Who Overreports?
Regime Support and Preference Falsification in Iran

Daniel L. Tavana[*] Penn State
Kevan Harris[†] UCLA
Gary Fong[‡] Penn State
Amir Farmanesh[§] IranPoll

June 15, 2026

Abstract

In authoritarian contexts, individuals often express greater support for the regime in public than they hold in private. This gap—overreporting—is typically treated as a measurement problem rather than a political behavior to be explained. We argue that overreporting deserves greater attention in survey-based studies of regime support and is itself shaped by which social referents are salient to respondents during survey interviews. We develop and test an original theoretical framework that identifies three sources of overreporting: expressive identification, behavioral conformity, and civic resources. Drawing on a pre-registered, nationally representative telephone survey of residents of Iran (N=1,999) with an embedded double list experiment, we estimate regime support from the list experiment at 47 percent—roughly ten percentage points below the 57 percent recorded by direct questioning. Our findings indicate that expressive identification matters most: national pride and religiosity both predict substantial overreporting. Preference falsification in Iran is shaped less by the coercion citizens face from the state than by the identity-based attachments the regime encourages—findings with implications for how scholars measure and interpret regime support in closed societies. [*181 words*]

Keywords: authoritarian regimes; preference falsification; regime support; list experiments; Iran

Word count: 9,182 words (excluding headings, figure and table text, and references)

[*]Assistant Professor, Department of Political Science, Pennsylvania State University. Email: tavana@psu.edu.
[†]Associate Professor, Department of Sociology, University of California, Los Angeles. Email: kevanharris@soc.ucla.edu.
[‡]Ph.D. Candidate, Department of Political Science, Pennsylvania State University. Email: cjf5914@psu.edu.
[§]CEO, IranPoll. Email: farmanesh@iranpoll.com.

Introduction

Our knowledge of public opinion in authoritarian regimes relies on surveys, and surveys rely on the assumption that respondents tell the truth. Where dissent is costly, that assumption may be fragile. Citizens may express public support for regimes they privately reject, and the answers researchers recover from direct survey questions may bear little resemblance to the beliefs they aim to measure (Kuran 1991, 1995). This matters for a variety of empirical questions scholars have asked about regime support—from the role of economic performance and patronage (Gehlbach & Keefer 2012; Huang & Zuo 2023; Rommel 2024), to identity-based attachments to political elites (Collins & Owen 2012; Hoffman 2020), to the constraining effects of state repression (Davenport 2007; Bhasin & Gandhi 2013). If our estimates of regime support are systematically inflated, so too are the inferences we draw from them. Researchers have responded with question techniques designed to measure sensitive attitudes under conditions of anonymity, of which the list experiment is by far the most widely used (Imai 2011; Blair & Imai 2012; Glynn 2013).

A growing body of work has used these techniques to document substantial gaps between direct and indirect estimates of regime support: a phenomenon we refer to as overreporting (Nicholson & Huang 2022). But the gap itself—the difference between what respondents say in public and what they reveal in private—is generally treated as a measurement problem to be diagnosed and corrected, not as behavior to be explained. The literature has not produced a systematic individual-level account of overreporting that tests competing mechanisms against each other or asks why particular citizens, rather than others, publicly affirm support they do not privately hold.

We argue that overreporting is itself a substantive form of political behavior worth greater attention. Following Blair *et al.* (2020), we treat overreporting as evidence of sensitivity bias: the gap between what respondents endorse when their answers are individually attributable, as in direct questions, and what they endorse under the anonymity of an aggregate response format. Sensitivity bias arises when

respondents have a social referent in mind whose perceived preferences and capacity to impose costs shape their answers in a survey. As a result, individual variation in overreporting reflects variation in which referents are most salient. We organize the predictors of overreporting into three theoretical sources (with social referents in parentheses): *expressive identification* (the religious community and the nation), *behavioral conformity* (the state and the enumerator), and *civic resources* (the cognitive and material capacity to form independent judgments). The first two push public expression away from private belief; the third brings the two closer together.

We test this framework using a pre-registered, nationally representative telephone survey of adults in Iran (N=1,999), conducted from February to March 2025 in collaboration with IranPoll. The survey embeds a double list experiment in which the sensitive item asks respondents how much they trust the authorities of the Islamic Republic. Each respondent receives one treatment list and one control list, yielding two observations per respondent and substantial efficiency gains over single list designs. Iran is a productive case for studying overreporting. It is a hybrid regime that combines elements of theocratic and democratic governance: elected representatives share power, often unevenly, with an unelected Supreme Leader and the clerical elite that comprise his ruling coalition. A number of surveys have estimated support at over 50 percent in national polls even as the country has experienced rising mass protest, declining electoral participation, and the recurring repression of dissent. We estimate aggregate regime support at approximately 47 percent in the double list experiment, against 57 percent in the direct question—a gap of roughly ten percentage points that accounts for nearly one in five respondents who publicly affirm support they do not privately hold. Subgroup analyses indicate expressive identification as the clearest source of overreporting, while findings related to variables we group under behavioral conformity and civic resources are more mixed.

The paper makes three contributions. First, it reframes overreporting from a measurement artifact to a substantive political act that can be theorized and tested. Existing research documents that public expression in authoritarian contexts diverges from private belief; we develop a theoretical framework

that explains who falsifies and why, putting the predictors of overreporting on equal theoretical footing with the predictors of regime support itself. Second, it synthesizes a broad and disconnected literature on the mechanisms behind preference falsification—fear, identity, social desirability, civic capacity—into a unified framework anchored to a common referent-based logic. The framework is consistent with Blair *et al.*'s (2020) social reference theory of sensitivity bias but extends it to the individual-level prediction of who has which referent in mind. In doing so, we test several competing mechanisms rather than considering them in isolation.

Third, the empirical findings invert the literature on regime support in authoritarian contexts. Predominant explanations suggest that coercion—fear of the security apparatus and exposure to repression—explains preference falsification (Kuran 1991, 1995; Wintrobe 1998; Davenport 2007). We find, in Iran, that the most internalized referents—identity-based attachments to the nation and the religious community—are the strongest predictors of overreporting, while variables capturing externally imposed pressure or independence from regime narratives are less important. Preference falsification in our data is shaped less by the coercion citizens face from the state than by the identities the regime encourages citizens to adopt (Guriev & Treisman 2019, 2020).

Theoretical framework

Understanding who supports non-democratic regimes and why is a central question in the comparative study of authoritarian politics. A common premise runs through this literature: when we ask citizens in closed societies how much they support or trust their leaders, they answer. And their answers inform the private beliefs scholars aim to explain. A parallel literature on preference falsification has called this premise into question. Where dissent is costly, public expressions of regime support may not reflect private belief (Kuran 1991), and the evidence scholars recover from direct questions may reflect systematic gaps between what citizens say and what they actually believe.

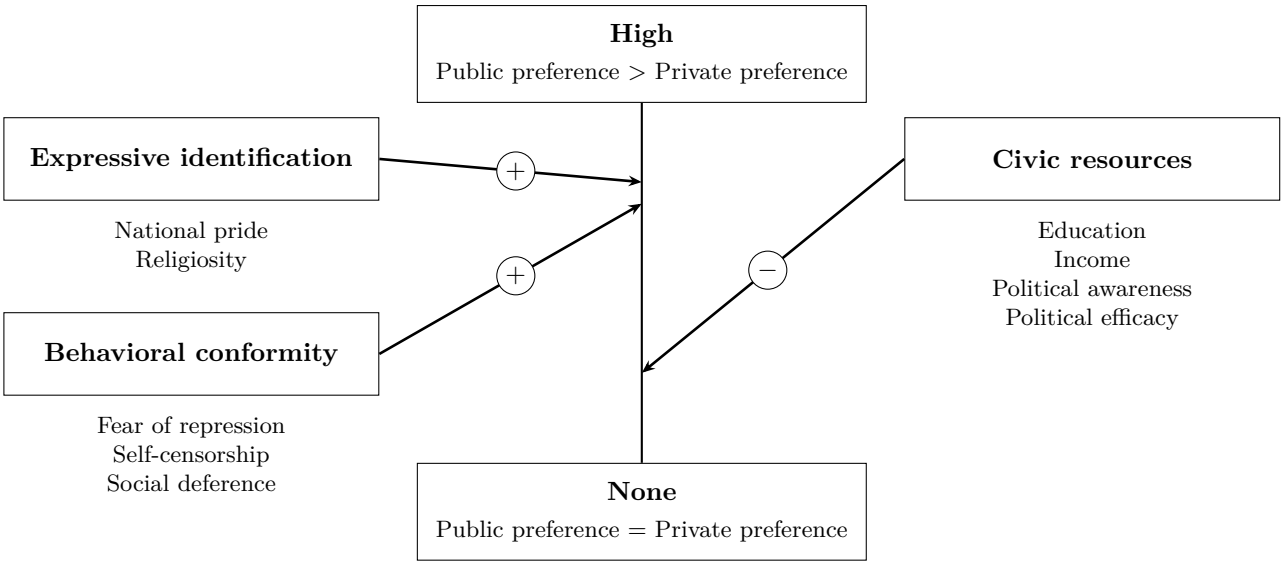
We treat overreporting as evidence of sensitivity bias: the difference between what respondents endorse when their answers are individually attributable, as in direct questions, and what they endorse when protected by the anonymity of an aggregate response format, as in list experiments (Blair *et al.* 2020). The substantive interpretation of this understanding of regime support is that citizens who overreport are not responding at random. They are managing their answers for reasons that may include fear of sanction, normative obligation, social deference, or rational calculation about the costs of candor (Corstange 2022; Hale 2022; Kasuya & Miwa 2023). Citizens do not falsify randomly (Jiang & Yang 2016; Robinson & Tannenberg 2019), and falsification can be critical to regime survival. When aggregated across a population, public affirmations of regime support produce the appearance of a consensus that authoritarian regimes rely on (Kuran 1991; Hale 2022; Buckley *et al.* 2024). Theorizing the predictors of overreporting is no less important than theorizing the predictors of regime support.

Scholars have responded to the pervasiveness of preference falsification with methodological innovations designed to measure sensitive attitudes under conditions of anonymity. Sensitive question techniques have been deployed across authoritarian and backsliding contexts to recover regime support estimates that direct questions cannot capture (Imai 2011; Blair & Imai 2012; Glynn 2013; Gingerich *et al.* 2016). Evidence documents substantial gaps between direct and indirect estimates in China (Jiang & Yang 2016; Robinson & Tannenberg 2019; Carter *et al.* 2024), Russia (Kalinin 2016; Frye *et al.* 2017; Hale 2022; Buckley *et al.* 2024), Syria (Corstange 2022), the Philippines (Kasuya & Miwa 2023; Iglesias 2025), Hong Kong (Kobayashi & Chan 2022; Kobayashi *et al.* 2025), Taiwan (Wu & Lin 2024), and across the broader authoritarian landscape (Tannenberg 2017; Carlson 2018; Tannenberg 2022; Shen & Truex 2021). Several studies move beyond aggregate estimates to document demographic correlates and offer mechanism-specific explanations (Jiang & Yang 2016; Kalinin 2016; Nicholson & Huang 2022; Kasuya & Miwa 2023). But this literature has not produced a systematic individual-level account of overreporting that places competing mechanisms on a common theoretical footing. Existing studies consider one or two pathways in isolation—fear of repression, social desirability, or economic

vulnerability—and demographic correlates are typically interpreted after the fact rather than derived from a theory that explains why particular individuals should be more or less likely to publicly affirm support they do not privately hold.

We organize the predictors of overreporting into three sources that capture different mechanisms through which overreporting may occur. We draw on Blair *et al.*'s (2020) social reference theory of sensitivity bias, which holds that survey responses diverge from privately held beliefs when respondents have a *social referent* in mind—a person or group whose perceived preferences and capacity to impose costs shape what the respondent is willing to publicly affirm. Individual variation in overreporting reflects variation in which referents are most salient and consequential. The first source, *expressive identification*, captures the most internalized referents: the self, the religious community, or the nation. Variables here push public expression away from private belief through identity-based obligations. The second, *behavioral conformity*, captures externally imposed referents—the state or the enumerator—that lead citizens to report support as a matter of survival, habit, or social deference. These variables also push public expression away from private belief, but through external pressure rather than internalized obligation. The third, *civic resources*, captures factors that reduce a respondent's dependence on a salient referent, equipping citizens to form and voice independent judgments. These invert the logic of the first two, predicting less overreporting. Although all three involve cognition, they differ in the type of pressure on public expression: identity-based attachments push expression away from private belief through felt obligation, externally imposed referents push it away through anticipated cost, and civic resources reduce dependence on salient referents. A respondent can be high on expressive identification and high on civic resources simultaneously: the two are not mirror images of the same cognitive process. Figure 1 presents the framework; we develop each source and its hypotheses below.

Figure 1: Sources of overreporting



Note: Predictors of overreporting are organized into three sources. Expressive identification and behavioral conformity are hypothesized to increase overreporting (+); civic resources are hypothesized to reduce it (−). The vertical spectrum represents the magnitude of overreporting, from high (public > private support) to none (public = private).

Expressive identification

The first source captures the most internalized social referents: identities through which citizens come to see regime support as an expression of who they are. These attachments take different forms in different contexts—religious communities, ethnic or sectarian groups, the nation, or a partisan in-group—but they share a common feature: identity-based obligations push public expression away from private belief. In hybrid regimes that rely on religious authority and nationalist narratives, for example, the religious community and the nation are the most consequential referents. Regime support is not just a response to coercion—but a felt obligation rooted in attachments to religious and national identities. Foundational work emphasizes that regime support rests on more than performance; a citizen's attachment to the political system can be a commitment formed through socialization, affective bonds, and shared narratives (Easton 1965). Where regimes successfully cultivate this support—drawing on religious tradition, charismatic authority, or national myths (Weber 1978)—citizens understand support not as coerced acquiescence but as a quasi-voluntary expression of

who they are (Levi 1997). Gerschewski (2013) synthesizes this as the "legitimation pillar" of authoritarian stability: a diffuse form of support that insulates regimes against performance shocks specific support alone cannot absorb. These identity-based attachments generate pressure to publicly affirm support even when private evaluation is ambivalent or negative. For respondents whose religious or national identity is intertwined with the regime's legitimating narratives, dissent feels like a betrayal of self, and direct survey questions activate identity-consistent affirmations. Variables in this source should therefore predict more overreporting.

The first variable we incorporate in our framework is national pride. The mechanism is the conflation of regime and country: for proud citizens, evaluating the regime becomes inseparable from evaluating the nation. Regimes cultivate this conflation through, for example, victimization narratives, performance-legitimacy claims, and controlled media narratives that bind national identity to incumbent leadership (Dimitrov 2009; Xu & Zhao 2023; Huang 2025); socialization links national identity to regime acceptance (Neundorf & Pop-Eleches 2020). The rally-effect literature documents these dynamics in detail. In Russia, high-intensity conflicts produce durable approval boosts (Bussmann & Iost 2024), patriotic priming raises presidential evaluations (Sirotkina & Zavadskaya 2020), and rallying extends to third-party publics whose leaders align with conflict narratives (Fukumoto & Tabuchi 2025). These effects are neither universal nor automatic: cross-nationally, the average rally effect is negative (Seo & Horiuchi 2024), and direct exposure to conflict costs can produce blame attribution rather than rallying (Hintson & Vaishnav 2023). The relevant referents are the nation as imagined community and fellow nationals as audience. For citizens who privately distinguish between regime and nation, criticism of one feels like disloyalty to the other when expressed publicly. We therefore expect higher national pride to predict more overreporting.

Our second variable is religiosity. Whether Islamic religiosity is inherently incompatible with democracy is a long-standing debate; the empirical record is mixed, with general measures of belonging, commitment, and orthodoxy predicting democratic attitudes inconsistently (Ciftci 2010; Spierings

2014) and most Arab Barometer respondents endorsing democracy regardless of religiosity (Jamal & Tessler 2008). Where religion is politically salient—where state legitimacy rests on religious authority, or where religious practice activates identity-based interests—religiosity predicts regime support. Religiosity raises support for Islamic over secular governance (Collins & Owen 2012), and communal religious practice heightens sectarian identity salience in ways that push citizens toward regime preferences aligned with their group's political position (Hoffman 2020). The politically consequential dimension is not piety *per se*, but the specific content of religious-political values (Ciftci 2013). In Iran, Tezcür *et al.* (2012) show that religiosity reduces democratic support and raises regime satisfaction, operating independently of demographic correlates that work only indirectly through performance evaluations. Where religious and political authority are fused, the religious community is a salient referent, and expressing regime support cannot easily be separated from expressions of piety. We therefore expect religiosity to predict more overreporting.

Our hypotheses consistent with this first source are as follows:

> H_{1a}: *Respondents who self-report higher levels of national pride are more likely than respondents who self-report lower levels of national pride to overreport support for the regime.*
>
> H_{1b}: *Religious respondents are more likely than non-religious respondents to overreport support for the regime.*

Behavioral conformity

The second source captures externally imposed referents: the state and the enumerator. Citizens here do not express support for the regime because it feels "right," but because honesty can be costly. Preference falsification is a rational response to external costs—fear of sanction, social ostracism, material loss—with each citizen having a threshold at which the cost of concealment is outweighed by the cost of dissent (Kuran 1995). The same mechanism that produces social-desirability bias on sensitive items in democratic contexts (such as attitudes towards race, immigration, and sexuality) operates in more demanding form in authoritarian contexts, where the audience capable of imposing

costs includes the state itself. The broader implication is known as the "dictator's dilemma": because loyalty under coercion is indistinguishable from genuine support, autocrats cannot reliably assess their own support (Wintrobe 1998). Coercion shapes what citizens are willing to say in public (Davenport 2007): fear causes overreporting because admitting the truth is costly, self-censorship internalizes this calculus into a durable disposition, and cultural norms of deference contribute non-political scripts that produce agreeableness toward an interlocutor. We expect fear, self-censorship, and norms of deference to predict more overreporting.

The first variable is fear of repression. When citizens believe that expressing opposition—even in a survey—could attract the attention of the security apparatus, the rational response is to report loyalty rather than voice dissent (Kuran 1991). Empirical work documents significant gaps between expressed and private support. Panel evidence shows that ralliers to wartime incumbents systematically misreport their prior vote, with list experiments confirming the gap reflects dissembling rather than recall error (Hale 2022). Enumerator-coded wariness spikes before elections in illiberal regimes and concentrates among opposition voters (Carlson 2018), and word-of-mouth awareness that surveys probe sensitive topics produces a grapevine effect that shifts later answers toward incumbent preferences (McCauley *et al.* 2022). Repression also reshapes expression indirectly. Where state media is pervasive, repression generates regime support rather than backlash (Pop-Eleches & Way 2023). During protest waves, preventive tools such as permit denial cause bystanders to view protesters as unreasonable (Tertytchnaya 2023). Repression is an implicit cue prompting bystanders to adopt the regime's stance on the punished behavior (Zhu *et al.* 2025). We therefore expect respondents with higher fear of repression to overreport more.

The second variable is self-censorship. Unlike fear, which is a response to acute perceived threat, self-censorship is a dispositional orientation toward concealment that citizens develop under sustained political pressure (Gueorguiev *et al.* 2017). The argument is not that self-censorship causes overreporting but that the two behaviors share an underlying disposition: citizens who routinely manage

their public expression in daily life carry that practice into survey interviews (Berinsky 1999). Cross-national item-nonresponse data identify systematic variation in self-censorship across authoritarian regimes, with closed regimes showing the most concealment (Shen & Truex 2021). In China, citizens strategically suppress political expression at the intersection of politically focal times and places, using silence to signal distance from potential troublemakers (Chang & Manion 2021). Social influence spreads self-censorship through family and peer networks (Yuen & Lee 2025), and citizens self-censor strategically where online repression is severe (Ong 2021). The behavioral signature extends to surveys: perceived government sponsorship can inflate estimates of trust, consistent with a disposition that activates when the social referent can punish (Tannenberg 2017; Isani & Schlipphak 2023). We therefore expect respondents who self-censor to overreport more.

The third variable is social deference: norms of politeness, ritual modesty, and conflict avoidance that shape how citizens respond to authority in face-to-face interaction. Unlike fear and self-censorship, this is a cultural script rather than a political response: it calls for the agreeable, socially harmonious answer in any interpersonal interaction, including a survey interview. Political culture—values and norms acquired through early socialization—shapes how citizens respond to authority independent of rational calculation. Shi (2001) shows that in both authoritarian and democratic Chinese-speaking contexts, cultural orientations emphasizing hierarchy and conflict avoidance make citizens more likely to voice support for governing authorities. An affective analogue operates in Russia, where collective engagement in political events generates pride, hope, and trust that mediate the relationship between engagement and approval (Greene & Robertson 2022). Survey interviews are not politically neutral: the interaction can call for deference, and a critical answer can feel inappropriate. Where norms of deference are more deeply ingrained, the interviewer becomes a referent demanding an agreeable response, and respondents add a layer of compliance on top of whatever political motivations are already operating. We therefore expect higher levels of social deference to predict more overreporting.

Our hypotheses consistent with this second source are as follows:

H_{2a}: *Respondents who are more fearful of repression are more likely than respondents who are less fearful of repression to overreport support for the regime.*

H_{2b}: *Respondents who self-censor are more likely than respondents who do not self-censor to overreport support for the regime.*

H_{2c}: *Respondents who self-report higher levels of social deference are more likely than respondents who self-report lower levels of social deference to overreport support for the regime.*

Civic resources

The third source captures the informational and cognitive capacity through which citizens perceive, evaluate, and respond to political life. Unlike the first two sources, this one describes resources that reduce a respondent's dependence on whichever referent is salient, equipping citizens to form and voice independent judgments. Foundational accounts of citizenship and political support frame this in different ways. Easton describes this as the cognitive substrate of specific support, in which citizens evaluate authorities against performance criteria rather than accepting regime claims at face value (Easton 1965). Others situate the same capacity in a participatory civic orientation in which citizens see themselves as competent to assess and influence political outcomes (Almond & Verba 1963). Judgments of one's own capability mediate the relationship between knowledge and action, determining whether information translates into independent expression or is suppressed in favor of deference (Bandura 1982). The same resources that enable critical evaluation lower the costs of expressing it candidly. A complementary mechanism may also be at work: better-resourced citizens are typically more aware of where the regime draws its line of permissible expression and can voice critical views close to (but not past) that line. Less-resourced citizens, uncertain about the exact location of the line and rarely following politics closely enough to track its shifts, may default to a baseline of risk-averse compliance. Variables in this source should therefore predict less overreporting.

The first pair of variables is education and income. Both are often treated as interchangeable proxies for socioeconomic status, but in authoritarian contexts they operate through distinct mechanisms.

Education equips citizens with cognitive tools to evaluate regime claims against lived experience (Cantoni *et al.* 2017; Kang & Zhu 2021; Kao 2021), though it can also demobilize or direct citizens into regime-friendly forms of participation (Croke *et al.* 2016; Wang 2018; Zeira 2019). Both arguments assume that education generates critical capacity. We extend that capacity to the formation and honest expression of regime evaluations: educated respondents are less dependent on official narratives and the referents that endorse them. Income operates differently: material resources reduce the costs of private deliberation and of honest public expression by loosening the respondent's economic dependence on the regime as a referent. Citizens whose economic position depends on the regime—through public-sector employment or patronage—develop weaker grievances and express less dissent (Rosenfeld 2017), while those exposed to economic precarity develop more critical evaluations of regimes that fail to protect them (Huang & Zuo 2023; Rommel 2024; Toklo *et al.* 2025). We expect more educated and higher-income respondents to overreport less.

The third variable is political awareness: citizens' baseline knowledge of and engagement with political life. The claim proceeds in two steps. First, politically aware citizens are better positioned to question official narratives. Citizens with greater information access are more likely to reject regime exaggerations of democratic commitment (Yeung 2023). In China, exposure to unofficial media decreases political trust through diminished subjective well-being, while state-media exposure inflates support for repressive institutions by concealing their political functions (Hu & Pu 2023; Xu *et al.* 2022). Internet exposure cultivates commitments to democratic values that can erode regime support, though censorship can mute direct effects on leadership approval (Huhe *et al.* 2018; Tang & Huhe 2020). Second, politically aware citizens update sincerely when they learn regime-discrediting information: informing supporters about electoral fraud, for example, reduces support through moral disapproval, and politicized social networks raise awareness in ways that translate into opposition (Reuter & Szakonyi 2015, 2021). Information produces private judgments that diverge from regime-supplied frames and reduces the salience of regime referents in survey responses. We therefore expect higher political

awareness to predict less overreporting.

The fourth variable is political efficacy: citizens' sense that they can form and express political views that matter, comprising both internal efficacy (confidence in one's own political competence) and external efficacy (belief that the system responds to citizen input). Judgments of one's own capability mediate the relationship between knowledge and action: high efficacy paired with an unresponsive environment produce protest or attempts at social change, while low efficacy combined with limited perceived outcomes produces apathy and withdrawal (Bandura 1982). In hybrid regimes, only opposition supporters revise perceptions of electoral integrity downward during autocratization, suggesting that efficacy filters information through prior orientation (Windecker *et al.* 2025; Robertson 2017). Conversely, low efficacy produces passive disengagement: citizens who perceive the regime as invincible do not mobilize even when local information reveals electoral competitiveness (Peisakhin *et al.* 2020), and citizens who distrust political institutions across the board withdraw through learned helplessness (Wen *et al.* 2025). Efficacious respondents are less dependent on regime referents to validate their political judgments. We therefore expect higher political efficacy to predict less overreporting.

Our hypotheses consistent with this third source are as follows:

> H_{3a}: *Less educated respondents are more likely than more educated respondents to overreport support for the regime.*
>
> H_{3b}: *Low-income respondents are more likely than high-income respondents to overreport support for the regime.*
>
> H_{3c}: *Respondents who exhibit lower levels of political awareness are more likely than respondents who exhibit higher levels of political awareness to overreport support for the regime.*
>
> H_{3d}: *Less efficacious respondents are more likely than more efficacious respondents to overreport support for the regime.*

Regime support in Iran

The Islamic Republic of Iran is a competitive authoritarian system in which "competition is real but unfair" (Kamrava 2023, p. 13). The president and legislature are popularly elected from a vetted slate of candidates, factional rivalry can generate substantive policy change, and a press operates within shifting limits on political speech. These institutions function, however, under the Supreme Leader, whose constitutional office holds absolute authority (*velayat-e faqih-e motlaq*) and extends beyond representative bodies into the judiciary and intelligence services. Power is distributed asymmetrically, and the boundaries between institutions shift depending on the issue and the moment. The state simultaneously solicits popular involvement and subordinates it to clerical review and approval. Iran is therefore neither a closed setting in which public opinion is presumed inaccessible nor a liberalizing autocracy in which open political expression can be taken for granted. Public affirmations of support coexist with recurring protest waves, repression, and declining turnout in national elections. This combination makes Iran a well-suited case for studying overreporting on sensitive political attitudes. Expressive identification, behavioral conformity, and civic resources are all visible in the Iranian case: religious and national idioms of legitimation, coercive institutions capable of punishing dissent, and variation in education, income, and access to political information.

The fusion of religious and state authority in Iran holds that political legitimacy (*mashru'iyat*) is divine in origin but "practically irrelevant without acceptability [*maqbuliyat*], which makes the system functional when people participate in it" (Kamrava 2024, p. 238). Disobedience to the *velayat-e faqih* is, in the state's own jurisprudential terms, religiously forbidden. Religion is therefore not just a demographic attribute: it is constitutive of state identity and the public compliance it stipulates. Shi'ite commemorations structure the civic calendar, and public piety functions as a marker of political loyalty. At the same time, survey data from inside Iran indicate a gap between official demands and private practice. Comparing a pre-revolutionary 1974 national survey with the 2000 National Survey

of Values and Attitudes, individual daily prayer was virtually unchanged (83 to 82 percent) while collective participation declined sharply, with congregational prayer falling from 59 to 20 percent (Kazemipur 2022, p. 122-124). Public religious performance and private religious practice diverge in ways that motivate close attention to religiosity as a predictor of overreporting.

National pride in the Iranian context also requires some elaboration. National consciousness in Iran reflects a tension between pre-Islamic Persian heritage and the Islamic revolutionary identity the state promotes (Maghen 2023, p. 256). The Iran-Iraq War (1980-1988), four decades of international sanctions, and recurring confrontations with external powers have reinforced sentiments of defensive nationalism that cut across political orientation. Pride can align with support for the state, but it is also expressed by citizens who oppose the government while refusing to concede their country's standing to outsiders. Both registers of pride are plausibly tied to public expressions of regime support, even when private evaluation diverges.

Two further features of the Iranian context make behavioral conformity plausible. The first is *ta'arof*, the ritualized exchange of politeness "fundamental to Iranian social life," which combines norms of face-saving, mutual deference, and ritual modesty in interpersonal interaction (Moruzzi 2025, p. 132). Survey interviews are not exempt from these conventions, and a regime-related question can activate pressure to offer an agreeable answer regardless of private belief, making *ta'arof* a natural empirical anchor for the social deference construct in our framework. The second is the coercive capacity of the Iranian state. Security institutions comprising the Islamic Revolutionary Guards Corps, the Basij paramilitary organization, the judiciary, and intelligence services exercise control "outside the scope of the authority of the president" and regularly suppress dissent through mass arrests and arbitrary detention (Enayat & Künkler 2025, p. 410). Repression in Iran is real, recurrent, and widely understood by citizens to be a feature of public life. Together, these conditions allow us to test whether norms of deference and coercive pressure contribute to preference falsification.

Civic resources also vary in ways that complicate a straightforward story of fear or conformity. The

post-revolutionary expansion of schooling, including university enrollment among women, means that formal education is widespread but still unevenly distributed across generation, gender, and place. Income is likewise heterogeneous, shaped by sanctions, inflationary cycles, and the divide between households tied to public-sector employment and those dependent on private or informal work. Access to news and public affairs also varies, with some citizens relying mainly on state media and others regularly consuming non-state sources. These forms of variation allow us to assess whether education, income, and political awareness narrow the gap between public expression and private evaluation. Iran is, in sum, a context in which expressive identification, behavioral conformity, and civic resources are simultaneously present and variable.

National opinion surveys have nevertheless been conducted in Iran since the mid-1970s, when the first nationwide values survey documented a growing religious undercurrent that the government dismissed and the research community largely ignored (Asadi & Tehranian 2016). Surveys have continued under the Islamic Republic, and practitioners with experience in the country have noted that Iranian surveys produce results "more conservative than the reality of society," attributing the gap to respondent caution and rising nonresponse (Gudarzi 2024). Overreporting is therefore plausible in Iran and acknowledged inside the country's survey establishment, motivating the indirect measurement approach we employ.

Research design

Our data are drawn from a pre-registered nationally representative telephone (CATI) survey of adults aged 18 and over in Iran, conducted from February 20, 2025, to March 18, 2025. A total of 1,999 respondents completed the survey. All interviews were monitored in real-time by call-center supervisors. We worked closely with IranPoll, one of the most reputable survey organizations in Iran, with a track record of accurately predicting past Iranian elections (Conduit & Akbarzadeh 2020). We

assess the representativeness and reliability of the data in Appendix A. The probability sample was drawn using a six-stage sampling procedure. All 31 Iranian provinces were included, with interviews allocated proportionally to each province's share of the national population. Within each province, the sample was stratified by settlement type into five tiers based on population size, from cities of one million or more to rural settlements, using Iran's 2016 National Population and Housing Census. Primary sampling units were randomly selected from each settlement tier. Within each unit, landline telephone exchanges were randomly selected as secondary sampling units, with no more than 10 interviews per exchange. Random digit dialing was used to reach households within each exchange. Finally, a single respondent was randomly selected from eligible adults in the household. Up to three contact attempts were made per household, no substitution of selected respondents was permitted, and callback appointments were arranged when the selected respondent was unavailable. No major political shocks or disruptive events occurred during the fieldwork period.[1]

The double list experiment

We use two list experiments (also known as the item count technique) to estimate the prevalence of regime support in Iran. As we describe above, in authoritarian contexts, direct questions about regime support are likely to produce inflated estimates. A large body of work shows that this is the case: respondents overreport their trust in or support for the regime due to social desirability bias. List experiments address this by embedding a sensitive item within a list of nonsensitive items and asking respondents how many items on the list apply to them—but not which ones. This provides respondents with a degree of anonymity, reducing the perceived cost of revealing a socially undesirable attitude. The list experiment is widely used in studies of sensitive political behavior. Our decision to use the list experiment was driven in part by our interest in comparison. To our knowledge, every experimental study of regime support in authoritarian contexts that relies on sensitive question techniques uses the

[1]Ethical approval was granted by institutional review boards at Pennsylvania State University (STUDY00025559) and the University of California, Los Angeles (IRB-24-5517).

list experiment. Table 1 summarizes these studies.

Table 1: Estimates of regime support from list and direct questions in authoritarian regimes

Country (year)	Attitude object	List	Direct	Diff	Source
China (2013)*	Central government leaders	88%	92%	-4%	Tang (2016)
China (2014)	Political regime	56	73	-17	Dai (2019)
China (2015)	Central government	30	31	-1	Chen & Yang (2019)
China (2016)	Central government leaders	62	90	-28	Li *et al.* (2018)
China (2017)	National government	66	91	-25	Robinson & Tannenberg (2019)
China (2018)	Central government	77	90	-13	Nicholson & Huang (2022)
China (2020)	Xi Jinping	73	96	-23	Carter *et al.* (2024)
Ethiopia (2020)*	Abiy Ahmed	14	53	-39	Norris (2022)
Kenya (2021)*	Uhuru Kenyatta	43	39	4	Norris (2022)
Kyrgyzstan (2020)*	Sooronbai Jeenbekov	47	50	-3	Norris (2022)
Nicaragua (2020)*	Jose Daniel Ortega	5	39	-34	Norris (2022)
Philippines (2019)*	Rodrigo Duterte	52	76	-24	Norris (2022)
Russia (2015)*	Vladimir Putin	79	88	-9	Frye *et al.* (2017)
Russia (2022)*	Vladimir Putin	63	84	-21	Frye *et al.* (2023)
Turkey (2021)	Recep Tayyip Erdoğan	34	42	-8	Neundorf *et al.* (2024)
Venezuela (2021)*	Nicolas Maduro	22	27	-5	Norris (2022)
Zimbabwe (2020)*	Emmerson Mnangagwa	36	41	-5	Norris (2022)

Note: "*" indicates nationally representative survey.

The sensitive item in our list experiments asks about the extent of trust in the "authorities of the Islamic Republic" (*mas'ulan-e jumhuri-ye islami*).[2] We chose this wording deliberately. Rather than asking about the Supreme Leader (Ali Khamenei) at the time of the survey or any other specific politician, we ask about the regime as a political system. This captures systemic regime support rather than approval of a specific politician. This distinction matters in a political system where citizens may hold different attitudes toward the institutional order and the individuals who occupy its offices—particularly since the Supreme Leader is a religious figure. In Iran, an elected president is charged with day-to-day government administration, making the distinction between the regime and its leaders unusually important in this context. In the list experiment, the item is framed as one of several groups the respondent may or may not trust: "How many of the following do you trust?" In the direct question, which uses the same term, the respondent is asked to agree or disagree with a

[2]The Persian word used for trust is *i'timad*, which can be translated as both trust for and confidence in. Our usage of this word reflects our aim to capture trust for and confidence in the regime while avoiding stronger words for support, such as *himayat*.

statement: "I trust the authorities of the Islamic Republic."

We implement a double list experiment (DLE) to improve the precision of our estimates. Underpowered list experiments are known to produce wide confidence intervals because they rely on indirect, differential estimation between treatment and control groups. Estimators typically have high variance because they inherit noise from both groups, often producing extremely wide confidence intervals (and imprecise estimates). Because we are specifically interested in estimating overreporting and in detecting subgroup differences in both regime support and overreporting, statistical power is critical. Implementing two independent list experiments with the same sensitive item and showing that they produce convergent estimates also helps us assess the reliability of our experiment—which should increase confidence in our findings. Each respondent is randomly assigned to receive one treatment list and one control list. Respondents assigned to the treatment condition for List 1 receive the control condition for List 2, and vice versa. This ensures that every respondent sees the sensitive item exactly once. All respondents receive List 1 first, followed by List 2. Our approach yields a doubled dataset in which each respondent contributes two observations—one from each list—which we pool in our analysis.

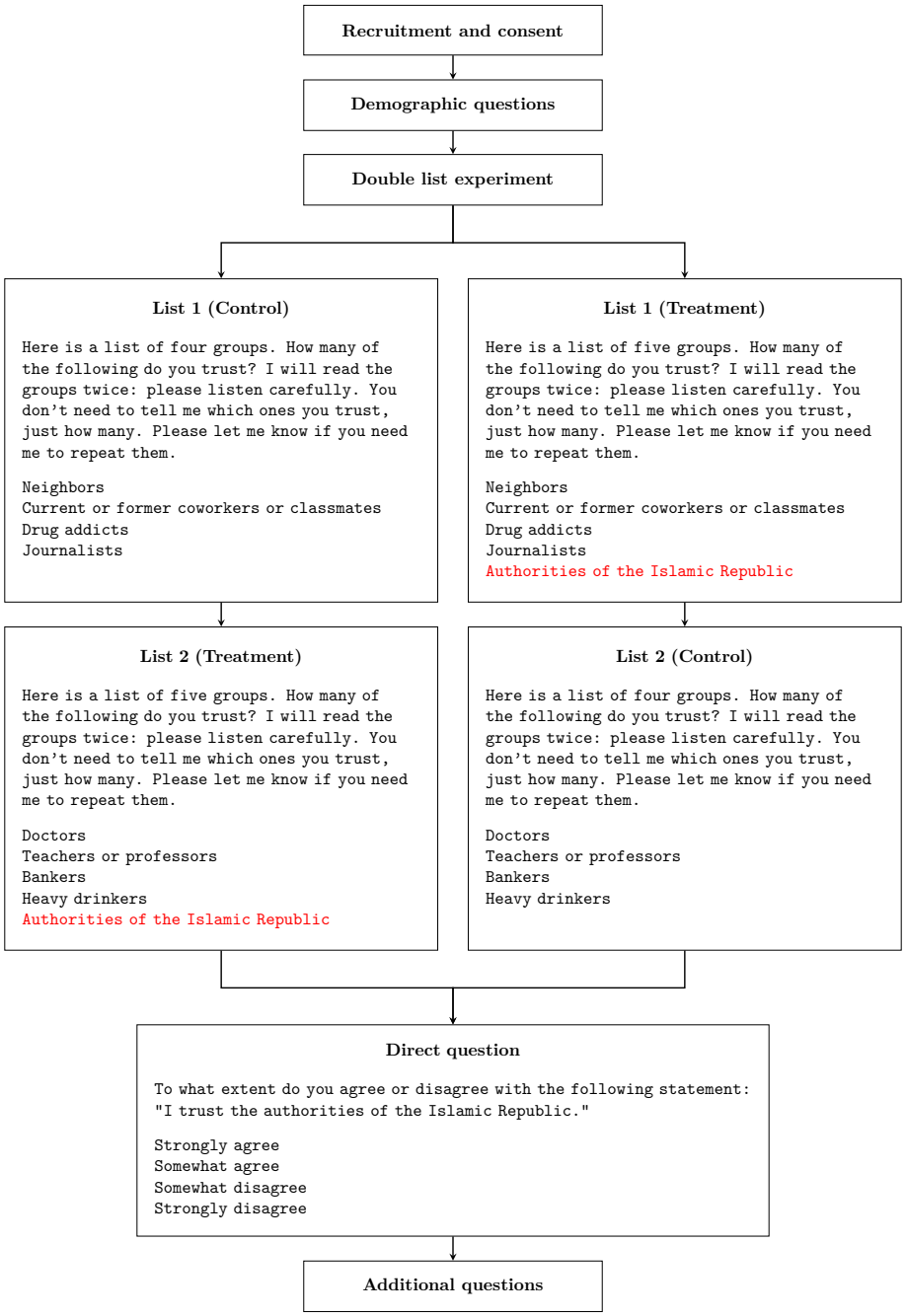
List 1 contains four control items (neighbors; current or former coworkers or classmates; drug addicts; journalists) plus the sensitive item in the treatment condition. List 2 contains four different control items (doctors; teachers or professors; bankers; heavy drinkers) plus the same sensitive item. Control items were selected following standard design principles for list experiments: the items were chosen to produce control group means of approximately 1 to 2 affirmative responses, minimizing the risk of ceiling and floor effects, and control items should be uncorrelated with attitudes towards the regime. Items within each list are read to respondents in randomized order to minimize order effects. Each list is read twice before respondents provide their count. We do this in an effort to minimize measurement error, since respondents cannot see the items and must hold them in memory while aggregating their responses. Enumerators explicitly instructed respondents to report only how many

groups they trust, not which ones, and offered to repeat the items if needed.

Following both list experiments, all respondents were asked a direct question: "To what extent do you agree or disagree with the following statement: 'I trust the authorities of the Islamic Republic,' " with four response options (strongly agree, somewhat agree, somewhat disagree, strongly disagree). The direct question is placed after the list experiments, as is standard practice, to avoid contaminating responses to the list experiment. We construct a binary measure of regime support from this question, coding respondents who "strongly agree" or "somewhat agree" as supporters and respondents who "strongly disagree" or "somewhat disagree" as non-supporters. Twenty-six respondents (1.3% of the sample) declined to answer the direct question. Figure 2 presents a visual overview of the survey.

Two assumptions are required for estimates from list experiments to be valid: the "no design effects" and "no liars" assumptions. The "no design effects" assumption holds that the addition of the sensitive item should not change responses to control items. The "no liars" assumption holds that respondents do not strategically or non-strategically misrepresent attitudes towards the sensitive item. In other words, the difference in mean item counts between treatment and control groups is an unbiased estimate of the prevalence of the sensitive attitude. We test both assumptions in detail in Appendices B.1 and B.2 and find no evidence that either is violated.

Figure 2: Survey flow



Note: Survey structure. In the double list experiment, each respondent receives one treatment list and one control list, ensuring every respondent sees the sensitive item (shown in red) exactly once. All respondents complete List 1 before List 2, and answer the direct question after both lists.

Independent variables

We organize our independent variables into a set of demographic variables and three sets of independent variables consistent with the sources we describe above. Our demographic variables are constructed as follows. `Female` is a binary indicator coded by the enumerator. `Age` is computed from year of birth and collapsed into five groups: 18-29, 30-39, 40-49, 50-59, and 60 and above. `Rural` is binary, indicating whether the respondent resides in a rural settlement. `Minority` is binary, coded 1 for any ethnicity other than Persian, including Turk/Azeri, Kurd, Lur, Arab, Baluch, Gilak/Mazandarani, and others.

Two variables capture expressive identification. `Pride` is measured on a four-point scale asking how proud the respondent is to be Iranian, coded so that higher values indicate greater pride. *Religiosity* is a binary measure based on frequency of prayer (*namaz*), coded 1 if the respondent reports praying several times a day (the maximum frequency on a seven-point scale), and 0 otherwise.

Three variables capture behavioral conformity. `Fear` of repression is the average of two items asking respondents how likely they would be to face punishment for participating in a protest and for criticizing the Islamic Republic on social media, coded so that higher values indicate greater perceptions of fear. `Self-censorship` is the average of two items: "It is difficult for me to express my opinion if I think others won't agree with what I say" (reverse-coded so that higher values indicate more self-censorship) and "If I disagree with others, I have no problem letting them know it" (coded so that agreement indicates less self-censorship). `Taarof` is our operationalization of social deference in the Iranian context: a culturally specific norm of ritualized politeness, modesty, and conflict-avoidance that maps onto the cultural-script mechanism in H_{2c}. We measure it with a single item asking respondents how often they use *ta'arof*, coded so that higher values indicate more frequent use. All three variables are grouped into three ordered categories using quantile-based cut points from the empirical distribution.

Four variables capture civic resources. `Education` collapses an 11-category classification into three groups: below high school, high school diploma, and post-secondary education. `Income` collapses eight monthly household income brackets (in Iranian tomans) into four groups. `Political awareness` is the average of three components, each rescaled to run from 0 to 1: political news consumption (four levels), political interest (four levels), and political knowledge (binary, coded 1 if the respondent correctly estimates the number of Majles-e Shura representatives as between 260 and 320, and 0 otherwise, including "don't know" responses). The resulting index is grouped into three ordered categories using quantile-based cut points from the empirical distribution. `Efficacy` is the average of four items: two internal efficacy items ("Sometimes politics and government seem so complicated that a person like me can't really understand what's going on" and "I feel that I have a pretty good understanding of the important political issues facing our country," the latter reverse-coded) and two external efficacy items ("People like me don't have any say about what the government does" and "Public officials don't care much what people like me think"). All items are coded so that higher values indicate greater efficacy; the index is grouped into three ordered categories.

Estimation

We estimate list experiment prevalence of regime support using nonlinear least squares (NLS) via the `ictreg` function in the `R list` package (Blair & Imai 2012). NLS requires only that the conditional mean functions are correctly specified and makes no distributional assumptions about the full response distribution. The alternative maximum likelihood (ML) estimator models each cell of the response distribution, including extreme cells where treatment group respondents score either zero or the maximum possible value. While ML offers greater statistical efficiency, it relies on observations in exactly those extreme cells. But well-designed list experiments that follow best practices for minimizing ceiling and floor effects actively reduce the number of respondents in those cells. In our data, the Hausman specification test proposed by Blair *et al.* (2019) yields negative test statistics for both lists,

indicating that the ML distributional assumptions are not satisfied. As a result, we prefer to use NLS for our analyses. A detailed discussion of the NLS-ML comparison and Hausman test results can be found in Appendix B. For the aggregate DLE estimates, standard errors are obtained from 1,000 bootstrap iterations clustered by respondent, as each respondent appears twice in the pooled dataset. List 1 and List 2 estimates are reported separately alongside the DLE estimates of regime support and overreporting.

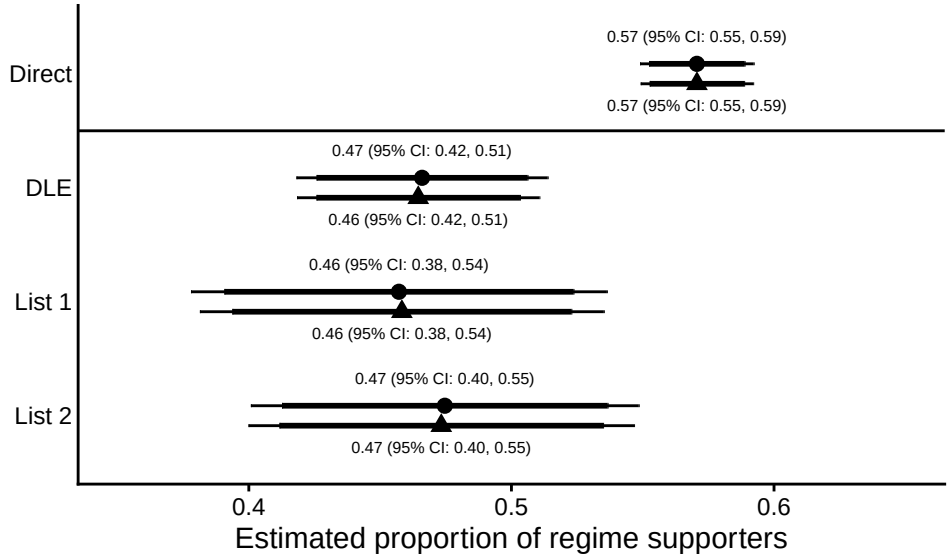
To assess which respondent characteristics predict regime support and overreporting, we estimate separate models for each independent variable. For each one, we fit an `ictreg` model on the pooled DLE data with that variable as the sole covariate, and a logistic regression on the direct question binary with the same variable. The variance-covariance matrix for the list experiment model is obtained via cluster bootstrap (1,000 iterations, clustered by respondent). Reporting subgroup differences in overreporting requires combining estimates from these two separately fit models. Because overreporting is the difference between two predicted probabilities—one from the list model and one from the direct question—and our key quantity is a difference of such differences across subgroups, it has no analytic standard error, and reporting the list and direct estimates separately would understate the uncertainty in their comparison. We therefore use parametric simulation (King *et al.* 2000) to carry the full sampling uncertainty of both models through to each quantity of interest, an approach also used in recent work comparing list and direct estimates across respondent subgroups (Nicholson & Huang 2022). For each model, we draw 10,000 parameter vectors from the multivariate normal distribution defined by the point estimates and the bootstrap variance-covariance matrix (for the list model) or the analytic variance-covariance matrix (for the direct question model). For each draw, we compute average predicted probabilities for each subgroup via the inverse logit, yielding a simulated list estimate, direct estimate, and their difference (overreporting) for each subgroup. Between-group differences are computed by comparing the highest and lowest values of the variable. Standard errors are the standard deviations of the 10,000 simulation draws.

Results

Main results

We begin by presenting aggregate estimates of regime support and overreporting before turning to subgroup analyses. Figure 3 displays estimated regime support from the direct question and the three list-based estimators (DLE, List 1, List 2), and Figure 4 displays estimated overreporting for each. All results are reported with both unadjusted and covariate-adjusted specifications; adjusted models control for `Female`, `Age`, `Rural`, and `Minority`.

Figure 3: Regime support: Direct and list estimates



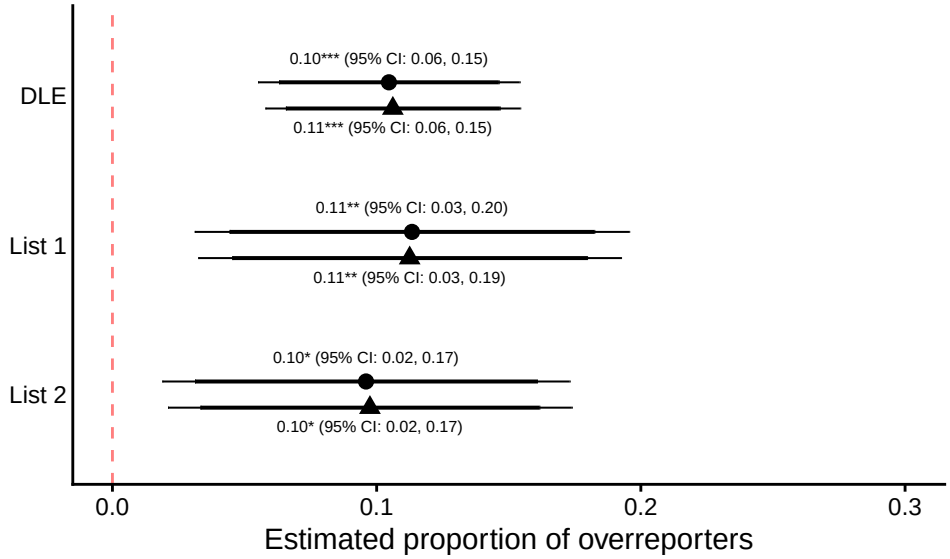
Note: Estimated proportion of regime supporters using the direct question, the double list experiment (DLE), List 1, and List 2. Point estimates are shown with 90% (thick) and 95% (thin) confidence intervals. Unadjusted models (circles) include no covariates; adjusted models (triangles) control for gender, age, settlement, and minority status. The DLE estimate pools both lists using the doubled dataset with standard errors clustered by respondent via 1,000 bootstrap iterations. The horizontal line separates the list-based estimates from the direct question estimate.

The direct question indicates that an estimated 57.06% of respondents support the regime (95% CI: [0.549, 0.592]). The list experiment estimates are substantially and consistently lower. The DLE, which pools both list experiments using the doubled dataset with standard errors clustered by respondent, estimates regime support at 46.60% unadjusted (95% CI: [0.418, 0.514]) and 46.45% adjusted (95% CI: [0.418, 0.511]).[3] The individual list experiments produce similar estimates: List 1

[3] In a separate scale robustness check, we experimentally varied the direct response format between a binary two-point

yields 45.72% unadjusted (95% CI: [0.378, 0.536]) and 45.83% adjusted (95% CI: [0.381, 0.535]), while List 2 yields 47.46% unadjusted (95% CI: [0.401, 0.549]) and 47.33% adjusted (95% CI: [0.400, 0.547]). The close correspondence between List 1 and List 2 is consistent with the expectation that the two list experiments are measuring the same thing. The narrower confidence intervals on the DLE reflect efficiency gains from pooling. Covariate adjustment has negligible effects on point estimates across all specifications.

Figure 4: Overreporting: Direct and list differences



Note: Estimated proportion of overreporters (difference between direct and list estimates). Positive values indicate that the direct question overstates support relative to the indirect measure. Point estimates are shown with 90% (thick) and 95% (thin) confidence intervals. Unadjusted models (circles) include no covariates; adjusted models (triangles) control for gender, age, settlement, and minority status. The dashed red line marks zero (no overreporting). $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

The difference between the direct question and list experiment estimates—the overreporting gap—is positive, substantively large, and statistically significant across all specifications. The DLE estimates that 10.46% of respondents overreport regime support in the unadjusted model (95% CI: [0.055, 0.154], $p < 0.001$) and 10.61% in the adjusted model (95% CI: [0.058, 0.154], $p < 0.001$). List 1 produces overreporting estimates of 11.34% unadjusted (95% CI: [0.031, 0.196], $p < 0.01$) and 11.25% adjusted (95% CI: [0.033, 0.193], $p < 0.01$). List 2 yields 9.60% unadjusted (95% CI: [0.019, 0.173], $p < 0.05$)

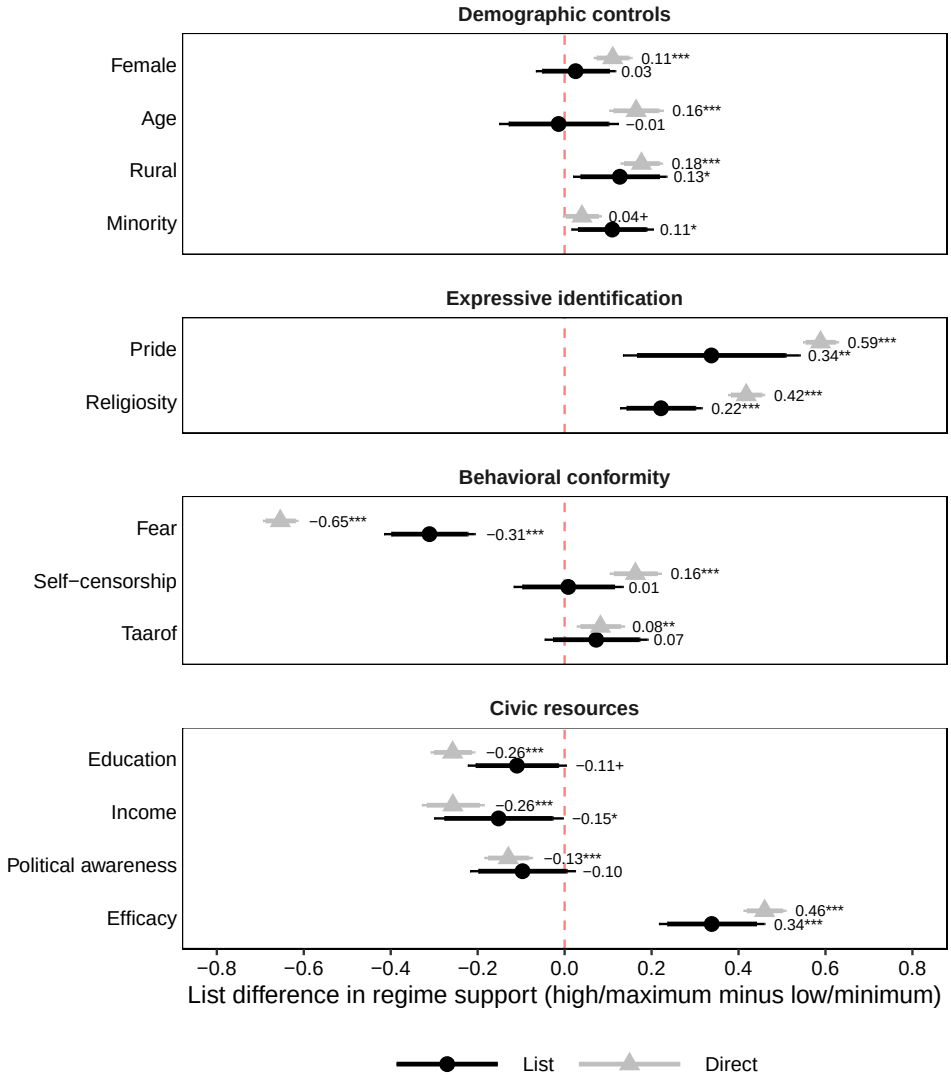
scale (agree or disagree) and the four-point scale we use above. Using the double list estimate of regime support at 46-47 percent as the benchmark, overreporting is about 8-9 percentage points with the binary scale and 10-11 percentage points with the four-point scale that we use in the main analysis. The binary option yields a lower direct support rate by roughly 2.5 points.

and 9.75% adjusted (95% CI: [0.021, 0.174], $p < 0.05$). The consistency of the overreporting gap across all three formats—ranging from approximately 10 to 11 percent—indicates that overreporting is not an artifact of a particular list or estimation strategy. In substantive terms, roughly one in ten respondents express trust in the authorities of the Islamic Republic when asked directly but do not do so when afforded the anonymity of the list experiment—accounting for approximately 18% of all individuals who express support for the regime when asked directly. Estimates for both list experiments and the double list experiment can be found in Tables A.7 and A.8.

Subgroup differences

We now turn from aggregate prevalence to subgroup variation. Figure 5 presents subgroup differences in regime support. Although our hypotheses concern overreporting rather than regime support itself, these descriptive patterns provide useful baseline context for the differences we examine in Figure 6. We focus here on the double list experiment estimates, which give the cleanest read of regime support across subgroups. The strongest positive predictors are the variables we group under expressive identification: respondents at the highest level of national pride are 34 percentage points more likely to privately support the regime than those at the lowest, and respondents who pray several times a day are 22 points more likely than those who do not. Civic resources tell a more muted story than existing work on Iran might suggest. Education, income, and political awareness all predict lower private support, but only income is statistically significant ($p < 0.05$); education's negative association is marginal, and political awareness is statistically indistinguishable from zero. Efficacy is an exception, predicting higher rather than lower support.

Figure 5: Subgroup differences and regime support



Note: Difference in predicted probability of regime support between respondents at the maximum and minimum values of each variable, estimated from bivariate NLS regressions (list) and logistic regressions (direct). Point estimates are shown with 90% (thick) and 95% (thin) confidence intervals. Predicted probabilities are computed via parametric simulation (10,000 draws); list model variance-covariance matrices are obtained from 1,000 bootstrap iterations clustered by respondent. $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

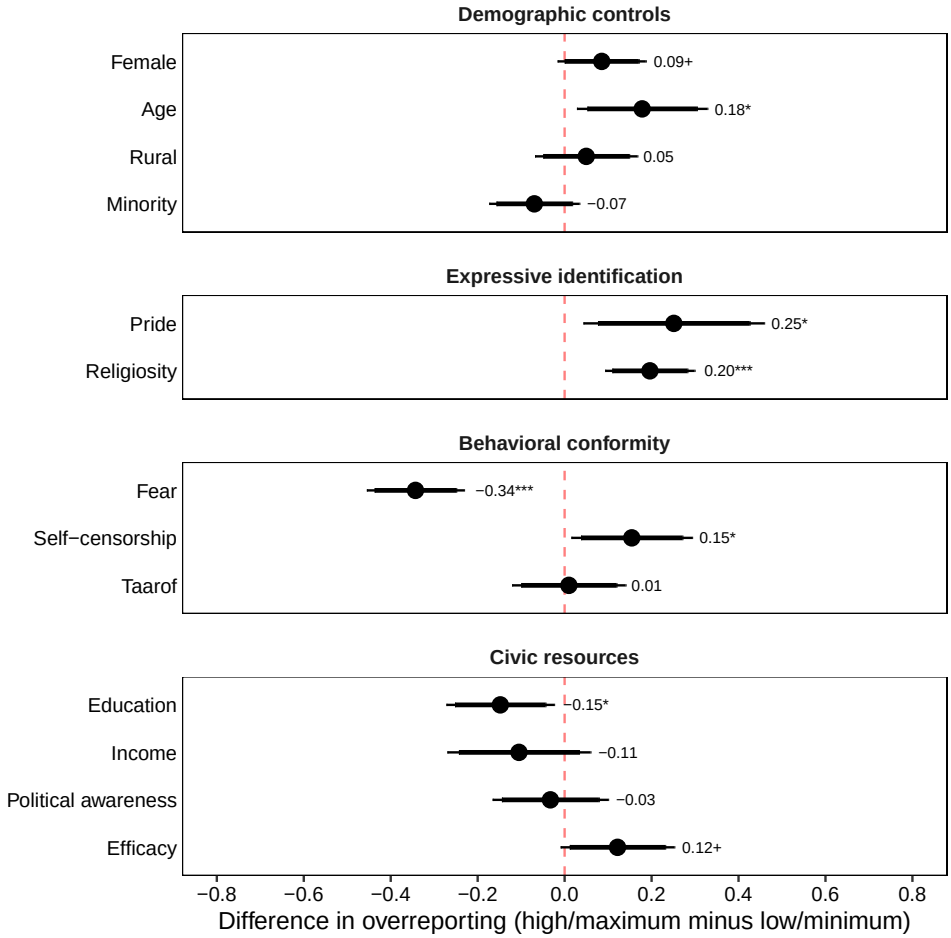
Behavioral conformity yields mixed findings. Fear of repression is the strongest negative predictor of regime support: the most fearful respondents are 31 percentage points less likely to privately support the regime than those at the lowest level of fear. Self-censorship and *ta'arof*, in contrast, are unrelated to regime support. The pattern suggests that our measurement of fear is not capturing perceptions of the regime's coercive capacity; if anything, fearful respondents are more critical. We return to this point in our discussion of overreporting below. Demographic patterns are more modest predictors of

regime support: rural respondents and ethnic minorities are somewhat more likely than urban and ethnic Persians to support the regime; gender and age differences are statistically indistinguishable from zero.

Figure 6 turns to overreporting. The pattern offers clear support for expressive identification as a predictor of preference falsification in Iran, but more mixed support for the other two sources in our framework. Both expressive identification variables (H_{1a} and H_{1b}) are positive and statistically significant. National pride is associated with a 25 percentage point increase in overreporting, and religiosity with a 20 point increase; both align with our hypothesized predictions and both are large relative to the aggregate overreporting rate of about 10 percent. Respondents who report the highest levels of national pride and those who pray several times a day are appreciably more likely than other respondents in our sample to express support for the authorities of the Islamic Republic that they do not privately hold. The internalized referents at the heart of this source—the nation and the religious community—appear to exert the kind of identity-based pressure on public expression that our framework predicts.

Behavioral conformity (H_{2a} to H_{2c}) yields a more complex picture. Self-censorship works as predicted: respondents who manage political expression in daily life overreport by 15 percentage points more than those who do not. *Ta'arof* is statistically indistinguishable from zero. Fear of repression, however, is significant in the opposite direction: respondents at the highest level of fear overreport by 34 percentage points less than those at the lowest. We read this not as evidence against the broader logic of behavioral conformity but as a reflection of what our measurement of fear captures. The scale's two items ask respondents to imagine being punished for participating in a protest or for criticizing the regime online—counterfactual scenarios that a respondent who sees themselves as a potential dissenter engages seriously, while a regime supporter dismisses as inapplicable. The variable therefore selects for an oppositional profile rather than for fear among ambivalent respondents, and respondents with high fear scores in our sample have little private support to conceal in the first place.

Figure 6: Subgroup differences and overreporting



Note: Difference in overreporting of regime support between respondents at the maximum and minimum values of each variable. Overreporting for each subgroup is defined as the bivariate direct question predicted probability minus the bivariate list experiment predicted probability; the plotted quantity is the difference-of-differences across subgroups. Positive values indicate that respondents at the high end of a variable overreport more than those at the low end. Point estimates are shown with 90% (thick) and 95% (thin) confidence intervals. $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Civic resources (H_{3a} to H_{3d}) are the most uneven of the three sources. Education works as predicted: more educated respondents overreport by 15 percentage points less than less educated respondents. Income runs in the predicted direction but does not reach conventional levels of statistical significance, and political awareness is statistically indistinguishable from zero. Efficacy is marginally positive—the opposite of the predicted sign—mirroring its positive association with regime support. We suspect this reflects how the efficacy items work in the Iranian context: agreeing that officials respond to people like oneself, for example, may track endorsement of the political system rather than the independent civic capacity the construct is intended to capture.

Demographic patterns are modest. Older respondents overreport by about 18 percentage points more than younger ones, women marginally more than men, and rural respondents and ethnic minorities do not overreport. Taken together, the subgroup results provide clear support for expressive identification as a predictor of preference falsification, partial support for behavioral conformity (through self-censorship, but not through *ta'arof* or our measure of fear), and still weaker support for civic resources (through education alone). The strongest mechanisms in our data are the most internalized—the variables most closely tied to the regime's legitimating narratives. Variables that capture externally imposed pressure or independence from regime narratives do less explanatory work, suggesting that preference falsification in Iran is shaped less by the coercion citizens face from the state than by the identities citizens bring to the survey setting. Complete estimates of both sets of findings are printed in Tables A.9 and A.10. Covariate-adjusted estimates—which do not change the size, direction, or significance of our estimates—can be found in Figures A.4 and A.5.

Robustness

In this section, we share the results of a series of diagnostic and robustness tests. Each test is reported in the Supplementary information appendices, but we summarize several key findings here. We begin with an assessment of the reliability of our data. In order to detect duplication and fraud, we compute the highest percentage of matching response values between each respondent and every other respondent across all survey variables (Kuriakose & Robbins 2016). No observation reaches the 85% threshold for suspected duplicates, providing no evidence of data fabrication via duplication (Figure A.1). To verify that treatment assignment is independent of respondent characteristics, we conduct χ^2 tests of independence between each variable used in the analysis and treatment assignment (Table A.3). Of the 28 tests, only one variable reaches significance at the 5% level.

Next, we evaluate the "no design effects" assumption, which requires that the sensitive item does not alter responses to the control items. Following Blair & Imai (2012), all estimated respondent-type

proportions are non-negative for both lists, and the Bonferroni-corrected p-value equals 1.00 for each (Table A.4). To test for mechanical inflation—the concern that treated respondents report higher counts simply because their list is longer—we administer a separate list experiment with a near-zero-prevalence placebo statement as suggested by Riambau & Ostwald (2021). The treatment-control difference is negligible (2.40 versus 2.41, $p = 0.41$), with no evidence of inflation among subgroups most theoretically vulnerable to satisficing behavior (Table A.5).

We also assess the "no liars" assumption by examining the observed response distributions for evidence of ceiling and floor effects (Table A.6). No treatment group respondent scored the maximum of 5 on either list, ruling out ceiling effects. Floor effects are a more substantive concern—9.52% of List 1 treatment respondents scored 0, but maximum likelihood models with floor corrections produce identical prevalence estimates. We discuss these corrections and provide additional discussion in Appendix B.2.

We also benchmark the list experiment against a non-sensitive item. A separate list experiment using favorability toward Russia—a topic that is not politically sensitive in Iran—produces an over-reporting gap that is small and statistically indistinguishable from zero in both unadjusted (0.05, $p = 0.27$) and adjusted (0.06, $p = 0.22$) specifications (Figure A.2). We also replicate the analysis presented in Figure 6 in Figure A.3: no subgroup exhibits overreporting estimates that are statistically distinguishable from zero. This null result provides evidence that the overreporting we observe for regime support reflects preference falsification—rather than an artifact of the list experiment format.

Last, because we test multiple subgroup hypotheses, we assess sensitivity to multiple comparisons using the Benjamini-Hochberg procedure applied to the nine hypothesized predictor variables (Benjamini & Hochberg 1995). For regime support, all four variables that reach significance without correction survive adjustment at $p < 0.05$: `Religiosity`, `Fear`, `Efficacy`, and `Pride`. For over-reporting, four of the five variables that reach significance without correction survive adjustment: `Fear`, `Religiosity`, `Pride`, and `Education`. `Self-censorship` is attenuated to marginal significance

($p = 0.053$). Full results are reported in Table A.11.

Conclusion

This paper has used a pre-registered, nationally representative telephone survey of Iranian adults with an embedded double list experiment to estimate regime support and overreporting in Iran. Our list-based estimate places regime support at approximately 47 percent, ten percentage points below the 57 percent recorded by direct questioning. The gap is statistically significant, robust across both list experiments and the pooled DLE, and unaffected by covariate adjustment, mechanical inflation, and a battery of additional diagnostics. Subgroup analyses provide clear support for one of our three theoretical sources—expressive identification, captured by national pride and religiosity—while behavioral conformity and civic resources yield more mixed findings. The most internalized referents are the strongest predictors of public expression that diverges from private belief.

Our first contribution to the literature is conceptual. Scholars have invested heavily in measuring around preference falsification, and we have learned that the gap between public and private expression in authoritarian regimes is real and substantial—though variable across contexts (Frye *et al.* 2017; Jiang & Yang 2016; Robinson & Tannenberg 2019; Hale 2022; Carter *et al.* 2024; Kasuya & Miwa 2023). We have learned much less about who falsifies and why. Treating overreporting as a substantive political behavior, rather than a measurement problem to be corrected, introduces a research agenda that intersects with longstanding questions about the sources of regime support in authoritarian politics, the nature of the dictator's dilemma, and the conditions under which citizens engage in preference falsification (Kuran 1991, 1995; Wintrobe 1998; Gerschewski 2013). Anchoring our framework to Blair *et al.*'s (2020) social reference theory of sensitivity bias allows us to ask which social referents—the religious community, the nation, the state, or the enumerator—are most salient to a given respondent, and to organize a broad set of mechanisms that have been studied separately into a single, referent-

based logic. The framework provides analytical scaffolding for what we believe is an under-theorized phenomenon.

The second contribution is substantive. The dominant prior in this literature, traceable to Kuran (1991, 1995) and reflected in much empirical work that followed (Wintrobe 1998; Davenport 2007), holds that fear of the security apparatus is the principal engine of preference falsification in authoritarian regimes. Our findings in Iran point in a different direction. Identity-based attachments to the nation and the religious community—the variables most closely tied to the legitimating narratives of the Islamic Republic (Collins & Owen 2012; Hoffman 2020; Tezcür *et al.* 2012; Neundorf *et al.* 2024)—are stronger predictors of preference falsification than variables connected to behavioral conformity or civic capacity. We are cautious in extending this claim beyond Iran. A single-country study cannot establish that identity-based legitimation is more important than coercion in shaping preference falsification across the authoritarian landscape, and Iran is unusual in the depth of fusion between religious and political authority. What our findings do suggest is that scholars who interpret cross-national variation in overreporting purely as a function of repression intensity may be overlooking an alternative source of variation: differences in how successfully regimes have cultivated normative attachments that citizens carry into survey settings.

We see two avenues for future research. The first is theoretical: if the magnitude and structure of preference falsification depend on which social referents are salient to citizens, then variation across regimes should track differences in narratives of legitimation as much as variation in repression. Regimes that have invested heavily in identity-based legitimation—through nationalist mobilization, religious authority, or charismatic appeals (Xu & Zhao 2023; Neundorf *et al.* 2024; Gerschewski 2013)—should produce different patterns of overreporting than regimes that rely primarily on patronage or coercion. Comparable list experiments fielded across regimes that vary along these sources, with shared sensitive items and a common set of theoretically motivated covariates, would allow scholars to test this prediction directly. The second is substantive and adjacent to the comparative study of

authoritarian stability. Recent scholarship on authoritarian persistence has increasingly emphasized the role of citizen attachments—to the nation, to religious tradition, or to a personalist leader—as a stabilizing force that operates independently of performance (Neundorf *et al.* 2024; Hale 2022; Buckley *et al.* 2024; Xu & Zhao 2023). Our findings provide individual-level evidence that these attachments shape not only what citizens privately believe but what they are willing to say in public, and that the gap between the two is itself politically meaningful. Aggregated across a population, public affirmations produce the appearance of consensus on which authoritarian stability partly rests (Kuran 1991; Hale 2022; Buckley *et al.* 2024).

Preference falsification in Iran is shaped less by the coercion citizens face from the state than by the identities the regime has encouraged them to hold. Whether this finding generalizes is an empirical question for future research, and one we cannot settle with evidence from a single country. What is clear is that overreporting is too consequential a political behavior, and too tractable an empirical target, to remain an artifact at the margins of advances in survey methodology.

References

Ahlquist, John S. 2018. List Experiment Design, Non-Strategic Respondent Error, and Item Count Technique Estimators. *Political Analysis* 26(1), 34–53.

Almond, Gabriel A. and Sidney Verba. 1963. *The Civic Culture: Political Attitudes and Democracy in Five Nations*. Princeton: Princeton University Press.

Asadi, Ali and Majid Tehranian. 2016. *Seda'i Ke Shenideh Nashod: Negarash-Ha-Ye Ejtema'i-Farhangi Va Towse'eh-Ye Namotavazen Dar Iran [A Voice That Was Not Heard: Social-Cultural Perspectives and Uneven Development in Iran]*. Tehran: Nashr-e Ney.

Bandura, Albert. 1982. Self-Efficacy Mechanism in Human Agency. *American Psychologist* 37(2), 122–147.

Benjamini, Yoav and Yosef Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* 57(1), 289–300.

Berinsky, Adam J. 1999. The Two Faces of Public Opinion. *American Journal of Political Science* 43(4), 1209–1230.

Bhasin, Tavishi and Jennifer Gandhi. 2013. Timing and Targeting of State Repression in Authoritarian Elections. *Electoral Studies* 32(4), 620–631.

Blair, Graeme and Kosuke Imai. 2012. Statistical Analysis of List Experiments. *Political Analysis* 20(1), 47–77.

Blair, Graeme, Kosuke Imai, and Jason Lyall. 2014. Comparing and Combining List and Endorsement Experiments: Evidence from Afghanistan. *American Journal of Political Science* 58(4), 1043–1063.

Blair, Graeme, Winston Chou, and Kosuke Imai. 2019. List Experiments with Measurement Error. *Political Analysis* 27(4), 455–480.

Blair, Graeme, Alexander Coppock, and Margaret Moor. 2020. When to Worry About Sensitivity Bias: A Social Reference Theory and Evidence from 30 Years of List Experiments. *American Political Science Review* 114(4), 1297–1315.

Buckley, Noah, Kyle L. Marquardt, Ora John Reuter, and Katerina Tertytchnaya. 2024. Endogenous Popularity: How Perceptions of Support Affect the Popularity of Authoritarian Regimes. *American Political Science Review* 118(2), 1046–1052.

Bussmann, Margit and Natalia Iost. 2024. Presidential Popularity and International Crises: An Assessment of the Rally-'Round-the-Flag Effect in Russia. *Post-Soviet Affairs* 40(2), 105–118.

Cantoni, Davide, Yuyu Chen, David Y. Yang, Noam Yuchtman, and Y. Jane Zhang. 2017. Curriculum and Ideology. *Journal of Political Economy* 125(2), 338–392.

Carlson, Elizabeth. 2018. The Perils of Pre-Election Polling: Election Cycles and the Exacerbation of Measurement Error in Illiberal Regimes. *Research & Politics* 5(2).

Carter, Erin Baggott, Brett L. Carter, and Stephen Schick. 2024. Do Chinese Citizens Conceal Opposition to the CCP in Surveys? Evidence from Two Experiments. *The China Quarterly* 259, 804–813.

Chang, Charles and Melanie Manion. 2021. Political Self-Censorship in Authoritarian States: The Spatial-Temporal Dimension of Trouble. *Comparative Political Studies* 54(8), 1362–1392.

Chen, Yuyu and David Y. Yang. 2019. The Impact of Media Censorship: 1984 or Brave New World? *American Economic Review* 109(6), 2294–2332.

Ciftci, Sabri. 2010. Modernization, Islam, or Social Capital: What Explains Attitudes Toward Democracy in the Muslim World? *Comparative Political Studies* 43(11), 1442–1470.

Ciftci, Sabri. 2013. Secular-Islamist Cleavage, Values, and Support for Democracy and Shari'a in the Arab World. *Political Research Quarterly* 66(4), 781–793.

Collins, Kathleen and Erica Owen. 2012. Islamic Religiosity and Regime Preferences: Explaining Support for Democracy and Political Islam in Central Asia and the Caucasus. *Political Research*

Quarterly 65(3), 499–515.

Conduit, Dara and Shahram Akbarzadeh. 2020. Pre-Election Polling and the Democratic Veneer in a Hybrid Regime. *Democratization* 27(5), 737–757.

Corstange, Daniel. 2018. Clientelism in Competitive and Uncompetitive Elections. *Comparative Political Studies* 51(1), 76–104.

Corstange, Daniel. 2022. Mass Dissimulation in the Syrian Civil War. *The Journal of Politics* 84(3), 1469–1481.

Croke, Kevin, Guy Grossman, Horacio A. Larreguy, and John Marshall. 2016. Deliberate Disengagement: How Education Can Decrease Political Participation in Electoral Authoritarian Regimes. *The American Political Science Review* 110(3), 579–600.

Dai, Yaoyao. 2019. Beloved Governments: Authoritarian Regimes' Toolkit for Building Popular Support. PhD diss., Pennsylvania State University.

Davenport, Christian. 2007. State Repression and Political Order. *Annual Review of Political Science* 10(1), 1–23.

Dimitrov, Martin K. 2009. Popular Autocrats. *Journal of Democracy* 20(1), 78–81.

Easton, David. 1965. *A Systems Analysis of Political Life*. New York: Wiley.

Enayat, Hadi and Mirjam Künkler, eds. 2025. *The Rule of Law in the Islamic Republic of Iran: Power, Institutions, and the Limits of Reform*. Cambridge: Cambridge University Press.

Frye, Timothy, Scott Gehlbach, Kyle L. Marquardt, and Ora John Reuter. 2017. Is Putin's Popularity Real? *Post-Soviet Affairs* 33(1), 1–15.

Frye, Timothy, Scott Gehlbach, Kyle L. Marquardt, and Ora John Reuter. 2023. Is Putin's Popularity (still) Real? A Cautionary Note on Using List Experiments to Measure Popularity in Authoritarian Regimes. *Post-Soviet Affairs* 39(3), 213–222.

Fukumoto, Kentaro and Takahiro Tabuchi. 2025. The Rally 'Round the Flag Effect in Third Parties: The Case of the Russian Invasion of Ukraine. *Journal of Elections, Public Opinion and Parties* 35(1), 171–180.

Gehlbach, Scott and Philip Keefer. 2012. Private Investment and the Institutionalization of Collective Action in Autocracies: Ruling Parties and Legislatures. *The Journal of Politics* 74(2), 621–635.

Gerschewski, Johannes. 2013. The Three Pillars of Stability: Legitimation, Repression, and Co-Optation in Autocratic Regimes. *Democratization* 20(1), 13–38.

Gingerich, Daniel W., Virginia Oliveros, Ana Corbacho, and Mauricio Ruiz-Vega. 2016. When to Protect? Using the Crosswise Model to Integrate Protected and Direct Responses in Surveys of Sensitive Behavior. *Political Analysis* 24(2), 132–156.

Glynn, Adam N. 2013. What Can We Learn with Statistical Truth Serum? Design and Analysis of the List Experiment. *Public Opinion Quarterly* 77(S1), 159–172.

Greene, Samuel A. and Graeme Robertson. 2022. Affect and Autocracy: Emotions and Attitudes in Russia After Crimea. *Perspectives on Politics* 20(1), 38–52.

Gudarzi, Muhsin. 2024. Natijih-I Paymayish-Ha Muhafazih-Karanih-Tar Az Vaqi'iyyat Ast [The Results of Surveys Are More Conservative Than Reality]. *Hammihan* 2(466), 12.

Gueorguiev, Dimitar D., Li Shao, and Charles Crabtree. 2017. *Blurring the Lines: Rethinking Censorship Under Autocracy*.

Guriev, Sergei and Daniel Treisman. 2019. Informational Autocrats. *The Journal of Economic Perspectives* 33(4), 100–127.

Guriev, Sergei and Daniel Treisman. 2020. The Popularity of Authoritarian Leaders: A Cross-National Investigation. *World Politics* 72(4), 601–638.

Hale, Henry E. 2022. Authoritarian Rallying as Reputational Cascade? Evidence from Putin's Popularity Surge After Crimea. *American Political Science Review* 116(2), 580–594.

Hintson, Jamie and Milan Vaishnav. 2023. Who Rallies Around the Flag? Nationalist Parties, National Security, and the 2019 Indian Election. *American Journal of Political Science* 67(2), 342–357.

Hoffman, Michael. 2020. Religion, Sectarianism, and Democracy: Theory and Evidence from Lebanon. *Political Behavior* 42(4), 1169–1200.

Hu, Qian and Yanping Pu. 2023. Using Unofficial Media, Less Trusting of Chinese Polity?—An Analysis Based on the Moderated Mediation Effect. *PLOS ONE* 18(6).

Huang, Haifeng. 2025. Reckoning with Reality: Correcting National Overconfidence in a Rising Power. *International Organization* .

Huang, Xian and Cai Zuo. 2023. Economic Inequality, Distributive Unfairness, and Regime Support in East Asia. *Democratization* 30(2), 215–237.

Huhe, Narisong, Min Tang, and Jie Chen. 2018. Creating Democratic Citizens: Political Effects of the Internet in China. *Political Research Quarterly* 71(4), 757–771.

Iglesias, Sol. 2025. Populist Appeal or Fearful Support? Examining the War on Drugs in the Philippines Under Duterte. *Pacific Affairs* 98(2), 217–244.

Imai, Kosuke. 2011. Multivariate Regression Analysis for the Item Count Technique. *Journal of the American Statistical Association* 106(494), 407–416.

Isani, Mujtaba and Bernd Schlipphak. 2023. Who Is Asking? The Effect of Survey Sponsor Misperception on Political Trust: Evidence from the Afrobarometer. *Quality & Quantity* 57(4), 3453–3481.

Jamal, Amaney A. and Mark A Tessler. 2008. The Democracy Barometers (Part II): Attitudes in the Arab World. *Journal of Democracy* 19(1), 97–110.

Jiang, Junyan and Dali L Yang. 2016. Lying or Believing? Measuring Preference Falsification From a Political Purge in China. *Comparative Political Studies* 49(5), 600–634.

Kalinin, Kirill. 2016. The Social Desirability Bias in Autocrat's Electoral Ratings: Evidence from the 2012 Russian Presidential Elections. *Journal of Elections, Public Opinion and Parties* 26(2), 191–211.

Kamrava, Mehran. 2023. *Righteous Politics: Power and Resilience in Iran.* Cambridge: Cambridge University Press.

Kamrava, Mehran. 2024. *How Islam Rules in Iran: Theology and Theocracy in the Islamic Republic.* Cambridge: Cambridge University Press.

Kang, Siqin and Jiangnan Zhu. 2021. Do People Trust the Government More? Unpacking the Distinct Impacts of Anticorruption Policies on Political Trust. *Political Research Quarterly* 74(2), 434–449.

Kao, Jay C. 2021. Family Matters: Education and the (Conditional) Effect of State Indoctrination in China. *Public Opinion Quarterly* 85(1), 54–78.

Karpowitz, Christopher F, Sarah Austin, Jacob Crandall, and Raquel Macias. 2023. Experimenting with List Experiments: Interviewer Effects and Immigration Attitudes. *Public Opinion Quarterly* 87(1), 69–91.

Kasuya, Yuko and Hirofumi Miwa. 2023. Pretending to Support? Duterte's Popularity and Democratic Backsliding in the Philippines. *Journal of East Asian Studies* 23(3), 411–437.

Kazemipur, Abdolmohammad. 2022. *Sacred as Secular: Secularization Under Theocracy in Iran.* Montreal: McGill-Queen's University Press.

King, Gary, Michael Tomz, and Jason Wittenberg. 2000. Making the Most of Statistical Analyses: Improving Interpretation and Presentation. *American Journal of Political Science* 44(2), 341–355.

Kobayashi, Tetsuro and Polly Chan. 2022. Political Sensitivity Bias in Autocratizing Hong Kong. *International Journal of Public Opinion Research* 34(4).

Kobayashi, Tetsuro, Jaehyun Song, and Polly Chan. 2025. Is a Snapshot Enough? Repeated List Experiments in Autocratizing Hong Kong Under the National Security Law. *Journal of Elections, Public Opinion and Parties* 1–15.

Kuhn, Patrick M and Nick Vivyan. 2018. Reducing Turnout Misreporting in Online Surveys. *Public Opinion Quarterly* 82(2), 300–321.

Kuran, Timur. 1991. Now Out of Never: The Element of Surprise in the East European Revolution of 1989. *World Politics* 44(01), 7–48.

Kuran, Timur. 1995. *Private Truths, Public Lies: The Social Consequences of Preference Falsification.* Cambridge: Harvard University Press.

Kuriakose, Noble and Michael Robbins. 2016. Don't Get Duped: Fraud Through Duplication in Public Opinion Surveys. *Statistical Journal of the IAOS* 32(3), 283–291.

Levi, Margaret. 1997. *Consent, Dissent, and Patriotism.* Cambridge: Cambridge University Press.

Li, Xiaojun, Weiyi Shi, and Boliang Zhu. 2018. The Face of Internet Recruitment: Evaluating the Labor Markets of Online Crowdsourcing Platforms in China. *Research & Politics* 5(1), 1–8.

Maghen, Ze'ev. 2023. *Reading Revolutionary Iran: The Worldview of the Islamic Republic's Religio-Political Elite.* Berlin: De Gruyter.

McCauley, John, Steven Finkel, Michael Neureiter, and Christopher Belasco. 2022. The Grapevine Effect in Sensitive Data Collection: Examining Response Patterns in Support for Violent Extremism. *Political Science Research and Methods* 10(2), 353–371.

Moruzzi, Norma Claire. 2025. *Tied Up in Tehran: Women, Social Change, and the Politics of Daily Life in Postrevolutionary Iran.* Cambridge: Cambridge University Press.

Neundorf, Anja and Grigore Pop-Eleches. 2020. Dictators and Their Subjects: Authoritarian Attitudinal Effects and Legacies. *Comparative Political Studies* 53(12), 1839–1860.

Neundorf, Anja, Aykut Ozturk, Ksenia Northmore-Ball, Katerina Tertytchnaya, and Johannes Gerschewski. 2024. A Loyal Base: Support for Authoritarian Regimes in Times of Crisis. *Comparative Political Studies* .

Nicholson, Stephen P. and Haifeng Huang. 2022. Making the List: Reevaluating Political Trust and Social Desirability in China. *American Political Science Review* 117(3), 1158–1165.

Norris, Pippa. 2022. *In Praise of Skepticism: Trust but Verify.* Oxford: Oxford University Press.

Ong, Elvin. 2021. Online Repression and Self-Censorship: Evidence from Southeast Asia. *Government and Opposition* 56(1), 141–162.

Peisakhin, Leonid, Arturas Rozenas, and Sergey Sanovich. 2020. Mobilizing Opposition Voters Under Electoral Authoritarianism: A Field Experiment in Russia. *Research & Politics* 7(4), 2053168020970746.

Pop-Eleches, Grigore and Lucan A. Way. 2023. Censorship and the Impact of Repression on Dissent. *American Journal of Political Science* 67(2), 456–471.

Reuter, Ora John and David Szakonyi. 2015. Online Social Media and Political Awareness in Authoritarian Regimes. *British Journal of Political Science* 45(1), 29–51.

Reuter, Ora John and David Szakonyi. 2021. Electoral Manipulation and Regime Support: Survey Evidence from Russia. *World Politics* 73(2), 275–314.

Riambau, Guillem and Kai Ostwald. 2021. Placebo Statements in List Experiments: Evidence from a Face-to-Face Survey in Singapore. *Political Science Research and Methods* 9(1), 172–179.

Robertson, Graeme. 2017. Political Orientation, Information and Perceptions of Election Fraud: Evidence from Russia. *British Journal of Political Science* 47(3), 589–608.

Robinson, Darrel and Marcus Tannenberg. 2019. Self-Censorship of Regime Support in Authoritarian States: Evidence from List Experiments in China. *Research & Politics* 6(3), 1–9.

Rommel, Tobias. 2024. Foreign Direct Investment and Political Preferences in Non-Democratic Regimes. *Comparative Political Studies* 57(8), 1276–1309.

Rosenfeld, Bryn. 2017. Reevaluating the Middle-Class Protest Paradigm: A Case-Control Study of Democratic Protest Coalitions in Russia. *American Political Science Review* 111(4), 637–652.

Seo, TaeJun and Yusaku Horiuchi. 2024. Natural Experiments of the Rally 'Round the Flag Effects Using Worldwide Surveys. *Journal of Conflict Resolution* 68(2-3), 269–293.

Shen, Xiaoxiao and Rory Truex. 2021. In Search of Self-Censorship. *British Journal of Political Science* 51(4), 1672–1684.

Shi, Tianjian. 2001. Cultural Values and Political Trust: A Comparison of the People's Republic of China and Taiwan. *Comparative Politics* 33(4), 401–419.

Sirotkina, Elena and Margarita Zavadskaya. 2020. When the Party's Over: Political Blame Attribution Under an Electoral Authoritarian Regime. *Post-Soviet Affairs* 36(1), 37–60.

Spierings, Niels. 2014. The Influence of Islamic Orientations on Democratic Support and Tolerance in Five Arab Countries. *Politics and Religion* 7(4), 706–733.

Tang, Min and Narisong Huhe. 2020. Parsing the Effect of the Internet on Regime Support in China. *Government and Opposition* 55(1), 130–146.

Tang, Wenfang. 2016. *Populist Authoritarianism: Chinese Political Culture and Regime Sustainability*. Oxford: Oxford University Press.

Tannenberg, Marcus. 2017. *The Autocratic Trust Bias: Politically Sensitive Survey Items and Self-Censorship*.

Tannenberg, Marcus. 2022. The Autocratic Bias: Self-Censorship of Regime Support. *Democratization* 29(4), 591–610.

Tertytchnaya, Katerina. 2023. "This Rally Is Not Authorized": Preventive Repression and Public Opinion in Electoral Autocracies. *World Politics* 75(3), 482–522.

Tezcür, Günes Murat, Taghi Azadarmaki, Mehri Bahar, and Hooshang Nayebi. 2012. Support for Democracy in Iran. *Political Research Quarterly* 65(2), 235–247.

Toklo, Sewordor, Precious Wedaga Allor, and Ruby Amanda Oboro-Offerie. 2025. Democracy Under Threat? The Influence of Perceived Insecurity on Public Support for Democratic Governance in Africa. *Democratization* .

Wang, Yuhua. 2018. *Are College Graduates Agents of Change? Education and Political Participation in China*.

Weber, Max. 1978. *Economy and Society: An Outline of Interpretive Sociology*. Berkeley: University of California Press.

Wen, Cheng, Qian Hu, and Sheng Chen. 2025. What Kinds of Government Trust Structures Affect Political Participation? Evidence from Chinese Youth Netizens. *PLOS ONE* 20(5), e0323981.

Windecker, Paula, Ioannis Vergioglou, and Marc S. Jacob. 2025. Living in Different Worlds: Electoral Authoritarianism and Partisan Gaps in Perceptions of Electoral Integrity. *British Journal of Political Science* 55, e44.

Wintrobe, Ronald. 1998. *The Political Economy of Dictatorship*. Cambridge: Cambridge University Press.

Wu, Chung-li and Alex Min-Wei Lin. 2024. Ambivalence and Perceptions of China: Two List Experiments. *Research & Politics* 11(1), 1–10.

Xu, Xu, Genia Kostka, and Xun Cao. 2022. Information Control and Public Support for Social Credit Systems in China. *The Journal of Politics* 84(4), 2230–2245.

Xu, Yiqing and Jiannan Zhao. 2023. The Power of History: How a Victimization Narrative Shapes National Identity and Public Opinion in China. *Research & Politics* 10(2).

Yeung, Eddy S. F. 2023. Overestimation of the Level of Democracy Among Citizens in Nondemocracies. *Comparative Political Studies* 56(2), 228–266.

Yuen, Samson and Francis L. F. Lee. 2025. Echoes of Silence: How Social Influence Fosters Self-Censorship Under Democratic Backsliding. *Democratization* 32(5), 1213–1238.

Zeira, Yael. 2019. From the Schools to the Streets: Education and Anti-Regime Resistance in the West Bank. *Comparative Political Studies* 52(8), 1131–1168.

Zhu, Jiangnan, Steve Bai, Siqin Kang, Juan Wang, and Kaixiao Liu. 2025. Authoritarian Cue Effect of State Repression. *American Journal of Political Science* 69(4), 1420–1434.

Supplementary information for: "Who Overreports? Regime Support and Preference Falsification in Iran"

Daniel L. Tavana* Penn State
Kevan Harris† UCLA
Gary Fong‡ Penn State
Amir Farmanesh§ IranPoll

Contents

*Assistant Professor, Department of Political Science, Pennsylvania State University. Email: tavana@psu.edu.
†Associate Professor, Department of Sociology, University of California, Los Angeles. Email: kevanharris@soc.ucla.edu.
‡Ph.D. Candidate, Department of Political Science, Pennsylvania State University. Email: cjf5914@psu.edu.
§CEO, IranPoll. Email: farmanesh@iranpoll.com.

A Survey details

The data come from a telephone survey (CATI) of residents of Iran conducted by IranPoll, a subsidiary of People Analytics, Inc. Data collection began on February 20, 2025, and continued until March 18, 2025. A total of 1,999 respondents completed the survey. All interviews were monitored in real-time by call-center supervisors.

The probability sample is nationally representative of the adult population aged 18 and above. Sampling proceeded in six stages. In the first stage, all 31 Iranian provinces were included with certainty, with interviews allocated proportionally to each province's share of the national population. In the second stage, each province was stratified by settlement type into five tiers: cities with populations of 1 million or more (Tier 1), cities with populations of 500,000 to 1 million (Tier 2), cities with populations of 100,000 to 500,000 (Tier 3), cities with populations below 100,000 (Tier 4), and rural locations (Tier 5), based on Iran's 2016 National Population and Housing Census. Interviews were allocated proportionally to the population residing in each settlement tier within each province. In the third stage, at least one primary sampling unit (PSU) was randomly selected from each settlement tier within each province, with no more than 20 interviews conducted in any single PSU where possible. In the fourth stage, landline telephone exchanges within each PSU were randomly selected as secondary sampling units, with no more than 10 interviews conducted per exchange. In the fifth stage, random digit dialing (RDD) was used within each selected exchange to reach random households. In the sixth and final stage, an informed household member was asked to provide the number of eligible adults residing in the household, and a single respondent was randomly selected using randomization software. If no one was home, interviewers attempted contact on two additional occasions (three attempts in total). If the selected respondent was unavailable, a mutually agreeable callback time was arranged. No substitution of selected respondents within households was permitted. Persons ineligible for interview included interviewers' relatives or acquaintances, those residing in hostels, camps, or similar temporary accommodations, and patients in hospitals, sanatoriums, or prisons.

We calculate outcome rates following the standards of the American Association for Public Opinion Research. The response rate (RR2) is the number of completed and partial interviews divided by the total number of interviews, eligible non-interviews, and cases of unknown eligibility, treating every case of unknown eligibility as eligible. The cooperation rate (COOP2) is the number of completed and partial interviews divided by the number of interviews plus eligible non-interviews involving contact (refusals, break-offs, and other non-interviews). The contact rate (CON2) is the proportion of all cases in which a responsible household member was reached, with cases of unknown eligibility weighted by an estimated eligibility proportion (e). The AAPOR2 variants of the response and cooperation rates count partial interviews as respondents in the numerator. The corresponding rates for our survey are as follows:

- the AAPOR2 contact rate was 86.5%
- the AAPOR2 cooperation rate was 74.1%
- the overall AAPOR2 response rate was 60.6%

A.1 Pre-registration

OSF project: https://osf.io/5j3gn
OSF preregistration: https://osf.io/yj6e7

A.2 Ethics

Ethical approval for this research was granted by Institutional Review Boards at Pennsylvania State University (Office of Research Protections reference code STUDY00025559) and the University of California, Los Angeles (Human Research Protection Program reference code IRB-24-5517).

We used the following recruitment and consent scripts:

متن انتخاب خانوار

سلام، اسم من _____ است و برای ایران پل کار می کنم. در حال حاضر در حال انجام یک پیمایش علمی در مورد مسائل مهم ملی و بین المللی هستیم. خانواده شما با استفاده از روش های علمی به صورت تصادفی برای شرکت در مطالعه ما انتخاب شده است. شرکت در این نظرسنجی کاملاً داوطلبانه است و به ما کمک می کند تا باورها و نگرش های مردم خود را در مورد مسائل مهم ملی و بین المللی بهتر درک کنیم. تمام پاسخ های این نظرسنجی کاملاً محرمانه خواهد بود و هرگز به گونه ای استفاده نمی شود که شما یا خانواده تان شناسایی شوید. هیچ اطلاعات شخصی جمع آوری نخواهد شد. شما می توانید در هر زمانی درخواست کنید که این مصاحبه پایان یابد یا کنار گذاشته شود.

[اگر پاسخ دهنده بپرسد در صورت داشتن هرگونه سوال با چه کسی تماس بگیرد:]

در صورت داشتن هرگونه سوال در مورد نظرسنجی، می توانید با آدرس contact@Iranpoll.com با ایران پل تماس بگیرید.

ممکن است با این سوال شروع کنیم: چند نفر بالای هجده سال با شما در این خانه زندگی می کنند؟

[اگر پذیرفته نشد: مصاحبه را خاتمه دهید و این فرد را از پژوهش حذف کنید. اگر گفته شد بله، بپرسید:]

میشه لطفا سن تقریبی اعضای خانواده خود را بگویید؟

[به طور تصادفی یکی از سنین ارائه شده را انتخاب کنید. در صورت عدم پذیرش: مصاحبه را خاتمه دهید و این فرد را از پژوهش حذف کنید.]

Household recruitment script

Hello, my name is _____ and I work for IranPoll. We are currently conducting a scientific survey about important national and international issues. Your household has been selected randomly using scientific methods to participate in our study. Participation in this survey is completely voluntary and will help us better understand the beliefs and attitudes of our people on important national and international issues. All responses to the survey will be kept completely confidential and never be used in a way that would identify you or your household. No personal information will be collected. You may ask for this interview to be ended or discarded at any point.

[If respondent asks who to contact if they have any questions:]

If you have any questions about the survey, you may contact IranPoll at contact@Iranpoll.com.

May we begin by this question: how many individuals over eighteen years of age reside in this house with you?

[If No/Refused: Terminate interview and remove from study. If Yes:]

Can you please give me their approximate ages?

[Randomly select one of the provided ages. If No/Refused: Terminate interview and remove from study.]

[Insert person selected] has/have been selected to participate in our survey from your household. Is [that person] available to do the survey now?

[If Yes: Continue with consent script. If No: Reschedule visit when selected respondent available.]

Participant consent script

Hello, my name is _____ and I work for IranPoll. We are currently conducting a scientific survey about important national and international issues.

The survey takes about 30 minutes and you may opt out of the survey at any time and request to be no longer contacted. Are you ready to begin the survey now?

[If No: Reschedule call back time. If Refuse or opt-out: Terminate interview and remove from study. If Yes:]

All responses to the survey will be kept completely confidential and never be used in a way that would identify you or your household. No personal information will be collected. You may ask for this interview to be ended or discarded at any point.

There will be no direct benefit to you for your participation in this study. However, the information obtained from this study may contribute to our understanding of Iranian society.

May we proceed with the interview?

[If Refuse/Opt-out: Terminate interview and remove from study. If Yes: Administer survey.]

[If respondent asks who to contact if they have any questions:]

[انتخاب شده] برای شرکت در نظرسنجی ما از خانواده شما انتخاب شده است. آیا [آن شخص] اکنون برای انجام نظرسنجی حضور دارد؟

[اگر بله: با متن رضایت نامه ادامه دهید. اگر خیر: برنامه ریزی کنید تا در زمانی که پاسخ دهنده انتخاب شده در دسترس بود، مجددا تماس بگیرید.]

متن رضایت نامه پاسخ دهنده

سلام، اسم من _____ است و در ایران پل کار می کنم. در حال حاضر در حال انجام یک بررسی علمی در مورد مسائل مهم ملی و بین المللی هستیم.

این نظرسنجی حدود ۳۰ دقیقه طول می کشد و شما می توانید در هر زمان از نظرسنجی انصراف دهید و درخواست کنید که دیگر با شما تماس گرفته نشود. آیا اکنون آماده شروع نظرسنجی هستید؟

[اگر نه: زمان تماس مجدد را برنامه ریزی کنید. در صورت امتناع یا انصراف: مصاحبه را خاتمه دهید و این فرد را از پژوهش حذف کنید. اگر بله:]

تمام پاسخ های این نظرسنجی کاملاً محرمانه خواهد بود و هرگز به گونه ای استفاده نمی شود که شما یا خانواده تان شناسایی شوید. هیچ اطلاعات شخصی جمع آوری نخواهد شد. شما می توانید در هر زمانی درخواست کنید که این مصاحبه پایان یابد یا کنار گذاشته شود.

شرکت شما در این مطالعه هیچ سود مستقیمی برای شما نخواهد داشت. با این حال، اطلاعات به دست آمده از این مطالعه ممکن است به درک ما از جامعه ایرانی کمک کند.

آیا می توانیم مصاحبه را شروع کنیم؟

[در صورت امتناع یا انصراف: مصاحبه را خاتمه دهید و این فرد را از پژوهش حذف کنید. اگر بله: نظرسنجی را انجام دهید.]

[اگر پاسخ دهنده بپرسد که در صورت داشتن هر گونه سوال با چه کسی تماس بگیرد:]

اگر در مورد نظرسنجی سوالی دارید، می توانید با آدرس ئنتئتاایرنپللئم با ایران پل تماس بگیرید.

If you have any questions about the survey, you may contact IranPoll at contact@Iranpoll.com.

Table A.1 compares the demographic composition of our sample (N = 1,999) to the 2016 National Population and Housing Census conducted by the Statistical Center of Iran, across five characteristics: gender, age cohort, educational attainment, settlement type, and employment status, as well as the distribution of respondents across all 31 Iranian provinces. Age cohorts are also compared to 2025 population projections from the United Nations Population Division's World Population Prospects data.

Table A.1: Sample demographics

		N	Percent	Census (2016)	WPP (2025)
Gender	Women	1024	51.2	49.9	
	Men	975	48.8	50.1	
	Missing	0	0.0	0.0	
Age	18-29	408	20.4	28.7	21.2
	30-39	390	19.5	27.3	24.3
	40-49	492	24.6	18.2	22.6
	50-59	319	16.0	12.8	14.8
	60+	390	19.5	13.1	17.1
	Missing	0	0.0	0.0	0
Education	Less than secondary	613	30.7	53.5	
	Secondary	606	30.3	22.9	
	BA and higher	778	38.9	22.8	
	Missing	2	0.1	0.8	
Settlement	Urban	1494	74.7	75.2	
	Rural	505	25.3	24.8	
	Missing	0	0.0	0.0	
Status	Employed	813	40.7	39.2	
	Unemployed	272	13.6	5.6	
	Student	113	5.7	7.0	
	Homemaker	564	28.2	35.7	
	Other	234	11.7	12.4	
	Missing	3	0.2	0.1	
Province	Alborz	66	3.3	3.5	
	Ardabil	30	1.5	1.6	
	Bushehr	26	1.3	1.3	
	Chaharmahal and Bakhtiari	23	1.2	1.1	
	East Azerbaijan	108	5.4	5.0	
	Isfahan	128	6.4	6.7	
	Fars	121	6.1	6.2	
	Gilan	56	2.8	3.5	
	Golestan	48	2.4	2.3	
	Hamedan	45	2.3	2.2	
	Hormozgan	45	2.3	2.0	
	Ilam	14	0.7	0.7	
	Kerman	72	3.6	3.8	
	Kermanshah	48	2.4	2.5	
	Khuzestan	119	6.0	5.6	
	Kohgiluyeh and Boyer-Ahmad	15	0.8	0.9	
	Kurdistan	39	2.0	2.0	
	Lorestan	45	2.3	2.2	
	Markazi	35	1.8	1.8	
	Mazandaran	83	4.2	4.4	
	North Khorasan	20	1.0	1.0	
	Qazvin	31	1.6	1.6	
	Qom	31	1.6	1.6	
	Razavi Khorasan	153	7.7	7.7	
	Semnan	20	1.0	0.9	
	Sistan and Baluchestan	75	3.8	2.7	
	South Khorasan	19	1.0	0.9	
	Tehran	339	17.0	17.7	
	West Azerbaijan	87	4.4	4.0	
	Yazd	31	1.6	1.3	
	Zanjan	27	1.4	1.3	
	Missing	0	0.0	0.0	

Note: Study sample comparison with the 2016 National Population and Housing Census conducted by the Statistical Center of Iran. We also compare age cohorts to 2025 age projections from the United Nations Population Division's World Population Prospects data.

Table A.2 presents summary statistics for each variable used in the analysis.

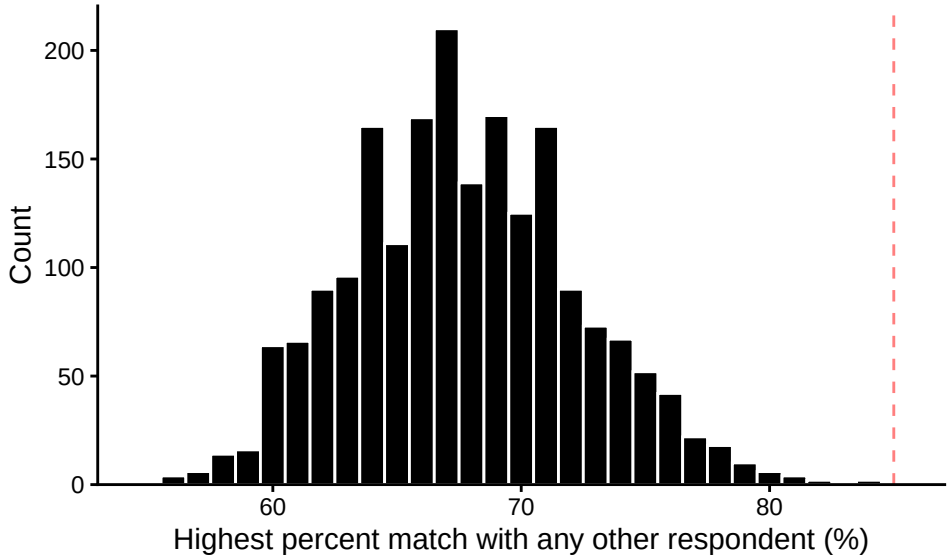
Table A.2: Summary statistics

	Min	Max	Mean	SD	N	Missing
Dependent variable						
Regime support	1	4	2.48	0.94	1973	26
Demographic controls						
Female	0	1	0.51	0.50	1999	0
Age	1	5	2.95	1.40	1999	0
Rural	0	1	0.25	0.43	1999	0
Minority	0	1	0.53	0.50	1998	1
Expressive identification						
National pride	1	4	3.68	0.69	1988	11
Religiosity	0	1	0.61	0.49	1978	21
Behavioral conformity						
Fear	1	3	1.75	0.85	1894	105
Fear (1)	1	4	2.55	1.14	1854	145
Fear (2)	1	4	2.36	1.06	1875	124
Self-censorship	1	3	1.83	0.73	1989	10
Self-censorship (1)	1	4	2.20	0.94	1965	34
Self-censorship (2)	1	4	1.95	0.95	1981	18
Taarof	1	3	1.98	0.81	1973	26
Civic resources						
Education	1	3	2.08	0.83	1997	2
Income	1	4	2.26	0.90	1952	47
Political awareness	1	3	1.86	0.79	1999	0
Political interest	0	1	0.44	0.36	1992	7
Political knowledge	0	1	0.40	0.49	1999	0
Political news	0	1	0.64	0.30	1998	1
Efficacy	1	3	1.88	0.79	1970	29
Internal efficacy (1)	1	4	2.24	1.11	1923	76
Internal efficacy (2)	1	4	2.55	0.98	1936	63
External efficacy (1)	1	4	2.43	1.07	1929	70
External efficacy (2)	1	4	2.26	1.11	1941	58
Additional variables						
Employed	0	1	0.41	0.49	1996	3
Unemployed	0	1	0.14	0.34	1996	3
Student	0	1	0.06	0.23	1996	3
Homemaker	0	1	0.28	0.45	1996	3
Public employee	0	1	0.11	0.31	1996	3

Note: Summary statistics for each variables (and their component parts) used in the analysis.

To assess the reliability of our data, following Kuriakose & Robbins (2016), we assess the prevalence of near-duplicate observations by computing, for each respondent, the highest percentage of matching values with any other respondent across all non-constant survey variables. Near-duplicate observations—defined as respondent pairs with 85% or more matching values—can indicate data fabrication by enumerators who copy or minimally alter previous responses. Figure A.1 displays the distribution of highest percent matches across all 1,999 respondents. No observation in our data reaches the 85% threshold, which suggests the low likelihood of data falsification or systematic duplication.

Figure A.1: Duplicate observation prevalence



Note: Distribution of highest percent match between each respondent and all other respondents across all variables. The dashed red line marks the 85% threshold for suspected duplicates.

B List experiment diagnostics

In this appendix, we report the results of several diagnostic tests that, taken together, increase confidence in our approach and resulting estimates.

We begin by discussing the choice between maximum likelihood (ML) and nonlinear least squares (NLS) estimators for list experiment regression. The choice reflects a tradeoff between statistical efficiency and robustness. Both estimators fit the same underlying model—one component for the control items and one for the sensitive item—but they optimize different objective functions (Blair & Imai 2012). NLS minimizes the sum of squared residuals and requires only that the conditional mean functions are correctly specified, making no distributional assumptions about the full response distribution. ML, in contrast, maximizes the complete observed-data likelihood function by modeling each cell of the response distribution, including the extreme cells where respondents in the treatment group score either 0 or the maximum possible value. This reliance on extreme cells gives ML greater statistical efficiency (tighter standard errors), but at a cost: Ahlquist (2018) demonstrates that ML is sensitive to the number of observations in extreme cells, but well-designed list experiments that follow best practices for minimizing strategic misrepresentation actively reduce the number of respondents in exactly those cells. When few observations fall in the extremes, small amounts of non-strategic measurement error (respondents misreading the question, rushing, or miscounting) can cause ML to produce severely biased estimates, even though NLS is comparatively more stable.

Blair *et al.* (2019) formalize this insight by showing that NLS is robust to non-strategic measurement error precisely because it does not model the extreme cells directly. They propose a Hausman specification test that detects divergence between ML and NLS estimates as a diagnostic for model misspecification. The test statistic is formed from the weighted difference between the ML and NLS coefficient vectors: under correct specification, both estimators are consistent and their difference should be statistically indistinguishable from zero, yielding a positive test statistic and a high p-value. A low p-value signals misspecification and suggests NLS should be preferred. However, Blair *et al.* (2019) note that a negative test statistic - which arises when ML fits the data worse than NLS - is itself evidence of misspecification, independent of the p-value, since ML should by construction never be less efficient than NLS when the model is correctly specified. In practice, when a list experiment is well-implemented and the sensitive trait is sufficiently common, ML, NLS, and ML with floor and ceiling corrections all produce substantively similar results, suggesting that convergence across estimators is itself evidence that distributional assumptions are approximately satisfied and that either estimator can be trusted (Karpowitz *et al.* 2023).

To assess whether the ML or NLS estimator is more appropriate for our data, we perform the Blair *et al.* (2019) Hausman specification test for both list experiments, fitting intercept-only models using the empty regression specification recommended for this diagnostic. For List 1, the test yields a statistic of -1.72 ($p = 0.42$), and for List 2 a statistic of -2.46 ($p = 0.29$). The p values for both lists are well above conventional significance thresholds: we fail to reject the null hypothesis of correct model specification for both lists. However, as Blair *et al.* (2019) note, the sign of the test statistic is itself informative: a negative value arises when ML fits the data worse than NLS, which should not occur under correct specification since ML is by construction at least as efficient as NLS when its distributional assumptions hold. The negative test statistics for both lists are therefore a sign that the distributional assumptions required by ML may not be satisfied in our data. As a result, we prefer NLS as the primary estimator. NLS requires only that the conditional mean functions are correctly specified and is robust to the kinds of non-strategic measurement error and distributional misspecification that can bias ML estimates. We report NLS estimates as our primary results throughout the paper.

Next, to verify that treatment assignment is independent of respondent characteristics, we conduct χ^2 tests of independence between each variable used in the analysis and treatment assignment (Table A.3). Because all respondents received List 1 first and treatment assignment for List 2 is the inverse of List 1 by design, a single balance check is sufficient. Under the null hypothesis of no association, the results confirm successful randomization: across the 28 variables tested, only `political awareness` yields a statistically significant result ($\chi^2 = 9.87$, $df = 2$, $p = 0.007$). Given the large number of tests conducted, a single rejection at this threshold is consistent with chance variation under the null. We proceed under the assumption that randomization was successful. Table A.3 prints the full results.

Table A.3: Randomization checks

Variable	df	χ^2	p-value
Demographic controls			
Female	1	0.55	0.459
Age	4	2.01	0.733
Rural	1	0.20	0.652
Minority	1	0.12	0.724
Expressive identification			
National pride	3	6.01	0.111
Religiosity	1	0.03	0.859
Behavioral conformity			
Fear	2	1.21	0.546
Fear (1)	3	2.04	0.564
Fear (2)	3	1.81	0.613
Self-censorship	2	1.24	0.537
Self-censorship (1)	3	1.54	0.672
Self-censorship (2)	3	7.63	0.054^{+}
Taarof	2	0.76	0.685
Civic resources			
Education	2	2.56	0.278
Income	3	0.12	0.989
Political awareness	2	9.87	0.007^{**}
Political interest	3	5.78	0.123
Political knowledge	1	3.29	0.070^{+}
Political news	3	7.60	0.055^{+}
Efficacy	2	2.16	0.340
Internal efficacy (1)	3	0.59	0.898
Internal efficacy (2)	3	2.34	0.505
External efficacy (1)	3	5.47	0.141
External efficacy (2)	3	3.79	0.286
Additional variables			
Employed	1	0.19	0.661
Unemployed	1	0.00	0.976
Student	1	0.31	0.576
Homemaker	1	0.00	1.000
Public employee	1	0.01	0.918

Note: χ^2 tests of independence (null hypothesis of no association between the variable and treatment assignment).
+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001.

B.1 "No design effects" assumption

We next assess the "no design effects" assumption, which requires that respondents' answers to the control items are not affected by the addition of the sensitive item to the list. Design effects can bias list experiment estimates in two ways. Following Blair & Imai (2012), Table A.4 reports the estimated proportion of each respondent type $\pi(Y_i(0), Z_i)$, defined by the latent count of affirmative answers to the control items and the latent response to the sensitive item. The "no design effects" assumption holds that none of these estimated proportions are negative. Negative values would indicate that the

observed response distributions are inconsistent with the assumption (Blair & Imai 2012). As Table A.4 shows, all estimated proportions are non-negative for both lists, providing no evidence of a design effect. For both List 1 and List 2, the Bonferroni-corrected p-value equals 1.00: we therefore fail to reject the null hypothesis of no design effect for either list.[5] These results are consistent with the assumption that treatment assignment did not alter respondents' answers to the control items.

Table A.4: Estimated proportion of respondent types

	List 1	List 2
$\pi(Y_i(0) = 0,\, Z_i = 1)$	0.058 (0.015)	0.026 (0.010)
$\pi(Y_i(0) = 1,\, Z_i = 1)$	0.179 (0.022)	0.146 (0.022)
$\pi(Y_i(0) = 2,\, Z_i = 1)$	0.172 (0.015)	0.244 (0.017)
$\pi(Y_i(0) = 3,\, Z_i = 1)$	0.047 (0.007)	0.058 (0.008)
$\pi(Y_i(0) = 4,\, Z_i = 1)$	0.000 (0.000)	0.000 (0.000)
$\pi(Y_i(0) = 0,\, Z_i = 0)$	0.096 (0.009)	0.036 (0.006)
$\pi(Y_i(0) = 1,\, Z_i = 0)$	0.246 (0.019)	0.270 (0.017)
$\pi(Y_i(0) = 2,\, Z_i = 0)$	0.195 (0.021)	0.209 (0.022)
$\pi(Y_i(0) = 3,\, Z_i = 0)$	0.007 (0.010)	0.009 (0.011)
$\pi(Y_i(0) = 4,\, Z_i = 0)$	0.000 (0.000)	0.002 (0.001)

Note: Estimated proportion of respondents of each type $\pi(Y_i(0), Z_i)$, where $Y_i(0) \in \{0, \ldots, 4\}$ denotes the latent count of affirmative answers to the control items and $Z_i \in \{0, 1\}$ denotes the latent response to the sensitive item. Standard errors in parentheses.

Design effects may also arise from non-strategic measurement error—not through deliberate concealment, but from the experimental setup itself. One way this might occur is through mechanical inflation: because the treatment group has more items on their list ($J + 1$ versus J), respondents may satisfice, selecting the perceived middle point rather than carefully evaluating each item. As a result, treated respondents will mechanically report more true items regardless of the sensitive item's content. This is a violation of the no design effects assumption not because the sensitive item changes how respondents think about control items, but because the unequal list lengths change how respondents count. It produces the same observational signature as a genuine treatment effect: a higher mean count in the treatment group. This means mechanical inflation biases the difference-in-means estimator upward, making the sensitive trait appear more prevalent than it truly is, independently of whether a respondent is lying strategically. Table A.4 would detect this in our experiment. But there is evidence that mechanical inflation may be concentrated among low education, low income, older, and less knowledgeable respondents (Riambau & Ostwald 2021).

Mechanical inflation introduces a source of bias that cannot be addressed by the existing diagnostic toolkit. But in order to assess the prevalence of mechanical inflation in our experiment, we introduce what Riambau & Ostwald (2021) describe as a "placebo" experiment: a separate list experiment prior to the double list experiment. Respondents were asked the following:

> `Here is a list of four/five items. How many of the following apply to you? I will read the items twice: please listen carefully. You don't need to tell me which ones apply to you, just how many. Please let me know if you need me`

[5] For List 1, the `ict.test` function produced a computational warning at the $j = 4$ boundary because no respondent in the treatment group scored 4 on the control items, rendering the comparison vector constant. This indicates that the boundary condition is trivially satisfied rather than a sign of a problem: a proportion that is exactly zero cannot be negative. The Bonferroni-corrected p-value of 1.00 is unaffected.

to repeat them.

I drink tea everyday
I like exercising
I am married
I go mountain climbing at least once a week
I moved to my current home less than a week ago

The neutral statements were chosen using generally accepted criteria: uncorrelated with each other and resistant to floor and ceiling effects. Treated respondents were presented with a placebo statement designed to be plausible, but likely false for all respondents: "I moved to my current home less than a week ago." Though we do not know the "true" percentage of those who moved in the past week, we assume it is false for all respondents.[6]

Table A.5 presents the results of the placebo experiment. For the whole sample, the mean item count is virtually identical across control and treatment groups (2.40 versus 2.41, $p = 0.41$), providing no evidence of mechanical inflation in our data. This pattern holds broadly across individual subgroups: across all variables and subgroups tested, the overwhelming majority of p-values fall well above conventional significance thresholds. There does not seem to be any systematic pattern of inflation concentrated among specific subgroups. Only one subgroup reaches statistical significance—respondents aged 40–49 ($p = 0.007$)—but given the large number of subgroup comparisons conducted, a single rejection of this kind is consistent with chance variation under the null. The subgroups most theoretically vulnerable to mechanical inflation (i.e., those with lower levels of education, income, and political awareness) show no evidence of inflation in our data. We see this experiment as strong evidence in support of the validity of our list experiment estimates.

[6] In the United States, between 8 and 13 percent of individuals change addresses each year: this would suggest a trivial "true" prevalence rate of weekly movers: between 0.15 and 0.25 percent on average.

Table A.5: Mean item counts by subgroup (Placebo statement experiment)

Variable	Subgroup	Control	Treatment	p-value
Sample		2.40 (1037)	2.41 (962)	0.4062
Demographic controls				
female	Men	2.40 (511)	2.41 (464)	0.4151
	Women	2.40 (526)	2.40 (498)	0.4496
age	18-29	1.88 (199)	1.88 (209)	0.5225
	30-39	2.63 (204)	2.69 (186)	0.2305
	40-49	2.62 (256)	2.78 (236)	0.0071**
	50-59	2.54 (186)	2.54 (133)	0.5084
	60+	2.26 (192)	2.16 (198)	0.9535
rural	Urban	2.38 (783)	2.39 (711)	0.3990
	Rural	2.46 (254)	2.46 (251)	0.5134
minority	Persian	2.35 (488)	2.42 (449)	0.0882+
	Minority	2.44 (549)	2.40 (512)	0.7990
Expressive identification				
pride	Low National Pride	2.43 (28)	2.56 (18)	0.2633
	Medium-low National Pride	2.34 (65)	2.28 (58)	0.6553
	Medium-high National Pride	2.40 (126)	2.39 (117)	0.5148
	High National Pride	2.40 (813)	2.41 (763)	0.3542
religiosity	Low Religiosity	2.30 (404)	2.31 (359)	0.4866
	High Religiosity	2.45 (620)	2.46 (595)	0.4163
Behavioral conformity				
fear	Low Fear	2.36 (505)	2.38 (462)	0.3683
	Medium Fear	2.39 (226)	2.38 (200)	0.5443
	High Fear	2.47 (257)	2.47 (244)	0.4745
self_censorship	Low Self-censorship	2.43 (369)	2.38 (352)	0.8216
	Medium Self-censorship	2.37 (477)	2.43 (407)	0.1050
	High Self-censorship	2.41 (186)	2.41 (198)	0.5233
taarof	Low Taarof	2.33 (350)	2.35 (311)	0.3954
	Medium Taarof	2.40 (353)	2.41 (329)	0.4233
	High Taarof	2.47 (319)	2.46 (311)	0.5869
Civic resources				
education	Less than secondary	2.39 (311)	2.33 (302)	0.8111
	Secondary	2.48 (322)	2.48 (284)	0.4736
	BA and higher	2.34 (403)	2.41 (375)	0.1472
income	0-8 mil	2.33 (206)	2.38 (182)	0.2239
	8-16 mil	2.39 (444)	2.38 (444)	0.5877
	16-24 mil	2.45 (234)	2.47 (223)	0.4277
	24+ mil	2.46 (131)	2.39 (88)	0.7438
political_awareness	Low Political Awareness	2.34 (401)	2.33 (390)	0.5418
	Medium Political Awareness	2.43 (370)	2.45 (328)	0.3313
	High Political Awareness	2.45 (266)	2.46 (244)	0.4089
efficacy	Low Efficacy	2.43 (367)	2.43 (371)	0.4567
	Medium Efficacy	2.39 (391)	2.41 (333)	0.3851
	High Efficacy	2.36 (261)	2.36 (247)	0.4977
Additional variables				
employed	Other (Not employed)	2.34 (606)	2.36 (577)	0.3581
	Employed	2.47 (429)	2.47 (384)	0.4965
unemployed	Other (Not Unemployed)	2.41 (889)	2.42 (835)	0.4568
	Unemployed	2.29 (146)	2.32 (126)	0.4039
student	Other (Not Student)	2.43 (983)	2.45 (900)	0.3159
	Student	1.69 (52)	1.72 (61)	0.4046
homemaker	Other (Not Homemaker)	2.34 (745)	2.33 (687)	0.6122
	Homemaker	2.53 (290)	2.59 (274)	0.1899
public_emp	Other (Not Public Employee)	2.38 (921)	2.38 (863)	0.4738
	Public Employee	2.54 (114)	2.60 (98)	0.2792

Note: Mean count by treatment and control groups (number of observations in parentheses). Reported p-values are from a one-sided difference in means t-test.
+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001.

B.2 "No liars" assumption

A critical identifying assumption of list experiment regression is the "no liars" assumption, which requires that respondents in the treatment group answer the sensitive item truthfully (Blair & Imai 2012). Formally, this means that the observed response to the sensitive item equals the respondent's true latent preference. Two specific violations of this assumption are of particular concern. Ceiling effects arise when "respondents' true preferences are affirmative for all the control items as well as the sensitive item" (Blair & Imai 2012, p. 50): because answering $J + 1$ would perfectly reveal their support for the sensitive item, such respondents may strategically report J instead. Floor effects arise in the mirror situation, where "the control questions are so uncontroversial that uniformly negative responses are expected for many respondents" (Blair & Imai 2012, p. 50): a respondent whose truthful answer is affirmative only for the sensitive item may fear that reporting $Y_i = 1$ would reveal their preference, and so reports 0 instead. Both effects lead to underestimation of the true proportion of affirmative responses to the sensitive item, as liars in both cases lower the observed mean response of the treatment group (Blair & Imai 2012).

The observed response distributions (Table A.6) allow us to assess the likelihood of ceiling and floor effects prior to estimation. Ceiling effects are indicated by the proportion of treatment group respondents scoring $J + 1 = 5$ in both List 1 and List 2: these cells are empty (0.00%), meaning that no respondent affirmed support for all five items. While this eliminates one form of ceiling concern, as we describe above, it also implies that the ML estimator has no direct observations in $\mathcal{J}(1, J + 1)$ from which to identify sensitive item prevalence at the top of the distribution. Floor effects are a more substantive concern. In the List 1 treatment group, 9.52% of respondents (95 individuals) scored 0, a non-trivial proportion that represents the pool of potential floor liars—respondents who may have suppressed a truthful response of 1 to avoid revealing support for the regime. The number is lower for List 2 (3.60%). Because floor liars and true zeros are observationally equivalent, these proportions cannot be decomposed without additional assumptions. These distributional features of our data reinforce our preference for the NLS estimator, which does not rely on extreme-cell observations for identification.

Table A.6: Observed response distributions

	List 1				List 2			
	Control		Treatment		Control		Treatment	
Response value	N	Percent	N	Percent	N	Percent	N	Percent
0	154	15.38	95	9.52	62	6.21	36	3.60
1	425	42.46	303	30.36	414	41.48	296	29.57
2	367	36.66	371	37.17	451	45.19	355	35.46
3	54	5.39	178	17.84	67	6.71	253	25.27
4	0	0.00	47	4.71	2	0.20	60	5.99
5	—	—	0	0.00	—	—	0	0.00
Nonresponse	1	0.10	4	0.40	2	0.20	1	0.10
Total	1001	100.00	998	100.00	998	100.00	1001	100.00

Note: Table displays counts and percentages across control and treatment groups for both list experiments. Percentages may not sum to 100 due to rounding.

How concerned should we be about the impact of floor effects on our estimates? In looking at the response distributions in Table A.6, our distribution is similar to or better than many recent and canonical list experiments in the literature (Blair *et al.* 2014; Corstange 2018; Kuhn & Vivyan 2018; Karpowitz *et al.* 2023; Buckley *et al.* 2024; Wu & Lin 2024). To assess whether floor lying is biasing

our estimates, we fit ML models with floor corrections for each list experiment (Blair & Imai 2012). We begin by comparing NLS and ML estimates, since floor corrections require the ML estimator. For List 1, the NLS estimate (45.7%) and ML estimate (48.8%) diverge by 3.1 percentage points. The divergence is larger for List 2: 47.4% (NLS) versus 54.9% (ML), a gap of 7.5 percentage points. As the Hausman specification test reported above indicates, both test statistics are negative (-1.72 for List 1; -2.46 for List 2), suggesting that ML fits the data worse than NLS, an outcome inconsistent with correct model specification. Since floor corrections are implemented entirely within the ML framework, they inherit this misspecification. The results below should be interpreted with caution - even though the different estimates between our ML and NLS estimates are similar in magnitude to those found in comparable studies (Karpowitz *et al.* 2023; Buckley *et al.* 2024).

First, both List 1 and List 2 floor models produced convergence warnings, indicating that the EM algorithm's monotone convergence property was violated. This is likely due to the flat likelihood surface arising from the near-separability of floor liars and genuine zero-scorers. The estimated conditional probability of floor lying is 0.016 for List 1 and 0.002 for List 2. Within the ML framework, floor-corrected regime support prevalence estimates do not change ML estimates: 47.4% versus 48.8% for List 1, and 54.7% versus 54.9% for List 2. Model fit statistics suggest the same. For List 1, the floor model log-likelihood (-2621.872) exceeds the baseline (-2621.882) by 0.011 units, and the floor model fits worse when comparing each model's Bayesian information criterion (change in BIC = 7.58 in favor of the baseline). For List 2, the floor model log-likelihood of -2569.133 is actually lower than the baseline of -2568.741 by 0.391 units—an outcome that is unlikely under a properly converging EM algorithm and constitutes additional evidence of convergence failure. The BIC comparison (ΔBIC = 8.39) again favors the baseline. Under standard BIC conventions, a ΔBIC exceeding 6 constitutes strong evidence against a more complex model. Taken together, these results suggest that boundary effects are unlikely to be a meaningful source of bias in the primary NLS estimates.

B.3 Benchmarking the list experiment

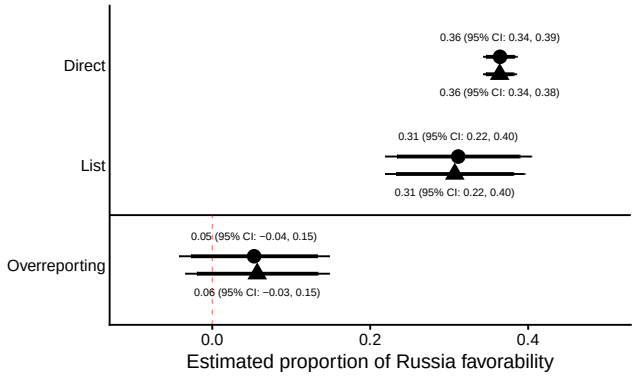
Is opposition to the regime politically sensitive? To assess the validity of our overreporting estimate and subgroup differences, we introduced an additional list experimental item after the double list experiment. Respondents were asked the following:

> `Here is a list of four/five countries. How many of the following do you have a favorable view of? I will read the countries twice: please listen carefully. You don't need to tell me which ones you have a favorable view of, just how many. Please let me know if you need me to repeat them.`
>
> `Japan`
> `China`
> `France`
> `United Kingdom`
> `Russia`

The neutral statements were again chosen using generally accepted criteria: uncorrelated with each other and resistant to floor and ceiling effects. We see this test as a benchmark validity check. Because favorability toward Russia is not a politically sensitive topic in Iran, we expect no difference between the list and direct estimates: any observed gap would suggest that the list experiment format itself introduces bias. The logic mirrors the placebo statement approach (Riambau & Ostwald 2021) but targets a substantive attitude rather than a near zero-prevalence behavior, making it a more

demanding test. If the list experiment inflates estimates even for a non-sensitive political attitude, the overreporting gap observed in the regime support experiment could also reflect inflation rather than preference falsification.

Figure A.2: Russia favorability and overreporting: List and direct estimates and differences

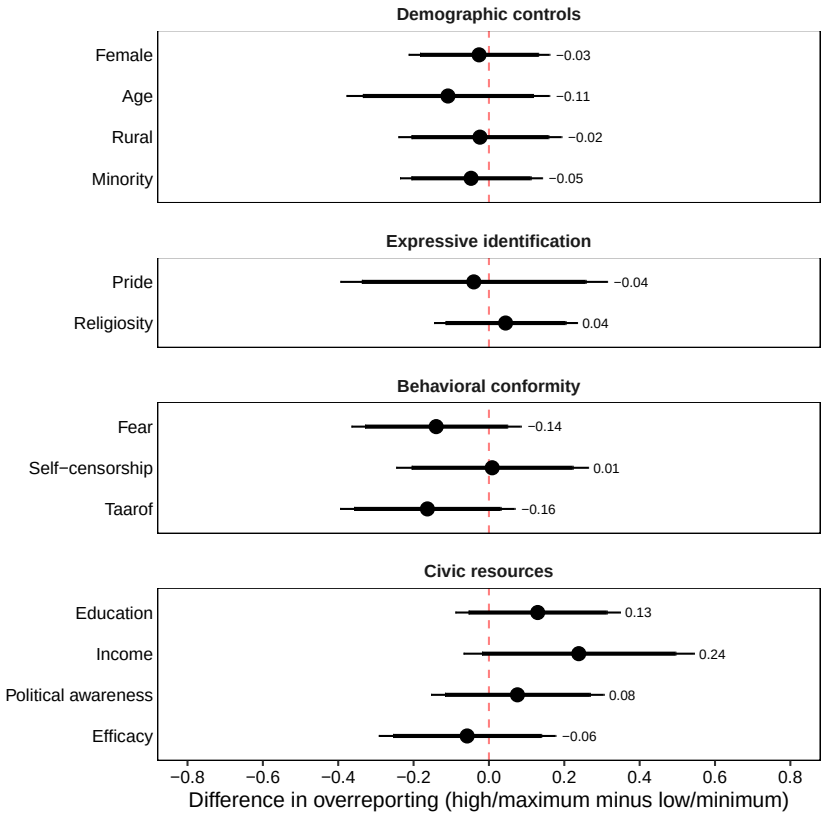


Note: Estimated proportion of respondents with a favorable view of Russia using the direct question and a single list experiment, and the estimated proportion of overreporters (difference between direct and list estimates). Point estimates are shown with 90% (thick) and 95% (thin) confidence intervals. Unadjusted models include no covariates; adjusted models control for gender, age, settlement, and minority status. $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Figure A.2 presents the main results from this experiment. The unadjusted list estimate of Russia favorability is 0.31 (95% CI: 0.22, 0.40) and the unadjusted overreporting gap is 0.05 (95% CI: -0.04, 0.15; p = 0.27). The covariate-adjusted estimates are substantively identical (estimate: 0.31, 95% CI: 0.22, 0.40; gap: 0.06, 95% CI: -0.03, 0.15; p = 0.22). Neither gap is statistically distinguishable from zero, consistent with the expectation that non-sensitive items should not generate preference falsification. This null result provides evidence that the overreporting we observe for regime support reflects genuine preference falsification rather than an artifact of the list experiment format.

Figure A.3 replicates the analysis presented in Figure 6 in the main text. As we describe above, our confidence intervals for these estimates are large (we did not implement a double list experiment for this question). But across all predictors, there is no evidence of overreporting among any of the theoretical subgroups of interest we analyze in the main paper.

Figure A.3: Subgroup differences and overreporting (Russia favorability)



Note: Difference in overreporting of Russia favorability between respondents at the maximum and minimum values of each variable. Overreporting for each subgroup is defined as the direct question predicted probability minus the list experiment predicted probability; the plotted quantity is the difference-of-differences across subgroups. Positive values indicate that respondents at the high end of a variable overreport more than those at the low end. Point estimates are shown with 90% (thick) and 95% (thin) confidence intervals. Predicted probabilities are computed via parametric simulation (10,000 draws) using model-based variance-covariance matrices. $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

C Additional analyses

Tables A.7 and A.8 report the NLS coefficient estimates for the sensitive and control item equations underlying the regime support estimates presented in Figure 3. Table A.7 reports coefficients for List 1 and List 2 separately; Table A.8 reports coefficients for the pooled double list experiment.

Table A.7: Coefficient estimates (List 1 and List 2)

	Sensitive items		Control items	
Variables	Est.	SE	Est.	SE
List 1, unadjusted				
`Intercept`	-0.174	0.164	-0.707	0.028
List 1, adjusted				
`Intercept`	-0.890	0.540	-0.844	0.089
`Female`	0.207	0.344	-0.089	0.058
`Age`	0.048	0.127	0.056	0.021
`Rural`	0.476	0.373	0.078	0.064
`Minority`	0.646	0.343	-0.010	0.058
List 2, unadjusted				
`Intercept`	-0.105	0.153	-0.478	0.024
List 2, adjusted				
`Intercept`	-0.312	0.477	-0.556	0.074
`Female`	0.127	0.315	-0.067	0.049
`Age`	-0.025	0.112	0.034	0.017
`Rural`	0.550	0.365	0.076	0.053
`Minority`	0.134	0.314	-0.012	0.049

Note: Estimates from nonlinear least squares item count technique regression (`ictreg`, `method = "nls"`). $J = 4$ control items.

Table A.8: Coefficient estimates (Double list experiment)

	Sensitive items		Control items	
Variables	Est.	SE	Est.	SE
DLE, unadjusted				
`Intercept`	-0.137	0.099	-0.591	0.018
DLE, adjusted				
`Intercept`	-0.584	0.311	-0.699	0.058
`Female`	0.149	0.197	-0.074	0.036
`Age`	0.010	0.075	0.045	0.014
`Rural`	0.474	0.221	0.081	0.041
`Minority`	0.406	0.199	-0.014	0.038

Note: Estimates from nonlinear least squares item count technique regression (`ictreg`, `method = "nls"`). $J = 4$ control items. Standard errors from 1,000 bootstrap iterations clustered by respondent.

Tables A.9 and A.10 report the estimates presented in Figures 5 and 6. Table A.9 reports subgroup differences in regime support from the list experiment and the direct question; Table A.10 reports subgroup differences in overreporting.

Table A.9: Subgroup differences in regime support (Figure 5)

Variable	List (SE)	Direct (SE)
Demographic controls		
Female	0.025 (0.047)	0.111*** (0.022)
Age	-0.014 (0.070)	0.164*** (0.031)
Rural	0.127* (0.055)	0.177*** (0.024)
Minority	0.110* (0.048)	0.040+ (0.022)
Expressive identification		
Religiosity	0.222*** (0.048)	0.418*** (0.021)
Pride	0.338** (0.104)	0.589*** (0.021)
Behavioral conformity		
Fear	-0.311*** (0.053)	-0.654*** (0.020)
Self-censorship	0.008 (0.064)	0.163*** (0.030)
Taarof	0.073 (0.060)	0.082** (0.028)
Civic resources		
Education	-0.110+ (0.058)	-0.258*** (0.026)
Income	-0.152* (0.076)	-0.257*** (0.036)
Political awareness	-0.097 (0.062)	-0.129*** (0.028)
Efficacy	0.338*** (0.062)	0.460*** (0.025)

Note: Difference in predicted probability of regime support between respondents at the maximum and minimum values of each variable. Standard errors in parentheses.
+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001.

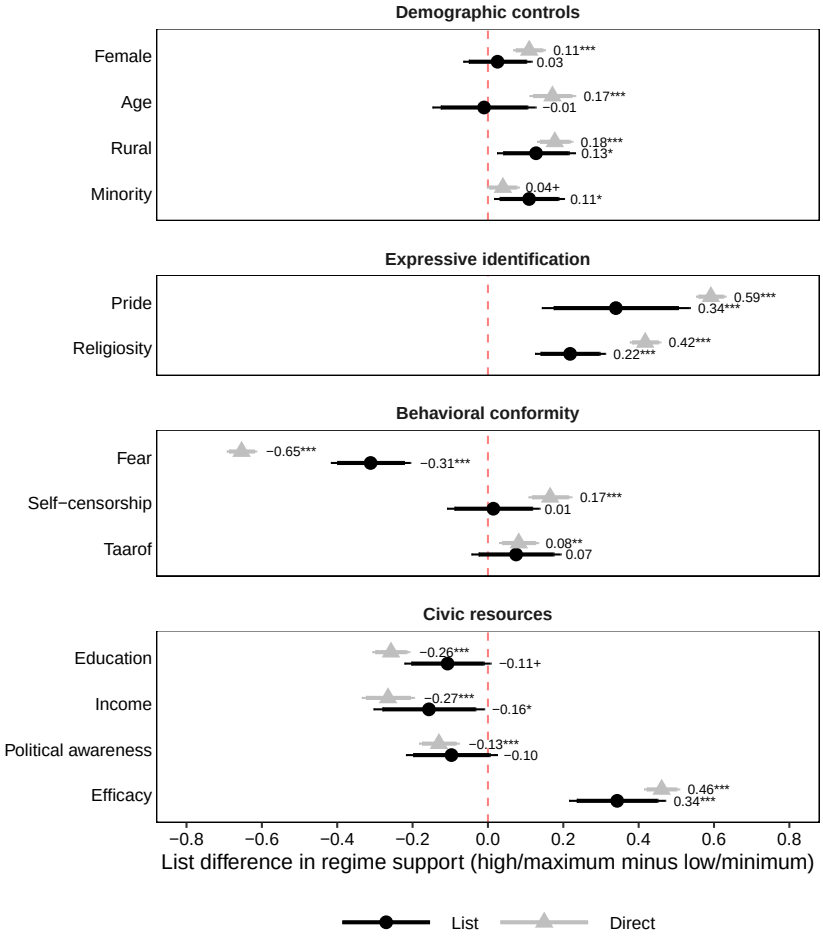
Table A.10: Subgroup differences in overreporting (Figure 6)

Variable	Overreporting (SE)
Demographic controls	
Female	0.085+ (0.052)
Age	0.178* (0.077)
Rural	0.050 (0.060)
Minority	-0.070 (0.053)
Expressive identification	
Religiosity	0.196*** (0.052)
Pride	0.251* (0.106)
Behavioral conformity	
Fear	-0.343*** (0.057)
Self-censorship	0.154* (0.071)
Taarof	0.010 (0.067)
Civic resources	
Education	-0.148* (0.063)
Income	-0.105 (0.084)
Political awareness	-0.033 (0.068)
Efficacy	0.122+ (0.067)

Note: Difference in overreporting between respondents at the maximum and minimum values of each variable. Standard errors in parentheses.
+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001.

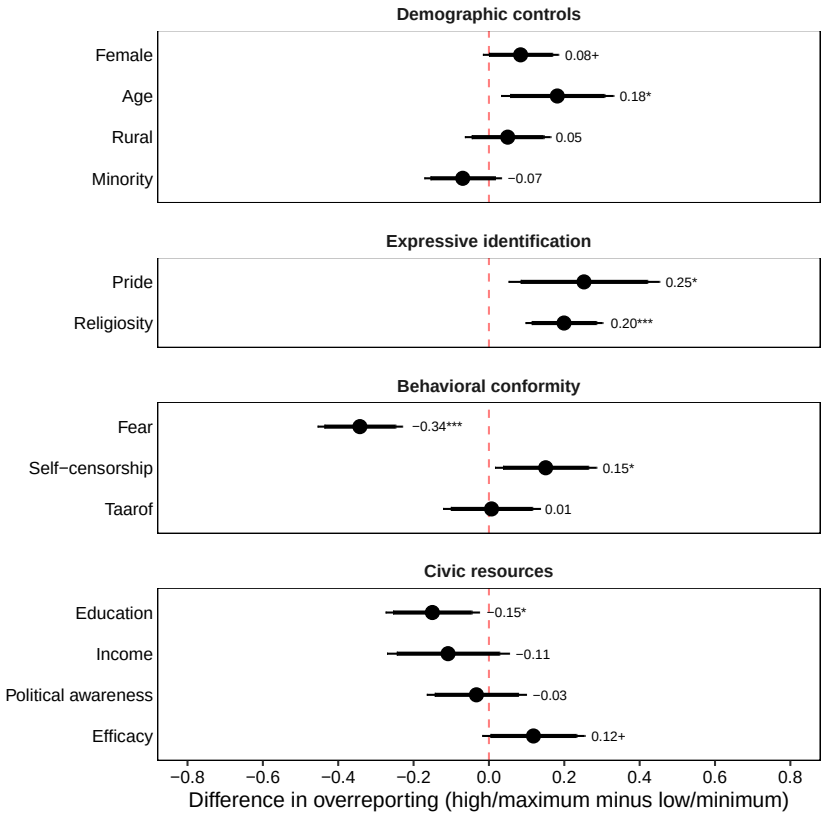
Figures A.4 and A.5 replicate the analyses presented in Figures 5 and 6 in the main text, with controls.

Figure A.4: Subgroup differences and regime support (adjusted)



Note: Difference in predicted probability of regime support between respondents at the maximum and minimum values of each variable, estimated from covariate-adjusted NLS regressions (list) and logistic regressions (direct). Point estimates are shown with 90% (thick) and 95% (thin) confidence intervals. Predicted probabilities are computed via parametric simulation (10,000 draws); list model variance-covariance matrices are obtained from 1,000 bootstrap iterations clustered by respondent. Estimation follows the same procedure as Figure 5. $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Figure A.5: Subgroup differences and overreporting (adjusted)



Note: Difference in overreporting of regime support between respondents at the maximum and minimum values of each variable. Overreporting for each subgroup is defined as the covariate-adjusted direct question predicted probability minus the covariate-adjusted list experiment predicted probability; the plotted quantity is the difference-of-differences across subgroups. Positive values indicate that respondents at the high end of a variable overreport more than those at the low end. Point estimates are shown with 90% (thick) and 95% (thin) confidence intervals. Estimation follows the same procedure as Figure 6. $^{+}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

Table A.11 reports Benjamini & Hochberg (1995) adjusted p-values for the nine hypothesized variables.

Table A.11: Benjamini-Hochberg adjusted p-values

	List		Overreporting	
Variable	Estimate (SE)	p	Estimate (SE)	p
Expressive identification				
`Religiosity`	0.222 (0.048)	0.0000^{***}	0.196 (0.052)	0.0008^{***}
`Pride`	0.338 (0.104)	0.0026^{**}	0.251 (0.106)	0.0432^{*}
Behavioral conformity				
`Fear`	-0.311 (0.053)	0.0000^{***}	-0.343 (0.057)	0.0000^{***}
`Self-censorship`	0.008 (0.064)	0.8957	0.154 (0.071)	0.0526^{+}
`Taarof`	0.073 (0.060)	0.2581	0.010 (0.067)	0.8830
Civic resources				
`Education`	-0.110 (0.058)	0.0854^{+}	-0.148 (0.063)	0.0432^{*}
`Income`	-0.152 (0.076)	0.0795^{+}	-0.105 (0.084)	0.2714
`Political awareness`	-0.097 (0.062)	0.1494	-0.033 (0.068)	0.7081
`Efficacy`	0.338 (0.062)	0.0000^{***}	0.122 (0.067)	0.1022

Note: Benjamini-Hochberg adjusted p-values for the nine hypothesized predictor variables. Estimates are differences in predicted probability between respondents at the maximum and minimum values of each variable (standard errors in parentheses). Demographic controls are excluded from the correction.
+ p<0.1; * p<0.05; ** p<0.01; *** p<0.001.