and Carlo Prato. 2024. “Theories of Democratic Backsliding.” *Annual Review of Political Science* 27:381–400.
- Grossman, Guy, Dorothy Kronick, Matthew Levendusky and Marc Meredith. 2022. “The Majoritarian Threat to Liberal Democracy.” *Journal of Experimental Political Science* 9(1):36–45.
- Haspolat, Evren. 2016. “Yasamadan yürütmeye iktidar devri olarak torba yasa (ma).” *Alternatif Politika* 8(2):404–449.
- Hollyer, James R, Marko Klasnja and Rocío Titiunik. 2023. “Charismatic leaders and democratic backsliding.” *Work. Pap., Univ. Minn./Georgetown Univ./Princeton Univ.* <https://titiunik.github.io/papers/KlasnjaHollyerTitiunik2023-wp.pdf>.
- Kadioğlu, Ayşe. 2021. “Autocratic legalism in new Turkey.” *Social Research: An International Quarterly* 88(2):445–471.

- Kahneman, Daniel and Amos Tversky. 1984. "Choices, values, and frames." *American psychologist* 39(4):341.
- Kendall, Chad and Ryan Oprea. 2024. "On the Complexity of Forming Mental Models." *Quantitative Economics* 15(1):175–211.
- Krishnarajan, Suthan. 2023. "Rationalizing democracy: the perceptual bias and (un) democratic behavior." *American Political Science Review* 117(2):474–496.
- Krutz, Glen S. 2001. "Tactical maneuvering on omnibus bills in Congress." *American Journal of Political Science* pp. 210–223.
- Lee, Barton E. 2022. "Gridlock, Leverage, and Policy Bundling." *Journal of Public Economics* 212:104687.
- Luo, Zhaotian and Adam Przeworski. 2019. "Democracy and its vulnerabilities: Dynamics of democratic backsliding." *Available at SSRN 3469373* .
- Lupia, Arthur. 1994. "Shortcuts Versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections." *American Political Science Review* 88(1):63–76.
- Maćkowiak, Bartosz, Filip Matějka and Mirko Wiederholt. 2023. "Rational inattention: A review." *Journal of Economic Literature* 61(1):226–273.
- Matějka, Filip and Guido Tabellini. 2021. "Electoral Competition with Rationally Inattentive Voters." *Journal of the European Economic Association* 19(3):1899–1935.
- Navot, Suzie. 2023. "An Overview of Israel's "Judicial Overhaul": Small Parts of a Big Populist Picture." *Israel Law Review* 56(3):482–501.
- Oprea, Ryan. 2020. "What Makes a Rule Complex?" *American Economic Review* 110(12):3913–3951.
- Özcan, Abdullah Sait and Fırat Kimya. 2024. "Executive decrees, omnibus bills, and the politics of abusive constitutionalism in Turkey." *Southeast European and Black Sea Studies* 24(4):881–903.
- Reporters Without Borders. 2018. "Doğan Media Group Sale Completes Government Control of Turkish Media." <https://rsf.org/en/do%C4%9Fan-media-group-sale-completes-government-control-turkish-media>. Accessed 17 July 2026.

- Sav, Atilâ. 2014. “Yasa yapma sanatının ölümü (torba yasalar).” *Ankara Barosu Dergisi* (4):595–606.
- Sebők, Miklós, Rebeka Kiss and Ádám Kovács. 2023. “The Concept and Measurement of Legislative Backsliding.” *Parliamentary Affairs* 76(4):741–772.
- Shirikov, Anton. 2024. “Rethinking Propaganda: How State Media Build Trust Through Belief Affirmation.”
- Sims, Christopher A. 2003. “Implications of rational inattention.” *Journal of monetary Economics* 50(3):665–690.
- Sinclair, Barbara. 2012. *Unorthodox Lawmaking: New Legislative Processes in the U.S. Congress*. 4 ed. CQ Press.
- Somer, Murat. 2019. “Turkey: The slippery slope from reformist to revolutionary polarization and democratic breakdown.” *The ANNALS of the American Academy of Political and Social Science* 681(1):42–61.
- Svolik, Milan. 2018. “When polarization trumps civic virtue: Partisan conflict and the subversion of democracy by incumbents.” *Available at SSRN 3243470* .
- Tversky, Amos and Daniel Kahneman. 1981. “The Framing of Decisions and the Psychology of Choice.” *Science* 211(4481):453–458.
- Yanatma, Servet. 2021. “Advertising and media capture in Turkey: how does the state emerge as the largest advertiser with the rise of competitive authoritarianism?” *The International Journal of Press/Politics* 26(4):797–821.
- Yeşil, Bilge. 2016. *Media in New Turkey: The Origins of an Authoritarian Neoliberal State*. Champaign, IL: University of Illinois Press.
- Zaller, John. 1992. *The nature and origins of mass opinion*. Cambridge university press.