

Clash of Ideologies:

How Citizens Respond to U.S. and China's Rhetoric About Domestic Politics in Southeast Asia*

Nico Ravanilla[†] & Damian Boldt[‡]

August 25, 2026

Abstract

Southeast Asia is an arena of U.S.–China competition for influence. How each great power addresses the region's domestic politics is central to that competition. Conventional wisdom holds that American criticism of illiberal policies should be punished, while Chinese praise should be rewarded, because the former clashes with Asian values and the latter affirms them. We test this claim with survey experiments among 5,168 respondents in Indonesia, Malaysia, the Philippines, and Vietnam, paired with a comprehensive measure of Asian values as defined by their proponents. American criticism of illiberal policies reduces favorability by an average of one point on an eleven-point scale in all countries, while Chinese praise yields no systematic gains. Contrary to pre-registered hypotheses, neither effect is moderated by Asian values. Instead, the penalty for American criticism grows with nationalism. Backlash reflects resistance to foreign judgment rather than ideological dissonance, with implications for great-power rhetoric about partners' domestic politics.

Word count: 9,684

*Pre-analysis plan lodged at OSF on November 26, 2023. We are grateful to Renard Sexton, collaborator and co-principal investigator on the survey project from which these data come, whose partnership made this study possible. For helpful comments and conversations, we thank Kelly Matush and Dotan Haim. We also thank senior research associate Alicia Chen and research assistants Rico La Vina, Rocky Intan, Mai Truong, Jenny Kim, and Yan Liu, and TGM Research for respondent recruitment. Research carried out under Emory University IRB exemption #STUDY00002300 and UCSD IRB exemption #804650. Funding from the Carnegie Corporation of New York, the National Science Foundation, and the Koch Foundation is gratefully acknowledged.

[†]Associate Professor, School of Global Policy & Strategy, University of California San Diego. nra-vanilla@ucsd.edu

[‡]Postdoctoral Research Associate, National Security Data and Policy Institute, University of Virginia. dboldt@virginia.edu

1 Introduction

In 1994, four years after stepping down as Singapore’s prime minister, Lee Kuan Yew told Foreign Affairs what he thought of American values advocacy in Asia. “It is not my business to tell people what’s wrong with their system,” he said. “It is my business to tell people not to foist their system indiscriminately on societies in which it will not work” (Zakaria, 1994, 110). The warning distilled what became known as the Asian values thesis: the claim that Asian societies rest on a moral foundation that prizes community over the individual, harmony over dissent, and order over liberty, a position pressed by Lee, by Malaysia’s Mahathir Mohamad, and by the signatories of the 1993 Bangkok Declaration (Mahathir and Ishihara, 1995; Regional Meeting for Asia of the World Conference on Human Rights, 1993). Critics answered that the thesis was a myth told by autocrats about their own citizens (Kim, 1994; Sen, 1997). The debate was never resolved, and was instead absorbed into the practice of great-power diplomacy where it carried an enduring behavioral prediction: if citizens hold strong Asian Values and evaluate foreign rhetoric against them, they should reward messages that affirm those values and punish those that violate them.

The two powers have adopted sharply different rhetorical strategies toward the domestic politics of their regional partners. Beijing emphasizes non-interference, pledging to “refrain from imposing” its “own values or models on others and from stoking ideological confrontation.”¹ Washington, by contrast, continues to criticize restrictions on press freedom, human rights, and due process, while often tempering those criticisms where strategic considerations are greatest, and insisting throughout that America is “not asking anyone to choose between the United States and China.”² The clash of these approaches in Southeast

¹Xi Jinping, “Join Hands on the Path Towards Modernization,” keynote address launching the Global Civilization Initiative, Beijing, March 15, 2023 (Xi, 2023). The same formulation appears in the State Council white paper *A Global Community of Shared Future* (State Council Information Office of the People’s Republic of China, 2023).

²Secretary of State Antony Blinken, joint press availability, Riyadh, June 8, 2023 (Blinken, 2023). On muting: when Washington upgraded relations with Hanoi in September 2023, rights organizations objected that human rights had been set aside for strategic gain (Human Rights Watch, 2023; Agence France-Presse, 2023).

Asia has real strategic implications. When the ISEAS *State of Southeast Asia* survey asked the region’s policy elites in early 2023 which power ASEAN should align with if forced to choose, the United States led by twenty-two points. One year later, China led for the first time in the survey’s history (Seah et al., 2023, 2024).³ This suggests that how citizens respond when great powers publicly criticize or praise their own country’s domestic politics could swing the balance of influence within the region.

Existing work shows that public diplomacy can shape foreign opinion. High-level visits improve host-country approval (Goldsmith, Horiuchi and Matush, 2021; Goldsmith and Horiuchi, 2009). Aggressive “wolf warrior” rhetoric backfires while conciliatory rhetoric persuades (Mattingly and Sundquist, 2023; Xu, 2025). Chinese state media can sell the China model to audiences on five continents (Mattingly et al., 2025). This literature has yet to directly examine what great-power competition in Southeast Asia most often turns on: what a foreign power says about the host country’s own domestic politics. We test two competing accounts of how citizens respond to such rhetoric. The first is *ideological resonance*, which predicts that citizens reward messages aligned with their values and punish those that conflict with them. The second, which we develop below, emphasizes foreign judgment rather than ideological agreement.

We fielded parallel survey experiments in December 2023 in Indonesia, Malaysia, the Philippines, and Vietnam, four countries chosen for diversity in regime type, religion, and strategic alignment. Roughly 2,600 respondents in each experiment read a vignette in which a senior American diplomat criticized, or a Chinese diplomat praised, a policy of their own government. We randomized only the policy content: either an ideologically neutral issue—the government’s handling of fishing and pollution—or an illiberal policy, such as a crackdown on journalists or a violent drug war. The neutral arm lets us distinguish reactions to criticism itself from reactions to its ideological content. Before the experiment, every respondent completed a twelve-item measure of affinity for Asian values, the key moderator implied by

³The lead has continued to change hands, returning to the United States in 2025 (52.3 percent against 47.7) and to China in 2026 (52.0 percent against 48.0) (Seah et al., 2025; Lin et al., 2026).

the ideological-resonance account. All quantitative analyses follow a pre-registered analysis plan lodged before data collection; the minimal deviations and registered hypotheses that failed are documented in full in Section A.9 of the Supporting Information (SI).

Our first result establishes that citizens distinguish between foreign criticism of domestic politics and criticism of ordinary public policy. When the American diplomat criticized an illiberal policy, respondents marked him down by about one point on an eleven-point favorability scale, roughly a seventh of the control group’s rating, and marked down the United States itself by half a point or more. The penalty appears in all four countries. However, when the same diplomat criticized the neutral policy, in the same words and the same setting, respondents did not punish him at all. In other words, what costs Washington is not criticism itself but criticism of the host country’s domestic politics. Up to this point, the ideological resonance account appears to fit the evidence.

The initial pattern proves misleading in three important respects. First, the penalty has little to do with the values of the citizen hearing the rhetoric. Across a variety of possible codings and specifications, respondents high in Asian-values affinity punished the American diplomat no more than respondents who score low in Asian-values affinity. Instead, the penalty grows with nationalism. Second, Chinese praise of the same illiberal policies earns no systematic reward. Respondents did not register the praise as an affirmation of shared values when asked after the experiment what had influenced their ratings. In the pooled data, Beijing’s only gain came from praising the neutral policy, and that modest premium does not survive corrections for multiple testing, so we treat it as suggestive. Finally, we asked whether these rhetorical effects extend beyond evaluations of the diplomat and the country to the broader strategic question that great-power competition ultimately cares about: which power respondents think their country should align with if forced to choose. They do not. Neither power’s rhetoric moved respondents’ stated preferences over strategic alignment.

We argue that these results locate the backlash against foreign rhetoric in a different place than the ideological-resonance account assumes. Citizens do not appear to evaluate such rhetoric primarily by asking whether they agree with the values it expresses. Instead, they appear to evaluate who is claiming the authority to judge their country’s domestic politics. Criticism therefore triggers what we call a *sovereignty tax*: a penalty that grows with nationalism and falls on foreign criticism of the nation’s own political affairs. Praise of those same policies earns no corresponding premium because respect for the host country’s domestic politics is already the baseline citizens expect from outsiders. Southeast Asians simultaneously hold communitarian values and strong commitments to universal human rights, and in our data the two are positively correlated. The evidence therefore points to a condemnation of foreign judgment rather than a clash of values.

Our paper makes three contributions. First, we directly examine how citizens respond when foreign governments publicly criticize or praise their own country’s domestic politics amid great power competition. Existing work shows that public diplomacy shapes foreign opinion through visits, messaging, and media, but it has not isolated this form of rhetoric (Goldsmith and Horiuchi, 2009; Goldsmith, Horiuchi and Matush, 2021; Matush, 2023; Xu, 2025; Mattingly et al., 2025). In doing so, we join the literature on backlash against external pressure and interference in domestic politics (Gruffydd-Jones, 2019; Terman, 2023; Abramson, Menon and Gitlin, 2025; Tomz and Weeks, 2020; Bush and Jamal, 2015; Corstange and Marinov, 2012). In turn, we extend this literature in multiple ways. Whereas much of previous attention was paid to domestic-facing consequences, we focus primarily on the external-facing soft power penalty paid due to such pressure. We also study the consequences of foreign criticism and interference in the context of U.S.–China competition. Finally, we extend our focus to include both foreign criticism *and* praise of a country’s domestic politics.

Second, we contribute to the measurement and evaluation of the Asian values thesis. Our twelve-item battery provides a direct large-scale measure of the thesis’s core claims across four Southeast Asian countries. Affinity for Asian values is high, broadly shared, and

internally structured, yet it does not govern citizens' responses to foreign rhetoric about their own country's domestic politics. We therefore provide the first experimental test of the thesis in its public-diplomacy form, and it fails, echoing earlier survey evidence that authority orientations do not translate into the political preferences often attributed to them (Dalton and Ong, 2005; Shin, 2012).

Third, our findings reshape how we think about diplomatic rhetoric in great power competition. The constraint on American values advocacy is not that Southeast Asians reject liberal values. Rather, they reject foreign judgment. The portion of the public most likely to punish American criticism is not those most attached to Asian values but those most attached to the nation. Our results also hint at hard limits to soft power in the context of U.S.–China competition in Southeast Asia. Our results do suggest that the risk profile of diplomatic interference in domestic politics is asymmetric to the extent that the U.S. bears more risk than China does. However, diplomatic interference is unlikely to generate a realignment of soft power within the region: neither U.S. criticism nor Chinese praise of domestic politics causes the public to ‘change sides’ when forced to pick between the two powers.

The rest of the paper proceeds as follows. In the next section, we develop the resonance and sovereignty accounts as rival predictions, and we state the hypotheses we registered before collecting the data. Section 3 describes the research design. Section 4 presents the results of the registered tests, and Section 5 presents the exploratory analyses that adjudicate between the two accounts. Section 6 concludes with scope conditions and implications.

2 Public Diplomacy and the Clash of Ideologies

2.1 What Public Diplomacy Does, and What We Do Not Know

Great powers pursue foreign policy goals by accumulating what Nye (2004) calls soft power: the ability to attract rather than coerce. Public diplomacy pursues soft power by targeting citizens using visits, speeches, and other diplomatic messaging. A growing empirical literature shows that these efforts can work. High-level visits improve host-country publics' approval of the visiting power (Goldsmith, Horiuchi and Matush, 2021), even though the same trips earn the traveling leader nothing at home (Goldsmith, Horiuchi and Matush, 2026). Earlier cross-national work found that public diplomacy moves foreign opinion when its content aligns with the receiving public's policy concerns (Goldsmith and Horiuchi, 2009). The effects are not confined to democracies or to Western senders. Chinese state media persuades audiences across five continents that the China model is superior (Mattingly et al., 2025), and development spending buys goodwill for both powers, though unevenly (Blair, Marty and Roessler, 2022).

The literature has also shown that *what* governments say matters a good deal. Conciliatory Chinese appeals improve foreign evaluations, while combative “wolf warrior” rhetoric sours them, both abroad and at home (Mattingly and Sundquist, 2023; Xu, 2025). Leaders sometimes provoke this backlash deliberately because antagonizing foreign audiences can rally domestic ones (Matush, 2023). The institutional architecture of public diplomacy also shapes foreign opinion at the margins (Brazys and Dukalskis, 2019; Repnikova, 2022).

This literature has largely treated the content of public diplomacy as messages about the *sender*: its policies, values, institutions, or international behavior. Yet one important category of diplomatic rhetoric has remained unstudied: what a foreign government says about the recipient country's own domestic politics. The omission matters because this is precisely where American and Chinese public diplomacy diverge in Southeast Asia. American diplomats publicly criticize restrictions on press freedom, due process, and human rights. Chinese

diplomats instead emphasize non-interference and often praise the same governments. Existing research cannot isolate how citizens respond to this contrast. Visit studies (e.g., Goldsmith, Horiuchi and Matush, 2021) bundle rhetoric together with pageantry, agreements, and news coverage. Messaging studies (e.g., Mattingly and Sundquist, 2023; Mattingly et al., 2025) vary tone or claims about the sender’s own political system. Neither varies what a foreign government says about the *recipient* country’s domestic politics while holding everything else constant. That is the variation our experiments isolate. Whether citizens reward or punish such rhetoric depends on how they hear it. We develop two competing accounts.

2.2 The Ideological-Resonance Account

The first account applies a familiar proposition from persuasion research to great-power diplomacy. Messages are most persuasive when they resonate with the receiver’s prior values and least persuasive when they conflict with them. In Southeast Asia, this general logic takes on a specific cultural content through the Asian values thesis introduced above: publics in the region are said to prize community, hierarchy, harmony, and order, and to resist liberal individualism (Zakaria, 1994; Mahathir and Ishihara, 1995; Regional Meeting for Asia of the World Conference on Human Rights, 1993). We refer to this account as *ideological resonance*. If foreign governments publicly criticize domestic institutions that reflect these values, they should incur a penalty. If they instead praise those same institutions, they should earn a reward.

The ideological resonance account generates predictions at two levels, and the second is its distinctive implication. At the aggregate level, American criticism of illiberal domestic policies should reduce favorability toward the United States, while Chinese praise of those same policies should increase favorability toward China. At the individual level, however, both effects should be strongest among citizens with the greatest affinity for Asian values. The penalty for American criticism and the premium for Chinese praise should therefore increase monotonically with citizens’ commitment to those values, because resonance and

dissonance are properties of the match between a message and its receiver. If average treatment effects appear but this individual-level gradient does not, then the account has failed at its proposed mechanism.

The empirical standing of the Asian values thesis remains contested. Survey research finds that authority-respecting orientations coexist with strong aspirations for democracy (Dalton and Ong, 2005; Park and Shin, 2006; Chu et al., 2008; Shin, 2012), while modernization research argues that the region’s long-run trajectory points toward increasingly emancipative values (Welzel, 2013). Yet the thesis has never been tested in the form that matters for contemporary public diplomacy: as a theory of how citizens respond to foreign rhetoric about their own country’s domestic politics. Our design provides a particularly demanding test. We study four countries where affinity for Asian values is high, measure that affinity before treatment with a pre-registered instrument, and manipulate rhetoric directed precisely at the values the thesis claims should resonate.

2.3 The Sovereignty Account

The second account begins from a different premise. Before evaluating what a foreign government says about their country’s domestic politics, citizens first evaluate who is claiming the authority to say it. We argue that this judgment unfolds in two stages. The first is relational. The listener asks whether the speaker has standing to judge the nation’s political affairs. Only if the message passes this appraisal does the second stage, an evaluation of the message’s substantive content, come into play. We call this process *positional hearing*. Its foundations are familiar. Citizens evaluate foreign actors through relational schemata before factual ones (Herrmann et al., 1997); they are sensitive to slights against national status (Powers and Renshon, 2023; Renshon, 2016; Gries, 2004); and unsolicited foreign pressure reliably provokes resistance rather than persuasion (Gruffydd-Jones, 2019; Terman, 2023; Abramson, Menon and Gitlin, 2025; Tomz and Weeks, 2020; Bush and Jamal, 2015; Corstange and Marinov, 2012).

The two-stage sequence generates distinct predictions about foreign criticism and praise of domestic politics. Criticism of how a nation governs itself can concern its treatment of journalists, administration of justice, or protection of civil liberties. Such criticism asserts that the speaker has standing to judge the nation's political affairs. It therefore triggers the relational stage before the content of the message is considered. Criticism of an ordinary policy issue, by contrast, makes no comparable claim to authority and proceeds directly to the content stage, where it is evaluated on its merits. Praise works differently. Because positional hearing treats non-interference as the baseline outsiders owe one another, endorsement of a nation's domestic arrangements is relationally unremarkable. One earns no premium for meeting a baseline expectation. Praise of an ordinary policy, by contrast, makes no claim about the nation's domestic arrangements at all. Like neutral criticism, it proceeds directly to the content stage, where approval can earn a modest premium on its merits.

The relational appraisal should vary most directly with national attachment. The more strongly citizens identify with the nation, the more consequential an outsider's claim to judge its domestic affairs becomes. The theory therefore predicts that the penalty for foreign criticism will grow with nationalism. That penalty may nevertheless weaken when citizens have an independent reason to grant outsiders standing to speak. One such case is commitment to universal human rights. Citizens who regard how governments treat their people as a matter of universal rather than purely domestic concern may be more willing to accept foreign criticism of those practices. The penalty should therefore be smaller among citizens more committed to human rights. Citizens' own positions on the communitarian-liberal spectrum, by contrast, should matter little because values-based evaluations are seldom reached once the relational appraisal has been triggered.

The two accounts also differ in what they imply about the outcome that great-power competition ultimately seeks to influence, namely strategic alignment. Under ideological resonance, rhetoric about domestic politics should shape the same underlying attitudes that determine geopolitical preferences. Effective criticism and praise should therefore shift not

only evaluations of the foreign power but also citizens’ preferences over which great power their country should align with. Positional hearing predicts otherwise. It distinguishes between expressive evaluations and strategic preferences. A foreign government’s rhetoric can change how citizens evaluate a diplomat or even a country without altering the deeper strategic judgments that reflect interests, geography, and security. This distinction is consistent with classic work showing that foreign-policy attitudes are hierarchically organized. Specific evaluations are relatively malleable, while broader orientations remain comparatively stable (Hurwitz and Peffley, 1987). In Southeast Asia, alignment preferences should be especially stable because they are additionally anchored in the region’s long-standing strategy of hedging between rival great powers (Kuik, 2008; Ciorciari, 2010).

The two accounts are not mutually exclusive. We expect positional hearing to dominate when a great power comments on the listener’s own country’s domestic politics because such rhetoric directly raises questions of authority and national standing. Ideological resonance should become more important when those relational concerns are absent. This may occur when governments speak about one another rather than about the host country, when the speaker belongs to the listener’s own identity community, or when domestic political conflict has already polarized the values under discussion (Corstange and Marinov, 2012). Our experiments examine the first set of conditions, which also characterize much of day-to-day great-power diplomacy in Southeast Asia.

2.4 Hypotheses and Observable Implications

We now state the hypotheses we pre-registered before collecting the data. They formalize the ideological-resonance account, which motivated the experimental design. The first pair concerns average treatment effects:⁴

⁴The numbering follows the pre-analysis plan, in which H1 and H2 concern the American experiment and H3 and H4 the Chinese experiment. Because we present the hypotheses paired by type—average effects first, then moderation—they appear here in the order H1, H3, H2, H4.

Hypothesis 1 (H1): U.S. criticism of illiberal policies in Southeast Asian countries will reduce favorability toward the United States.

Hypothesis 3 (H3): Chinese praise of illiberal policies in Southeast Asian countries will increase favorability toward China.

The second pair concerns the individual-level moderation that distinguishes the resonance account:

Hypothesis 2 (H2): The negative effect of U.S. criticism on favorability toward the United States will be larger among citizens with stronger affinity for Asian values.

Hypothesis 4 (H4): The positive effect of Chinese praise on favorability toward China will be larger among citizens with stronger affinity for Asian values.

The sovereignty account was developed after pre-registration, and we therefore treat its predictions as exploratory throughout. Even so, the experimental design permits both accounts to be evaluated against the same evidence. Where the two accounts agree, the data cannot distinguish between them: both predict that American criticism of illiberal domestic policies will incur a penalty, and neither predicts much response to criticism or praise of an ideologically neutral policy. Where they diverge, the data can adjudicate between them. Ideological resonance predicts that Chinese praise should earn a premium and that both the penalty for American criticism and the premium for Chinese praise should grow with citizens' affinity for Asian values. The sovereignty account instead predicts no premium for Chinese praise, little role for Asian values, stronger moderation by nationalism than by values, and no movement in broader strategic alignment despite changes in favorability. We report the pre-registered tests first in Section 4 and then evaluate these competing implications in the exploratory analyses of Section 5.

3 Research Design

3.1 Cases and Sampling

We fielded the experiments in Indonesia, Malaysia, the Philippines, and Vietnam in December 2023. Together, the four countries span the principal dimensions along which Southeast Asia varies: regime type, from the Philippines’ competitive democracy to Vietnam’s single-party state; dominant religion; and baseline geopolitical orientation toward the United States and China. At the same time, they share the conditions required for our test. All four are arenas of active U.S.–China competition for public opinion, and all four, as we show below, are home to publics with high affinity for Asian values.

The experiments were embedded in a wave of a larger annual online survey on regional foreign-policy attitudes. TGM Research recruited respondents from its online panels using quotas for age, gender, and location, yielding samples that are broadly representative of each country’s internet-connected adult population. Internet penetration ranges from roughly 71 percent in Indonesia to 89 percent in Malaysia. Those beyond the reach of online surveys are disproportionately rural and older, so our estimates describe the segment of the public that great-power public diplomacy can realistically reach. Table 1 summarizes the samples.

After completing the demographic and baseline-attitude questions, respondents were randomly assigned to the American-rhetoric experiment, the Chinese-rhetoric experiment, or unrelated experimental modules. No respondent participated in more than one experiment, making the American and Chinese samples independent by design. We exploit this feature later when comparing the effectiveness of the two powers’ rhetoric directly. The study was classified as minimal risk by the institutional review boards of [Institution A] and [Institution B]. SI Section A.8 describes the ethics procedures, including the protections adopted for respondents in the region’s more restrictive political environments.

Table 1: Samples and design

	Survey	U.S.	China	Asian-values index			Female	Age
		experiment	experiment	Mean	SD	α		
Indonesia	1,507	679	681	6.40	1.79	0.90	0.48	34.7
Malaysia	1,408	634	636	6.71	1.87	0.93	0.49	34.5
Philippines	1,267	570	569	6.45	2.07	0.93	0.49	37.2
Vietnam	1,567	699	700	6.91	2.08	0.94	0.46	34.5

Note: Respondents not routed to either experiment completed unrelated survey modules. The Asian-values index averages twelve items scored 0–10; α is Cronbach’s alpha. Female and age are sample shares and means. Within each experiment, respondents were randomized evenly across the four arms; arm-level counts and balance tests are reported in SI Sections A.2 and A.3.

3.2 Treatments

Each experiment presents a two-paragraph vignette modeled on real diplomatic events of the study period. In the American experiment, the first paragraph, held constant across arms, reports that the U.S. Vice President spoke at the U.S.-ASEAN summit about the region’s importance and the need for cooperation, language adapted from Kamala Harris’s September 2023 summit remarks. The second paragraph reports that a senior American diplomat then met with the respondent country’s representative, and here the arms diverge. In the control arm the diplomat echoes the summit’s cooperative themes. In the neutral-policy arm (T1) the diplomat criticizes the government’s failure to address overfishing and pollution in its waters. In the two illiberal-policy arms (T2 and T3) the diplomat criticizes, in the same format and register, a domestic policy that conflicts with liberal-democratic norms. The Chinese experiment mirrors this structure exactly, with a senior Chinese diplomat who either echoes the ASEAN-China summit’s cooperative themes (control), praises the government’s handling of the same neutral policy (T1), or praises the same illiberal policies the American diplomat criticized (T2 and T3).

The illiberal policies are country-specific, chosen so that the rhetoric refers to a real and salient controversy in each country rather than a hypothetical one. Table 2 lists them. Within each country, the two policies were also chosen with an ordering in mind: policy A

Table 2: Country-specific illiberal policies in the T2 and T3 arms

	T2: Illiberal policy A	T3: Illiberal policy B
Indonesia	Discrimination against LGBTQ+ citizens	Blasphemy laws
Malaysia	Sedition Act restrictions on expression	Application of Sharia law
Philippines	Crackdown on journalists and media	Drug-war violence and killings
Vietnam	Internet surveillance and censorship	Persecution of political dissidents

Note: In the American experiment the diplomat criticizes the policy; in the Chinese experiment the diplomat praises the government’s handling of the same policy. The neutral-policy arm (T1) addresses overfishing and marine pollution in every country. By design, policy A is the comparatively less severe and policy B the more severe illiberal policy within each country’s context (see the text for discussion).

is the comparatively less severe of the pair, and policy B the more severe, in that country’s own context. In the Philippines, for example, restrictions on journalists (A) are set against the drug war’s killings (B). We recognize that reasonable readers may rank a given pair differently, and nothing in our main conclusions depends on the ranking. But the ordering gives the pooled arm coefficients a common interpretation: T2 pools the less severe policies and T3 the more severe ones. Full vignette texts for every country and arm, together with the exact wordings of every outcome and pre-treatment measure, appear in SI Section A.1.

The design does its work through two comparisons. The neutral-policy arm separates the cost of criticism itself from the cost of its ideological content, because the T1 diplomat criticizes with identical directness but touches no values controversy. And the mirrored structure prices the same policies twice, once as the object of American criticism and once of Chinese praise, which lets us ask whether the penalty and the premium are mirror images.

3.3 Outcomes

We measure three outcomes immediately after the vignette. Respondents rate the diplomat, from 0 (extremely unfavorable) to 10 (extremely favorable), and then rate the diplomat’s country on the same scale. The diplomat rating captures the messenger’s reception. The country rating captures how far the rhetoric’s effect travels up to the power itself. Respondents then report, for each of four factors, whether it pleased, displeased, or did not affect

their rating: the power’s stance toward their country’s domestic politics, its criticism or approval of their values, its role in regional security, and its economic role, with item wordings adapted to each power (exact wordings in SI Section A.1). These attribution items let respondents tell us which consideration the rhetoric activated, and we use them below when we distinguish between mechanisms. Finally, respondents place their country on a seven-point alignment scale anchored by complete alignment with China on the left and complete alignment with the United States on the right, an item modeled on the forced-choice question that regional elite surveys have tracked for years (Seah et al., 2023).

3.4 Measuring Affinity for Asian Values

Before the experiments, every respondent completed twelve items measuring affinity for Asian values. Each item is a trade-off statement scored from 0 (strongly disagree) to 10 (strongly agree). Our measurement strategy draws on the two traditions that have treated the Asian-values thesis as an object of survey research. The six dimensions we measure—collectivism, hierarchy, social harmony, family obligation, traditional governance, and authoritarian leadership—are those identified both in the thesis’s foundational texts (Zakaria, 1994; Regional Meeting for Asia of the World Conference on Human Rights, 1993) and in the Asian Barometer literature on authority orientations and Confucian social values (Dalton and Ong, 2005; Park and Shin, 2006; Shin, 2012; Chu et al., 2008). The trade-off format follows the classic approach to measuring values, asking respondents to prioritize between competing goods rather than endorse each independently, because most people endorse both (Welzel, 2013). One item, for example, asks respondents whether “Preserving social harmony and stability should be prioritized, even if it restricts freedom of speech and expression.” Each dimension appears in two versions, one abstract and one grounded in everyday life, so that the index does not depend on respondents’ comfort with abstract reasoning. The index averages the twelve items and is internally consistent in every country, with Cronbach’s alpha ranging from 0.90 to 0.94.

What the measure reveals is important because it speaks to the oldest objection to the Asian-values thesis: that these values are a story autocrats tell about their own citizens (Kim, 1994; Sen, 1997). Figure 1 suggests otherwise. Affinity for Asian values is real and widespread. Country means range from 6.40 in Indonesia to 6.91 in Vietnam, and only one respondent in six falls below the midpoint of the scale (panel A). It is also structured rather than indiscriminate. Respondents distinguish among the six dimensions: pooled across countries, collectivism receives the strongest support (7.05) and authoritarian leadership the weakest (6.19). Such differentiation is the hallmark of a coherent belief system rather than indiscriminate agreement with survey statements (panel B). Finally, affinity for Asian values coexists with strong commitments to universal human rights. The same respondents average between 8.5 and 9.2 on the statement that human rights are universal, and the two measures are positively correlated (panel C). Southeast Asians therefore hold the values identified by the thesis, but they do not hold them as the negation of liberal commitments. The terrain presupposed by the ideological-resonance account is therefore in place, making the test that follows a demanding one.

One measurement concern remains. Because all twelve items are coded so that agreement indicates greater affinity for Asian values, the index could partly capture a general tendency toward acquiescence. We return to this issue when testing moderation, where our conclusions are unchanged across four alternative constructions of the index, including one that explicitly removes acquiescence. SI Section A.5 reports additional construct-validity tests. The items are unidimensional; the index separates Malay and Chinese Malaysians by more than one and a half standard deviations; it declines with education, as theories of value formation predict; and it correlates selectively rather than uniformly with the survey's other agreement items, the pattern expected of a substantive construct rather than acquiescence.

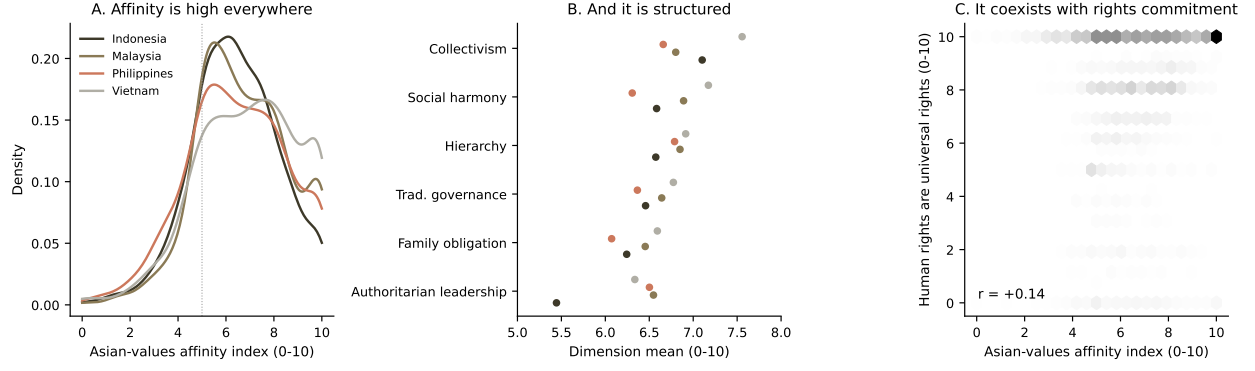


Figure 1: The Asian-values landscape. Panel A: kernel densities of the twelve-item affinity index by country; the dotted line marks the scale midpoint. Panel B: means of the six value dimensions by country (each dimension averages its abstract and everyday items). Panel C: joint distribution of the affinity index and agreement that “human rights are universal rights,” pooled; darker cells hold more respondents.

3.5 Estimation

We estimate treatment effects by ordinary least squares, comparing each treatment arm with the control group. For respondent i in country c , we estimate, separately for each experiment and outcome:

$$Y_{ic} = \alpha + \sum_{a=1}^3 \beta_a T_{ic}^a + \Gamma' X_{ic} + \delta_c + \varepsilon_{ic}, \quad (1)$$

where Y_{ic} is the outcome, T_{ic}^a indicates assignment to arm a (the control group is the omitted category, so each β_a is that arm’s effect relative to no policy rhetoric), δ_c are country indicators that keep every comparison within country, and X_{ic} is a vector of controls: age, gender, and category indicators for education, religion, ethnicity, and household income, each defined within country. Country-level models drop δ_c and estimate Equation 1 within each sample. When we test moderation, we add the moderator and its interactions with the arm indicators:

$$Y_{ic} = \alpha + \sum_{a=1}^3 \beta_a T_{ic}^a + \theta M_{ic} + \sum_{a=1}^3 \phi_a (T_{ic}^a \times M_{ic}) + \Gamma' X_{ic} + \delta_c + \varepsilon_{ic}, \quad (2)$$

where M_{ic} is the continuous Asian-values index (or, in exploratory analyses, nationalism, human-rights commitment, or another pre-treatment moderator), and each ϕ_a measures how arm a 's effect changes with a one-point increase in the moderator. Throughout, we report standard errors that are robust to unequal variances. Randomization achieved balance: across forty-eight covariate-by-arm tests, four reject equality at the five percent level, slightly more than the 2.4 that chance predicts, and every result we report is unchanged without controls. Balance tables and arm-level counts appear in SI Sections A.2 and A.3, and all robustness analyses appear in SI Section A.5. Every analysis can be reproduced from the replication package described in SI Section A.7.

4 Results

Drawing on two parallel vignette experiments in four countries, we find that American criticism of a host government's illiberal policies generates a large and consistent penalty, that identical criticism of a neutral policy generates none, and that Chinese endorsement of the same policies generates nothing at all. The average treatment effects for both experiments are summarized in Figure 2, and the full regression results appear in Tables 4 and 5. We walk through the results in the order of the registered hypotheses, beginning with where the two powers stood before any rhetoric was delivered.

Before any policy rhetoric, the two powers start from very different positions, and the control-group means in Table 3 show these starting positions. Respondents in the Philippines and Vietnam rate the United States about two points higher than China (7.16 against 5.26, and 7.15 against 4.97), Malaysia leans the other way (5.22 against 6.06), and Indonesia is nearly even (6.27 against 6.13). The forced-alignment item points in the same direction: leaning toward the United States in Manila and Hanoi, close to neutral in Jakarta, and tilting toward Beijing in Kuala Lumpur. These baselines echo the geography of the regional

elite surveys (Seah et al., 2024), and they matter for what follows, because the effects we report are remarkably uniform across these very different starting positions.

Hypothesis 1 predicted that American criticism of illiberal policies would depress favorability toward the United States, and it does, substantially. Pooling the four countries, respondents who read the American diplomat criticizing an illiberal policy rated him about one point lower than respondents in the control group: -1.10 points on the eleven-point scale for the less severe illiberal policies (T2) and -0.97 for the more severe ones (T3), each about 15 percent below the control-group mean of 6.75, or roughly four tenths of a standard deviation. The penalty travels beyond the messenger: ratings of the United States itself fall by 0.72 points in T2 and 0.52 in T3, about 8 to 11 percent below the control mean. Both estimates are precise and essentially identical with the full battery of controls.

The neutral-policy arm is what allows us to interpret this penalty, because it tells us what the penalty is for. When the American diplomat criticized the host government’s handling of overfishing and pollution, with the same directness, in the same meeting, respondents did not punish him: the effect is -0.18 points on the diplomat rating and -0.12 on the country rating, neither distinguishable from zero. In the pooled data, the gap between the neutral arm and each illiberal arm is statistically significant in every comparison. In other words, criticism as such costs the United States very little. What costs is the ideological content of the criticism. We take this as strong support for Hypothesis 1, with one addendum: the hypothesis appears to be right for a reason the registered account did not anticipate, a thread we pick up in the next section.

Hypothesis 3 predicted the mirror image for Beijing: endorsement of the same illiberal policies should raise favorability toward China. It does not. Pooled across countries, the Chinese diplomat who praised an illiberal policy gained nothing, $+0.01$ points in T2 and $+0.02$ in T3, with confidence intervals ruling out premiums larger than about a quarter of a point. Ratings of China itself are just as flat. The only positive movement anywhere in the Chinese experiment comes from the neutral arm: praising the host’s environmental

management earns the diplomat +0.32 points, about 5 percent above the control mean of 6.11, and significantly more than the illiberal endorsements earn. That premium, however, is the weakest result in the study. It is the one estimate whose significance does not survive a correction for the eighteen main tests we run (Benjamini-Hochberg $q = 0.08$), so we treat it as suggestive. Its direction, however, is consistent with the content-stage logic of positional hearing, under which praise that raises no relational stakes can be evaluated on its substantive merits. The evidence rejects Hypothesis 3. In the pooled data, Chinese praise of illiberal policies yields no measurable gains.

Hypotheses 2 and 4 contain the ideological-resonance account’s distinctive prediction: that responses to foreign rhetoric should vary with citizens’ own affinity for Asian values. They do not. Figure 4 plots the treatment effects across the full distribution of the Asian-values index. In the American experiment, the penalty for criticizing illiberal policies is essentially constant. The interaction between the illiberal treatments and the index is +0.11 points per index point for T2 and +0.06 for T3, neither statistically significant, and both opposite in sign to the registered prediction. At the 25th percentile of Asian-values affinity, the estimated penalty is -1.26 points; at the 75th percentile, it is -0.95 . Both estimates are statistically significant. Citizens with strong affinity for Asian values punish the American diplomat no more than those with weaker affinity, and if anything slightly less. The Chinese experiment tells the same story. The corresponding interactions are -0.11 and -0.12 per index point, again statistically insignificant and again opposite in sign to the registered prediction. Whatever modest premium Chinese praise receives among respondents with low affinity disappears rather than grows as affinity increases, the opposite of what ideological resonance predicts.

Because all twelve items are coded in the same direction, we asked whether the nulls could be a measurement artifact. They are not. The results are unchanged when we use an index purged of acquiescence, an index net of human-rights commitment, a political subindex based on the authoritarian-leadership and traditional-governance items, or a within-country

standardized index. Under the acquiescence-corrected codings, the wrongly signed interactions become slightly larger, and a few reach statistical significance, strengthening rather than rescuing the conclusion. Nor is the specification responsible. In SI Section A.5, we search for the missing gradient wherever the ideological-resonance account predicts it should appear: matching each vignette to the value it directly implicates, relaxing the functional form (Hainmueller, Mummolo and Xu, 2019), comparing extreme deciles, estimating each item separately, and restricting the sample by value coherence, reading time, and country. Across roughly thirty specifications, the interaction is never statistically significant in the predicted direction. We therefore reject Hypotheses 2 and 4. The penalty for American criticism and the absence of a premium for Chinese praise are unrelated to citizens’ affinity for Asian values.

The forced-alignment outcome is a different kind of null, and an informative one. Across both experiments, no treatment moves the choose-sides item: the largest estimate in the pooled data is about one-tenth of a point on the seven-point scale, and the confidence intervals rule out shifts larger than about a quarter of a point in either direction. Favorability toward a diplomat can fall by a full point while respondents’ answer to the question, “Which power should our country align with, if forced to choose?,” remains unchanged. We do not regard this as a failure of the treatments. Foreign-policy attitudes are hierarchically organized, with relatively stable general orientations and more malleable specific evaluations (Hurwitz and Peffley, 1987), and strategic alignment is among the broadest orientations our survey measures. It is therefore unsurprising that when the ISEAS regional elite surveys recorded large swings on this same question, observers attributed them to wars and trade disruptions rather than to diplomatic rhetoric (Seah et al., 2024). Our treatments move the evaluations they directly target, but not the broader strategic orientation beneath them. This distinction is also consistent with the sovereignty account’s contrast between expressive evaluations and strategic preferences.

Figure 3 reports the treatment effects separately by country (full estimates appear in SI Section A.4). The country-level results reinforce both headline findings. The penalty for American criticism appears in all four countries, ranging from -0.64 points in Malaysia to -1.79 in Indonesia for the pooled illiberal treatments. These differences are meaningful (a joint test rejects equal effects at $p < 0.001$), with Indonesia’s response to criticism of its LGBTQ+ policies (-2.5 points, more than one-third of the control mean) the largest effect in the study. The ordering of the policy pairs also provides some evidence of a dose response. In Malaysia, the Philippines, and Vietnam, the more severe policy (B) produces the larger penalty. Indonesia is the clear exception, suggesting that what made policy A especially costly was not its severity but its subject matter. Vietnam is the only country where criticism of the neutral policy also incurs a penalty (about -0.56 points), eliminating the difference between the neutral and illiberal treatments. We interpret this as a scope condition rather than an anomaly: in the region’s most closed polity, criticism of the government appears costly regardless of its subject.

The Chinese experiment shows a very different pattern. Treatment effects are statistically indistinguishable across countries ($p = 0.15$), indicating that the pooled null does not mask offsetting national responses. The one country-level estimate that departs from this pattern is Indonesia, where praise of illiberal policies is associated with a gain of about $+0.50$ points ($p < 0.05$)—the same country that punished American criticism most severely. Because the joint test does not reject homogeneity, we treat this estimate as suggestive rather than a demonstrated exception, but it is consistent with Indonesia being the most responsive audience in our sample in both directions.

These results are robust in every way we know how to check: ordered models reproduce the pattern, dropping the fastest decile of vignette readers sharpens the American penalties rather than weakening them, randomization inference returns $p < 0.001$ for the American illiberal arms and $p > 0.9$ for the Chinese ones, and the four American penalty estimates survive the multiple-testing correction at $q \leq 0.001$ (full tables in SI Section A.5).

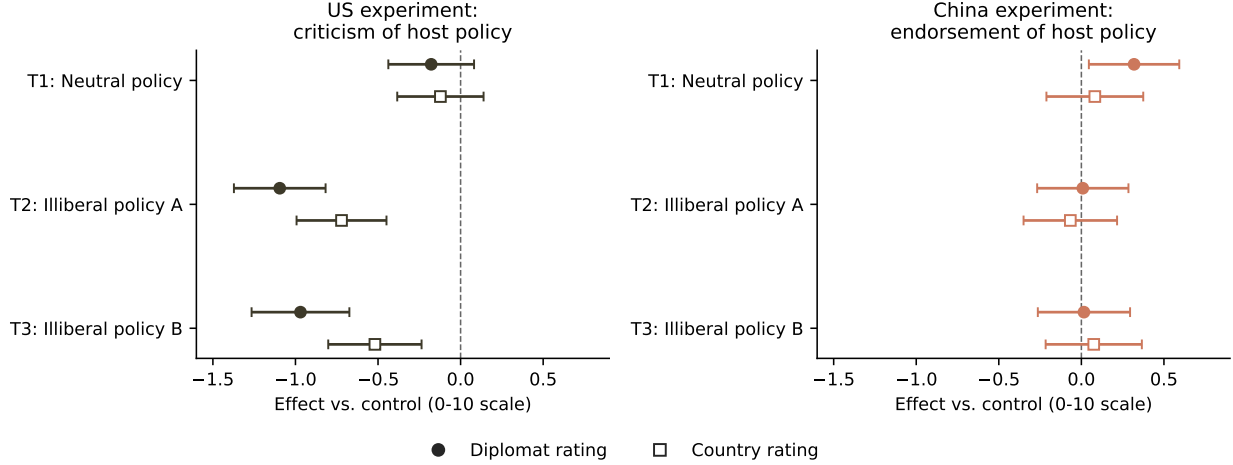


Figure 2: Average treatment effects of great-power rhetoric, pooled across countries. Markers are point estimates from Equation 1 without controls; horizontal bars are 95 percent confidence intervals. Filled circles: rating of the diplomat (0–10). Hollow squares: rating of the diplomat’s country (0–10). $N = 2,551$ (U.S. experiment) and 2,557 (China experiment) for the diplomat outcome.

Taken together, the registered tests deliver one confirmation and three rejections. The penalty is real, large, and everywhere (H1). The praise premium does not exist (H3). Neither is governed by the audience’s Asian values (H2, H4). What remains is to establish what the penalty is actually made of, and that is the work of the next section.

5 Exploratory Analysis: Locating the Clash

The registered analyses establish two facts. American criticism of illiberal domestic policies incurs a substantial penalty, while that penalty is unrelated to citizens’ affinity for Asian values. Those findings reject the ideological-resonance account’s distinctive mechanism, but they do not by themselves establish an alternative. In this section, we evaluate the additional implications of the sovereignty account developed above. If citizens respond first to the authority of a foreign government to judge their country’s domestic politics, rather than to the values expressed in the message itself, then we should observe four further patterns: criticism and praise should become asymmetric, respondents should react to foreign judgment

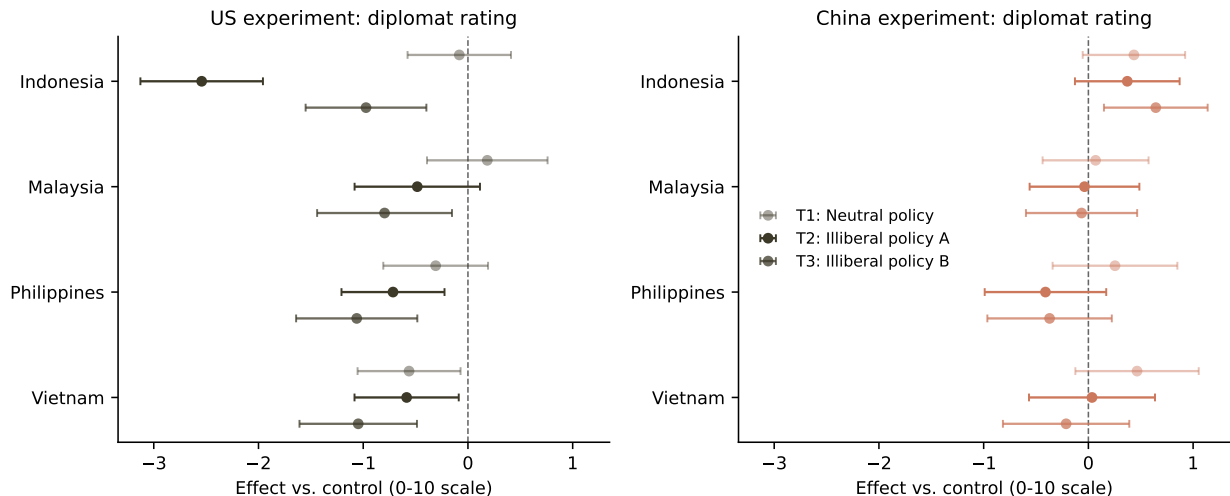


Figure 3: Effects on the diplomat rating by country and arm. Markers are point estimates against the within-country control group; horizontal bars are 95 percent confidence intervals. Shading distinguishes the neutral-policy arm (lightest) from the two illiberal-policy arms.

rather than ideological affirmation, nationalism rather than Asian values should govern the penalty, and changes in favorability should not extend to broader strategic alignment. We treat these analyses as exploratory throughout.

We begin with the asymmetry between criticism and praise, the sovereignty account's central prediction. Because respondents were randomly assigned to either the American or the Chinese experiment, the two samples are comparable by design and can be stacked: the outcome is favorability toward the treating power's diplomat, while treatment effects are allowed to differ across powers. Table 6 reports the resulting estimates.⁵ For the neutral-policy treatment, the two powers' returns are essentially symmetric: American criticism costs -0.21 , Chinese praise earns $+0.30$, and we cannot reject the hypothesis that the two effects are equal and opposite ($p = 0.62$ for the diplomat rating; $p = 0.57$ for the power rating). Once the rhetoric turns to the host country's domestic politics, however, the symmetry disappears. For the illiberal-policy treatments, the equal-and-opposite hypothesis is decisively rejected (both arms $p < 0.001$ for the diplomat rating), because American criticism incurs roughly

⁵The stacked specification estimates common country fixed effects across the two samples and omits controls, so its arm effects differ slightly from the single-experiment estimates in Tables 4 and 5 (for example, -0.21 rather than -0.18 for American criticism of the neutral policy).

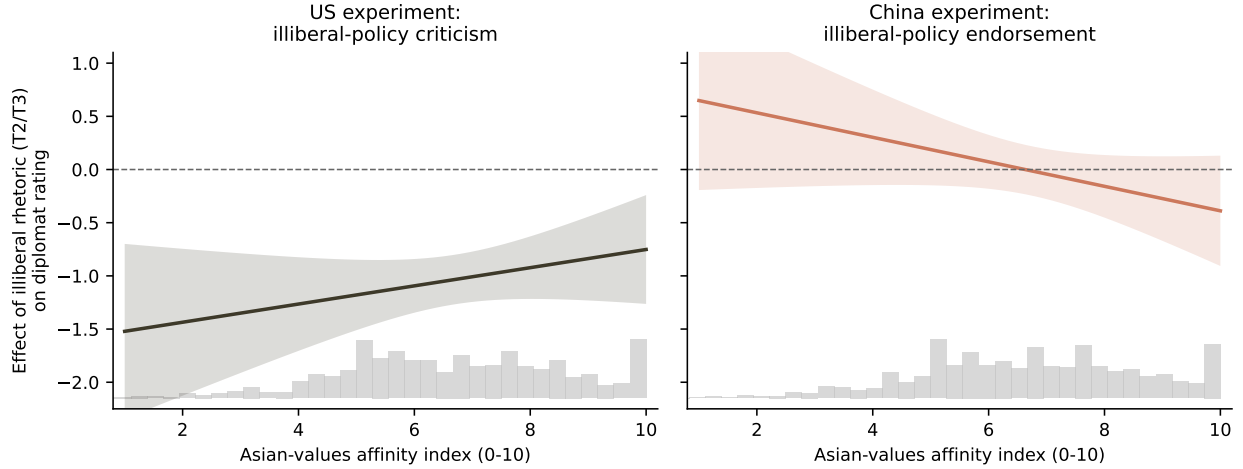


Figure 4: The effect of illiberal-policy rhetoric (T2 and T3 pooled) on the diplomat rating across the Asian-values affinity index, with 95 percent confidence bands. Gray bars show the distribution of the index. The registered prediction is a downward slope in the left panel and an upward slope in the right panel; neither appears.

a one-point penalty while Chinese praise earns essentially no premium. The asymmetry is therefore specific to rhetoric about domestic politics. Criticism is taxed, while praise merely meets the baseline expectation of non-interference.

Respondents’ own explanations point in the same direction. After rating the diplomat and the country, respondents reported whether each of four considerations pleased, displeased, or did not affect them. The baseline pattern among control-group respondents is itself revealing. The power’s economic role is viewed positively (a net +0.52 on the -1 to $+1$ scale), as is its security role (+0.24). By contrast, both “American criticism of the country’s values” (-0.29) and “telling the country how to run its domestic politics” (-0.25) are viewed negatively. The treatments amplify only one of these reactions. American criticism of an illiberal policy makes respondents significantly more displeased about being told how to run their country’s domestic politics (-0.08 , $p < 0.05$), while attitudes toward liberal-democratic values themselves change little (full attribution tables appear in SI Section A.6). The Chinese experiment is equally informative for what does not happen. Praise of illiberal policies does not make respondents any more positive toward “China’s approval of policies

Table 3: Favorability and alignment absent policy rhetoric (control-group means)

	Indonesia	Malaysia	Philippines	Vietnam
<i>US experiment, control group</i>				
Opinion of US diplomat (0–10)	6.69	5.83	7.15	7.29
Opinion of the US (0–10)	6.27	5.22	7.16	7.15
Choose sides (1=China ... 7=US)	4.82	4.35	5.31	4.92
<i>China experiment, control group</i>				
Opinion of Chinese diplomat (0–10)	6.36	6.44	6.04	5.61
Opinion of China (0–10)	6.13	6.06	5.26	4.97
Choose sides (1=China ... 7=US)	4.50	4.08	5.37	5.02

Note: Entries are control-group means. Favorability runs from 0 (extremely unfavorable) to 10 (extremely favorable). The choose-sides item runs from 1 (complete alignment with China) to 7 (complete alignment with the United States).

in line with Asian values” (+0.03, not significant). Citizens respond to foreign judgment, not to ideological affirmation.

Who, then, pays the sovereignty tax? The answer depends on whether the penalty reflects citizens’ values or their attachment to the nation. Before comparing moderators, it is worth noting how closely they are related. Affinity for Asian values correlates with nationalism at +0.32 (ranging from 0.22 in Indonesia to 0.38 in Malaysia), with perceived military strength at +0.40, and with government satisfaction at +0.39. Nationalism itself is nearly at ceiling, averaging 8.6. These correlations raise an obvious question: do these measures capture the same underlying disposition? To answer it, we estimate all four moderators jointly. Figure 5 plots the resulting interactions with illiberal-policy rhetoric, with all moderators standardized, and Table 7 reports the arm-specific estimates.

Nationalism dominates. Each additional point on the nationalism scale deepens the T2 penalty by -0.22 ($p < 0.05$) and the T3 penalty by -0.30 ($p < 0.01$). With a standard deviation of about 1.7 points, a one-standard-deviation increase in nationalism deepens the sovereignty tax by roughly four-tenths to half a point. Commitment to universal human rights pulls in the opposite direction, reducing the penalty by +0.14 per scale point ($p < 0.01$) for the T2 policies, consistent with citizens who grant the foreign critic standing being more

Table 4: Average treatment effects of American criticism of host policies (H1)

	Opinion of US diplomat		Opinion of United States		Choose sides (1=China...7=US)	
	(1)	(2)	(3)	(4)	(5)	(6)
T1: Neutral policy	-0.178 (0.132)	-0.176 (0.138)	-0.122 (0.133)	-0.121 (0.136)	-0.102 (0.072)	-0.088 (0.075)
T2: Illiberal policy A	-1.095* (0.142)	-1.075* (0.145)	-0.721* (0.139)	-0.698* (0.142)	-0.038 (0.072)	-0.011 (0.075)
T3: Illiberal policy B	-0.970* (0.151)	-1.016* (0.154)	-0.519* (0.144)	-0.580* (0.147)	-0.082 (0.071)	-0.068 (0.073)
Controls	No	Yes	No	Yes	No	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Control-group mean	6.75	6.75	6.45	6.45	4.84	4.83
p: T1=T2	0.000	0.000	0.000	0.000	0.380	0.297
p: T1=T3	0.000	0.000	0.008	0.002	0.784	0.785
p: T1=T2=T3	0.000	0.000	0.000	0.000	0.666	0.554
Observations	2,551	2,549	2,529	2,528	2,460	2,458
Adj. R ²	0.063	0.080	0.091	0.105	0.071	0.073

Note: OLS estimates of Equation 1, pooled across countries, U.S. experiment. The control group (no policy rhetoric) is the omitted category. Robust standard errors in parentheses. Controls: age, gender, and within-country category indicators for education, religion, ethnicity, and household income. * $p < 0.05$.

forgiving. The corresponding T3 interaction is positive but small and insignificant (+0.03), so we treat this attenuation as suggestive. Asian-values affinity provides no moderation in the direction the resonance account predicts. Once its correlation with nationalism is taken into account, the estimated interactions move to the opposite side of zero (+0.18 and +0.20, the latter significant at $p < 0.05$), implying that the region's liberals punish slightly *more* than those with stronger affinity for Asian values. Two measures that are positively correlated in the population therefore push the treatment effect in opposite directions once they are separated. This is the clearest evidence in favor of positional hearing: what governs the penalty is attachment to the nation, not attachment to the values under discussion. In the Chinese experiment, by contrast, none of the moderators meaningfully conditions the treatment effects (SI Section A.6).

The nationalism result raises one final alternative explanation. Perhaps respondents are not defending the nation at all, but simply defending their own government from foreign criticism. The government-satisfaction measure allows us to distinguish between these possibilities, and it points toward the former. The penalty remains essentially unchanged

Table 5: Average treatment effects of Chinese endorsement of host policies (H3)

	Opinion of Chinese diplomat		Opinion of China		Choose sides (1=China...7=US)	
	(1)	(2)	(3)	(4)	(5)	(6)
T1: Neutral policy	0.319* (0.140)	0.384* (0.139)	0.082 (0.150)	0.156 (0.149)	0.116 (0.072)	0.073 (0.071)
T2: Illiberal policy A	0.009 (0.141)	0.044 (0.142)	-0.067 (0.144)	-0.038 (0.146)	-0.005 (0.072)	-0.027 (0.073)
T3: Illiberal policy B	0.016 (0.143)	0.098 (0.143)	0.075 (0.149)	0.156 (0.149)	0.075 (0.072)	0.034 (0.072)
Controls	No	Yes	No	Yes	No	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Control-group mean	6.11	6.11	5.62	5.62	4.69	4.69
p: T1=T2	0.024	0.013	0.307	0.184	0.095	0.172
p: T1=T3	0.029	0.038	0.965	1.000	0.574	0.588
p: T1=T2=T3	0.036	0.026	0.504	0.303	0.237	0.389
Observations	2,557	2,555	2,526	2,524	2,446	2,444
Adj. R ²	0.027	0.063	0.050	0.081	0.115	0.138

Note: OLS estimates of Equation 1, pooled across countries, China experiment. The control group (no policy rhetoric) is the omitted category. Robust standard errors in parentheses. Controls as in Table 4. $*p < 0.05$.

among respondents dissatisfied with their national government, and the interactions with government satisfaction are small and statistically insignificant (SI Section A.6). Citizens who disapprove of the very governments whose policies the American diplomat criticized nevertheless punish the foreign criticism. Political knowledge does not moderate the penalty either. By contrast, attention to domestic news amplifies it: respondents who follow domestic politics most closely react most strongly to the T3 policies (-0.24 per attention point, $p < 0.01$). The sovereignty tax is therefore not concentrated among government supporters or the politically uninformed. It is strongest among nationalists and the politically attentive, consistent with citizens defending the standing of the nation rather than the reputation of the government.

Across these exploratory analyses, the evidence consistently favors the sovereignty account. The penalty for American criticism is broad rather than concentrated among citizens with strong affinity for Asian values. It grows with nationalism. Criticism and praise are treated asymmetrically once they concern the host country's domestic politics: criticism incurs a sovereignty tax, while praise earns no corresponding premium because non-interference

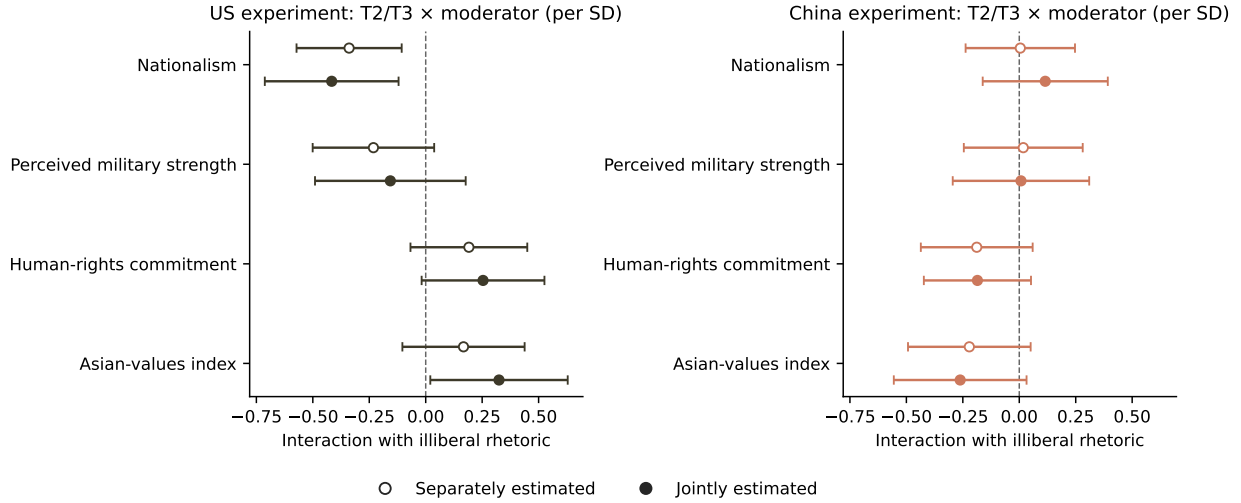


Figure 5: Interactions between illiberal-policy rhetoric (T2 and T3 pooled) and each standardized moderator, estimated separately (hollow) and jointly (filled), with 95 percent confidence intervals. In the American experiment, nationalism deepens the penalty and human-rights commitment attenuates it, while Asian-values affinity sits on the wrong side of zero for the resonance account. In the Chinese experiment nothing moderates.

is already the expected baseline. Respondents themselves attribute the penalty to foreign judgment rather than ideological disagreement, and the treatments shift evaluations of diplomats without altering broader strategic preferences. The country-level patterns fit the same account. Vietnam, the region's most closed polity, taxes criticism regardless of its content, while Indonesia, the most responsive audience in the sample, both punishes American criticism most severely and is the only country that appears to reward Chinese praise. These analyses are exploratory, and we have not directly observed the relational appraisal that positional hearing posits. Nevertheless, a single mechanism organizes every major pattern that the ideological-resonance account leaves unexplained.

Table 6: The asymmetry of returns: stacked estimates across experiments

	(1) Diplomat	(2) Power
T1: Neutral policy	-0.206 (0.133)	-0.168 (0.137)
T2: Illiberal policy A	-1.095* (0.146)	-0.716* (0.145)
T3: Illiberal policy B	-0.976* (0.153)	-0.531* (0.149)
China experiment	-0.637* (0.137)	-0.830* (0.143)
T1 $\times$ China	0.509* (0.194)	0.220 (0.205)
T2 $\times$ China	1.092* (0.204)	0.621* (0.208)
T3 $\times$ China	0.981* (0.211)	0.582* (0.214)
p: equal-and-opposite, T1	0.620	0.570
p: equal-and-opposite, T2	0.000	0.000
p: equal-and-opposite, T3	0.000	0.024
Observations	5,108	5,055
Adj. R ²	0.017	0.016

Note: OLS on the stacked samples: the outcome is the rating of the treating power’s diplomat (column 1) or of the treating power itself (column 2), both 0–10. Arm effects are for the American experiment; interactions with the China-experiment indicator give the difference in each arm’s effect across powers. Country fixed effects included; robust standard errors in parentheses. The equal-and-opposite tests ask whether the American effect and the Chinese effect sum to zero. * $p < 0.05$.

Table 7: Moderator horse-race, American experiment (diplomat rating)

	(1) Nationalism	(2) Mil. strength	(3) Human rights	(4) Asian values	(5) Horse-race
T1: Neutral policy	0.502 (0.685)	0.261 (0.498)	-0.467 (0.406)	0.456 (0.555)	0.510 (0.802)
T2: Illiberal policy A	0.240 (0.708)	-0.623 (0.509)	-2.062* (0.446)	-1.836* (0.580)	-1.000 (0.856)
T3: Illiberal policy B	1.456* (0.707)	0.193 (0.529)	-1.104* (0.442)	-1.498* (0.588)	0.699 (0.822)
T1: Neutral policy $\times$ Nationalism	-0.076 (0.080)				-0.038 (0.096)
T2: Illiberal policy A $\times$ Nationalism	-0.153 (0.083)				-0.218* (0.103)
T3: Illiberal policy B $\times$ Nationalism	-0.286* (0.084)				-0.295* (0.107)
T1: Neutral policy $\times$ Mil. strength		-0.063 (0.072)			-0.009 (0.086)
T2: Illiberal policy A $\times$ Mil. strength		-0.071 (0.075)			-0.073 (0.090)
T3: Illiberal policy B $\times$ Mil. strength		-0.178* (0.077)			-0.107 (0.096)
T1: Neutral policy $\times$ Human rights			0.037 (0.050)		0.036 (0.051)
T2: Illiberal policy A $\times$ Human rights			0.124* (0.053)		0.142* (0.055)
T3: Illiberal policy B $\times$ Human rights			0.011 (0.054)		0.034 (0.057)
T1: Neutral policy $\times$ Asian values				-0.086 (0.081)	-0.079 (0.090)
T2: Illiberal policy A $\times$ Asian values				0.114 (0.083)	0.177 (0.091)
T3: Illiberal policy B $\times$ Asian values				0.072 (0.086)	0.196* (0.099)
Observations	2,528	2,526	2,535	2,477	2,448
Adj. R ²	0.090	0.093	0.086	0.146	0.153

Note: OLS estimates of Equation 2, pooled with country fixed effects and controls, U.S. experiment. Columns 1–4 interact the arms with each moderator separately; column 5 includes all four moderators and their interactions jointly. Moderators are on their original 0–10 scales. Robust standard errors in parentheses. * $p < 0.05$.

6 Conclusion

In this paper, we tested a longstanding idea about how citizens respond when great powers publicly discuss their country's domestic politics. We called this idea ideological resonance. It underlies both Beijing's strategy of non-interference and the caution Washington often exercises when criticizing human rights practices in Southeast Asia. We tested the account under conditions that should have favored it: in four countries where affinity for Asian values is high, using a measure derived from the thesis's own canon, and with hypotheses pre-registered before data collection. The account correctly predicts one empirical regularity: American criticism of illiberal domestic policies incurs a substantial penalty, while criticism of an ideologically neutral policy does not. But it fails at its distinctive mechanism. The penalty does not vary with citizens' affinity for Asian values, and China earns no systematic reward for praising those same policies. Instead, the penalty grows with nationalism, appears among critics and supporters of the government alike, and never extends to strategic alignment.

What do these results imply for great-power diplomacy? They suggest that the costs of public criticism stem from a different source than is usually assumed. The penalty does not appear to reflect a clash between liberal and Asian values. Instead, it reflects resistance to foreign judgment. Citizens react not because they reject the values being expressed, but because they reject an outsider's claim to authority over their country's domestic politics. For Washington, this distinction matters. The constraint on American values advocacy is not that Southeast Asians reject liberal values. Indeed, respondents expressed stronger commitment to universal human rights than to Asian values. The constraint is foreign judgment. Unlike a deep conflict over values, that constraint may admit different diplomatic strategies. Our findings suggest, though they do not test, that criticism voiced by domestic actors, regional organizations, or multilateral institutions may incur a smaller penalty because it carries a weaker claim of great-power standing. For Beijing, the implication runs in the opposite direction. Respect for domestic sovereignty does not appear to be a source of diplomatic advantage. It is closer to a baseline expectation that earns little reward when met

but carries costs when violated. In the competition over rhetoric about domestic politics, Washington mostly risks losses while Beijing has little to gain.

Our findings also speak to the older debate over Asian values, but in a different way than that debate has usually been framed. To test the ideological-resonance account, we first had to measure affinity for Asian values, and our twelve-item battery, fielded to more than five thousand respondents in four countries, is, to our knowledge, the most direct large-scale measurement of the thesis's own claims to date. The measure suggests that the values are real. They are widely held across all four countries, internally structured rather than indiscriminate, and coexist with strong commitments to universal human rights. What the Asian-values thesis appears to have misidentified is not the values themselves but where they matter. In our data, the clash of ideologies does not take place within citizens, whose communitarian and liberal commitments coexist. It takes place between citizens and foreign states. Citizens respond less to the ideological content of foreign rhetoric than to the foreign government's claim to authority over their country's domestic politics. The returns to public diplomacy, at least in this setting, therefore appear to be governed more by sovereignty than by ideological agreement. They also suggest limits to what diplomatic rhetoric alone can accomplish: in our experiments, favorability toward a diplomat could fall by a full point without any detectable movement in strategic alignment.

Our evidence also has clear limits. First, we study internet-connected publics in four Southeast Asian countries that, while diverse in regime type and geopolitical orientation, share a region, a colonial past, and sustained exposure to U.S.–China competition. Second, our outcomes are immediate attitudinal responses to a single experimental vignette. Extended campaigns delivered through real media environments may produce different dynamics. Third, we do not observe the relational appraisal at the heart of positional hearing directly. Instead, we infer it from a consistent pattern of asymmetries, moderation effects, attributions, and strategic null effects.

The sovereignty account also generates clear scope conditions for future research. We expect positional hearing to be strongest when a great power comments on the listener's own country's domestic politics, where questions of authority and national standing are most salient. Ideological resonance may play a larger role when those relational concerns are absent—for example, when governments speak about one another rather than about the host country, when the speaker belongs to the listener's own identity community, or when domestic political conflict has already polarized the values under discussion. These are natural settings in which to distinguish the two accounts.

Thirty years after the region's leaders argued that Asia's values were different, the citizens of Southeast Asia offer a more precise answer. The values are real, and they are widely shared. But they do not determine how citizens respond to foreign governments that publicly judge their country's domestic politics. That response appears to be governed first by sovereignty, and only then, if at all, by ideology.

References

- Abramson, Yehonatan, Anil Menon and Abir Gitlin. 2025. “Whose critique matters? The effects of critic identity and audience on public opinion.” *American Journal of Political Science* 69:355–370.
- Agence France-Presse. 2023. “Biden Accused of Sidelining Vietnam and India Rights Over Strategic Interests.” *U.S. News & World Report*, September 11. <https://www.usnews.com/news/us/articles/2023-09-11/biden-accused-of-sidelining-vietnam-and-india-rights-over-strategic-interests>.
- Blair, Robert A., Robert Marty and Philip Roessler. 2022. “Foreign Aid and Soft Power: Great Power Competition in Africa in the Early Twenty-first Century.” *British Journal of Political Science* 52(3):1355–1376.
- Blinken, Antony J. 2023. “Secretary Antony J. Blinken and Saudi Foreign Minister Prince Faisal bin Farhan Al Saud at a Joint Press Availability.” U.S. Department of State, Riyadh, June 8. <https://2021-2025.state.gov/secretary-antony-j-blinken-and-saudi-foreign-minister-prince-faisal-bin-farhan-al-saud-at-a-joint-press-availability/>.
- Brazys, Samuel and Alexander Dukalskis. 2019. “Rising Powers and Grassroots Image Management: Confucius Institutes and China in the Media.” *Chinese Journal of International Politics* 12(4):557–584.
- Bush, Sarah Sunn and Amaney A. Jamal. 2015. “Anti-Americanism, Authoritarian Politics, and Attitudes about Women’s Representation: Evidence from a Survey Experiment in Jordan.” *International Studies Quarterly* 59(1):34–45.
- Chu, Yun-han, Larry Diamond, Andrew J. Nathan and Doh Chull Shin, eds. 2008. *How East Asians View Democracy*. New York: Columbia University Press.

- Ciorciari, John D. 2010. *The Limits of Alignment: Southeast Asia and the Great Powers since 1975*. Washington, DC: Georgetown University Press.
- Corstange, Daniel and Nikolay Marinov. 2012. "Taking Sides in Other People's Elections: The Polarizing Effect of Foreign Intervention." *American Journal of Political Science* 56(3):655–670.
- Dalton, Russell J. and Nhu-Ngoc T. Ong. 2005. "Authority Orientations and Democratic Attitudes: A Test of the "Asian Values" Hypothesis." *Japanese Journal of Political Science* 6(2):211–231.
- Goldsmith, Benjamin E. and Yusaku Horiuchi. 2009. "Spinning the Globe? U.S. Public Diplomacy and Foreign Public Opinion." *Journal of Politics* 71(3):863–875.
- Goldsmith, Benjamin E., Yusaku Horiuchi and Kelly Matush. 2021. "Does Public Diplomacy Sway Foreign Public Opinion? Identifying the Effect of High-Level Visits." *American Political Science Review* 115(4):1342–1357.
- Goldsmith, Benjamin E., Yusaku Horiuchi and Kelly Matush. 2026. "Does Public Diplomacy Sway Domestic Public Opinion? Presidential Travel Abroad and Approval at Home." *International Studies Quarterly* 70(2):sqag024.
- Gries, Peter Hays. 2004. *China's New Nationalism: Pride, Politics, and Diplomacy*. Berkeley: University of California Press.
- Gruffydd-Jones, Jamie J. 2019. "Citizens and Condemnation: Strategic Uses of International Human Rights Pressure in Authoritarian States." *Comparative Political Studies* 52(4):579–612.
- Hainmueller, Jens, Jonathan Mummolo and Yiqing Xu. 2019. "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice." *Political Analysis* 27(2):163–192.

- Herrmann, Richard K., James F. Voss, Tonya Y. E. Schooler and Joseph Ciarrochi. 1997. "Images in International Relations: An Experimental Test of Cognitive Schemata." *International Studies Quarterly* 41(3):403–433.
- Human Rights Watch. 2023. "Vietnam: Biden Should Raise Rights on Hanoi Visit." News release, September 9. <https://www.hrw.org/news/2023/09/09/vietnam-biden-should-raise-rights-hanoi-visit>.
- Hurwitz, Jon and Mark Peffley. 1987. "How Are Foreign Policy Attitudes Structured? A Hierarchical Model." *American Political Science Review* 81(4):1099–1120.
- Kim, Dae Jung. 1994. "Is Culture Destiny? The Myth of Asia's Anti-Democratic Values." *Foreign Affairs* 73(6):189–194.
- Kuik, Cheng-Chwee. 2008. "The Essence of Hedging: Malaysia and Singapore's Response to a Rising China." *Contemporary Southeast Asia* 30(2):159–185.
- Lin, Joanne, Melinda Martinus, Kristina Fong, Thi Phuong Thao Pham, Indira Zahra Aridati and Sasitharan Gauri. 2026. The State of Southeast Asia: 2026 Survey Report. Technical report ASEAN Studies Centre, ISEAS-Yusof Ishak Institute Singapore: . <https://www.iseas.edu.sg/frontpage-featured/the-state-of-southeast-asia-2026-survey-report/>.
- Mahathir, Mohamad and Shintaro Ishihara. 1995. *The Voice of Asia: Two Leaders Discuss the Coming Century*. Tokyo: Kodansha International.
- Mattingly, Daniel C. and James Sundquist. 2023. "When Does Public Diplomacy Work? Evidence from China's "Wolf Warrior" Diplomats." *Political Science Research and Methods* 11(4):921–929.
- Mattingly, Daniel, Trevor Incerti, Changwook Ju, Colin Moreshead, Seiki Tanaka and Hikaru Yamagishi. 2025. "Chinese State Media Persuades a Global Audience That the "China

- Model” Is Superior: Evidence from a 19-Country Experiment.” *American Journal of Political Science* 69(3):1029–1046.
- Matush, Kelly. 2023. “Harnessing Backlash: How Leaders Can Benefit from Antagonizing Foreign Actors.” *British Journal of Political Science* 53(3):902–918.
- Nye, Joseph S., Jr. 2004. *Soft Power: The Means to Success in World Politics*. New York: PublicAffairs.
- Park, Chong-Min and Doh Chull Shin. 2006. “Do Asian Values Deter Popular Support for Democracy in South Korea?” *Asian Survey* 46(3):341–361.
- Powers, Ryan and Jonathan Renshon. 2023. “International status concerns and domestic support for political leaders.” *American Journal of Political Science* 67(3):732–747.
- Regional Meeting for Asia of the World Conference on Human Rights. 1993. “Final Declaration of the Regional Meeting for Asia of the World Conference on Human Rights (Bangkok Declaration).” UN Doc. A/CONF.157/ASRM/8–A/CONF.157/PC/59.
- Renshon, Jonathan. 2016. “Status Deficits and War.” *International Organization* 70(3):513–550.
- Repnikova, Maria. 2022. *Chinese Soft Power*. Elements in Global China Cambridge: Cambridge University Press.
- Seah, Sharon, Joanne Lin, Melinda Martinus, Kristina Fong, Indira Aridati, Thi Phuong Thao Pham and Damon Chee. 2024. The State of Southeast Asia: 2024 Survey Report. Technical report ASEAN Studies Centre, ISEAS-Yusof Ishak Institute Singapore: <https://www.iseas.edu.sg/wp-content/uploads/2024/03/The-State-of-SEA-2024.pdf>.
- Seah, Sharon, Joanne Lin, Melinda Martinus, Kristina Fong, Thi Phuong Thao Pham and Indira Zahra Aridati. 2025. The State of Southeast Asia: 2025 Survey Report. Technical

- report ASEAN Studies Centre, ISEAS-Yusof Ishak Institute Singapore: . <https://www.iseas.edu.sg/wp-content/uploads/2025/03/The-State-of-SEA-2025-1.pdf>.
- Seah, Sharon, Joanne Lin, Melinda Martinus, Sithanoxay Suvannaphakdy and Thi Phuong Thao Pham. 2023. The State of Southeast Asia: 2023 Survey Report. Technical report ISEAS-Yusof Ishak Institute Singapore: .
- Sen, Amartya. 1997. “Human Rights and Asian Values.” Sixteenth Morgenthau Memorial Lecture on Ethics and Foreign Policy. New York: Carnegie Council on Ethics and International Affairs.
- Shin, Doh Chull. 2012. *Confucianism and Democratization in East Asia*. New York: Cambridge University Press.
- State Council Information Office of the People’s Republic of China. 2023. A Global Community of Shared Future: China’s Proposals and Actions. Technical report State Council Information Office Beijing: . http://english.scio.gov.cn/whitepapers/2023-09/26/content_116710660.htm.
- Terman, Rochelle. 2023. *The Geopolitics of Shaming: When Human Rights Pressure Works—and When It Backfires*. Princeton, NJ: Princeton University Press.
- Tomz, Michael and Jessica LP Weeks. 2020. “Public opinion and foreign electoral intervention.” *American Political Science Review* 114(3):856–873.
- Welzel, Christian. 2013. *Freedom Rising: Human Empowerment and the Quest for Emancipation*. New York: Cambridge University Press.
- Xi, Jinping. 2023. “Join Hands on the Path Towards Modernization.” Keynote address to the CPC in Dialogue with World Political Parties High-Level Meeting, Beijing, March 15. https://www.mfa.gov.cn/eng/wjb/zzjg_663340/xws_665282/xgxw_665284/202303/t20230317_11043656.html.

- Xu, Weifang. 2025. "Pride and Prejudice: The Dual Effects of "Wolf Warrior Diplomacy" on Domestic and International Audiences." *Journal of Conflict Resolution* 69(5):839–867.
- Zakaria, Fareed. 1994. "Culture Is Destiny: A Conversation with Lee Kuan Yew." *Foreign Affairs* 73(2):109–126.

Supporting Information for:
“Clash of Ideologies:
How Citizens Respond to U.S. and China’s Rhetoric
About Domestic Politics in Southeast Asia”

Contents

A.1	Survey Experimental Details	SI-2
A.2	Descriptive Statistics	SI-6
A.3	Balance Tests	SI-6
A.4	Main Results: Full and Per-Country Estimates	SI-7
A.5	Robustness	SI-14
A.6	Mechanisms	SI-20
A.7	Replication	SI-23
A.8	Ethics & IRB	SI-23
A.9	Pre-registration and Deviations	SI-23

A.1 Survey Experimental Details

A.1.1 Overview

The experiments were embedded in a round of a larger annual online survey project on regional foreign policy attitudes, fielded in December 2023 in Indonesia, Malaysia, the Philippines, and Vietnam. TGM Research recruited respondents from its online panels under quotas for age, gender, and location. Respondents consented, completed demographic and baseline attitudinal modules (including the twelve Asian-values items and the nationalism, human-rights, government-satisfaction, political-knowledge, and news-attention measures used in the paper), and were then randomly routed to the American-rhetoric experiment, the Chinese-rhetoric experiment, or other embedded modules unrelated to this study. No respondent saw both experiments. Qualtrics page timers recorded time spent on each vignette; we use them for the speeder analyses in Section A.5.

A.1.2 Vignettes

Every vignette has two paragraphs. The first paragraph is constant across arms within each experiment and across countries. The second paragraph varies by arm; its opening and closing sentences are held parallel across countries, and the policy content is country-specific (Table 2 of the paper). Bracketed country names substitute for each sample. The texts below are transcribed verbatim from the fielded Qualtrics instruments.¹

American experiment

First paragraph (all arms): “At the US-ASEAN summit, US Vice President Kamala Harris spoke about the US’s commitment to protecting Southeast Asia, especially with the growing influence of China. She talked about the importance of the region and the need for working together to keep it stable and prosperous.”

Control: “Later, a senior American diplomat met with [country]’s representative. The diplomat talked about how important [country] is to the US and agreed with Vice President Harris’ points on teamwork and working towards a stable and prosperous region.”

T1, neutral policy (all countries): “Later, a senior American diplomat met with [country]’s representative. The diplomat criticized [country]’s inability to address too much fishing and pollution in the [country]’s waters. The diplomat urged the [country] government to promote responsible fishing, highlighting this as crucial for protecting the environment.”

T2 and T3, illiberal policies, by country:

Indonesia, T2: “The diplomat criticized Indonesia’s discriminatory policies and societal attitudes towards LGBTQ+ individuals. The diplomat urged the Indonesian government to uphold and protect the rights of all its citizens regardless of sexual orientation or gender identity, highlighting this as crucial for a healthy democracy.”

Indonesia, T3: “The diplomat criticized Indonesia’s blasphemy laws for restricting freedom of speech and religious expression, often leading to discrimination and persecution of minorities. The diplomat urged the Indonesian government to respect religious freedom and pluralism, stressing that such laws are at odds with international norms for human rights and the principles of a democratic and inclusive society.”

Malaysia, T2: “The diplomat criticized Malaysia’s application of the Sedition Act as a tool to curb freedom of expression and dissent. The diplomat urged the Malaysian government to uphold the core values of free speech and open political dialogue, highlighting this as crucial for a healthy democracy.”

Malaysia, T3: “The diplomat criticized Malaysia’s implementation of Sharia law, highlighting how certain aspects of this law, especially those affecting women’s rights and religious freedom, are viewed as being in conflict with international human rights standards. The diplomat urged the Malaysian government to respect individual liberties and equality, that such laws stand in stark contrast to the principles of democracy and universal human rights.”

¹Occasional grammatical errors in the fielded English texts are reproduced verbatim and marked with *[sic]*. The professionally translated local-language versions of the instruments, in which respondents took the survey in the local vernacular, do not carry these errors.

Philippines, T2: “The diplomat criticized the Philippines’ recent crackdown on journalists and media freedom. The diplomat urged the Philippine government to protect press freedom and allow journalists to work without fear, highlighting this as crucial for a healthy democracy.”

Philippines, T3: “The diplomat criticized the Philippines’ violent methods used in the Philippines’ drug war, like excessive force and extrajudicial killings. The diplomat urged the Philippine government to respect human rights and due process, stressing that these harsh tactics go against international human rights norms and the basic principles of justice and law.”

Vietnam, T2: “The diplomat criticized Vietnam’s restrictive internet surveillance measures and censorship. The diplomat urged the Vietnamese government to protect citizens’ fundamental right to digital freedom, emphasizing that such actions are antithetical to the core principles of open and free internet.”

Vietnam, T3: “The diplomat criticized Vietnam’s human rights record, and specifically targeted the suppression of freedom of expression and the persecution of political dissidents. The diplomat urged the Vietnamese government to respect human rights and due process, stressing that these harsh tactics go against international human rights norms and the basic principles of justice and law.”

Chinese experiment

First paragraph (all arms): “At the ASEAN-China summit, Chinese Foreign Minister Wang Yi talked about China’s strong commitment to Southeast Asia. He mentioned the economic partnership between ASEAN and China and how China treats all countries equally, aiming to help ASEAN grow.”

Control: “Later, a senior Chinese diplomat met with [country]’s representative. The diplomat stressed how important [country] is to China and agreed with Wang Yi on the need to work together for the region’s stability and growth.”

T1, neutral policy (all countries): “Later, a senior Chinese diplomat met with [country]’s representative. The diplomat spoke positively about [country]’s efforts in managing its environmental challenges, including issues with fishing and pollution in its waters. The diplomat expressed support for the [country] government’s approach to these issues, emphasizing the importance of each nation addressing environmental concerns in ways that best suit their circumstances and capabilities.”

T2 and T3, illiberal policies, by country:

Indonesia, T2: “The diplomat spoke positively about Indonesia’s efforts in managing societal issues in a way that respects traditional values and cultural practices. The diplomat expressed support for the Indonesian government’s approach to LGBTQ+ issues, and said that Indonesia is taking the right approach in maintaining social harmony and cultural norms.”

Indonesia, T3: “The diplomat spoke positively about Indonesia’s approach to its blasphemy laws. The diplomat expressed support for the Indonesian government’s enforcement of these laws is [*sic*] a suitable way to maintain social harmony and respect cultural norms.”

Malaysia, T2: “The diplomat spoke positively about about [*sic*] Malaysia’s use of the Sedition Act in balancing societal stability and information flow. The diplomat expressed support for the Malaysian government’s approach to these issues, and said that Malaysia’s application of the Sedition Act is an appropriate measure fostering social cohesion and national development.”

Malaysia, T3: “In a subsequent private meeting with Malaysia’s representative at the meeting, a senior Chinese diplomat expressed understanding for Malaysia’s nuanced approach to Sharia law. The diplomat acknowledged Malaysia’s efforts in integrating Sharia law within its legal framework, reflecting the nation’s commitment to its cultural and religious norms. This approach was commended as an example of respecting traditional values while managing a diverse society, aligning with China’s perspective on respecting national sovereignty and cultural practices in governance.”

Philippines, T2: “The diplomat spoke positively about the Philippines’ handling of the media and journalist [*sic*] critical of the government. The diplomat expressed support for the Philippine government’s approach to these issues, and said that the steps the Philippines took against some journalists and media were needed to keep the country peaceful and orderly.”

Philippines, T3: “The diplomat spoke positively about the Philippines’ strong actions in its drug war. The diplomat expressed support for the Philippine government’s approach to these issues, and noted that tough measures are sometimes necessary to keep society safe and orderly, aligning with principles of societal stability valued in many Asian cultures.”

Vietnam, T2: “The diplomat spoke positively about Vietnam’s regulated approach to digital freedom and internet governance. The diplomat expressed support for the Vietnamese government’s approach to these issues, highlighting the congruence of these policies [*sic*] a vision of harmonious societal development through guided internet regulation.”

Vietnam, T3: “The diplomat spoke positively about Vietnam’s pragmatic stance on human rights issues. The diplomat expressed support for the Vietnamese government’s approach to these issues, praising the alignment of Vietnam’s strategies with the principles of stability and collective societal goals.”

A.1.3 Vignette Case Selection

In each country we tailored the two illiberal-policy arms to controversies that were prominent and values-laden at the time of fielding, so that the rhetoric would refer to a real object of domestic contention rather than a hypothetical one. Within each country, the pair was chosen with an ordering in mind: policy A is the comparatively less severe of the two and policy B the more severe, in that country’s own context. This ordering is a design-intent claim rather than a registered one (see Section A.9), and readers who would rank a given pair differently can rely on the per-country estimates, where each policy stands on its own. The neutral-policy arm addresses overfishing and marine pollution in every country, a matter of technical governance that is salient in all four but not a values controversy in any of them. Within each experiment the arms hold the diplomatic setting, the meeting format, and the sentence structure parallel, so that arms differ only in the policy content addressed.

A.1.4 Outcome Question Wordings

Following the vignette, respondents answered, in order: (1) “Based on the previous vignette, please rate your overall opinion of the senior [American/Chinese] diplomat. Use a scale from 0 to 10, where 0 indicates an extremely unfavorable opinion and 10 indicates an extremely favorable opinion.” (2) “Please rate your overall opinion of [the United States/China] on a scale from 0 to 10.” (3) The attribution battery: “For each of the following factors that may have influenced your rating, please indicate your stance: ‘Displeased,’ ‘Unaffected,’ or ‘Pleased.’” The four factors, adapted per experiment, cover the power’s criticism or approval of the country’s adherence to values standards, its role in regional security and stability, its role as an economic and trading partner, and its stance toward how the country runs its domestic politics. (4) “If [country] were forced to align itself with one of the two strategic rivals, which should it choose? Please indicate your preference on the scale below, where the far left represents a complete alignment with China, the far right represents a complete alignment with the United States, and the center represents a neutral stance.”

A.1.5 Pre-Treatment Measures

All items below were asked before the experimental module, on 0–10 scales unless noted.

Asian-values battery

Respondents rated their agreement with twelve statements, each scored from 0 (strongly disagree) to 10 (strongly agree). Six general statements pair each dimension of the Asian-values canon against its liberal-democratic rival:

1. *Collectivism:* “The collective interests of the community should be prioritized, even if it means limiting individual rights and freedoms.”
2. *Hierarchy:* “Maintaining established traditions and hierarchical social orders is important, even if it conflicts with the promotion of equality.”
3. *Social harmony:* “Preserving social harmony and stability should be prioritized, even if it restricts freedom of speech and expression.”
4. *Family obligation:* “Adhering to family expectations and societal norms is essential, even if

it compromises individual autonomy and personal life choices.”

5. *Traditional governance*: “Upholding traditional governance is crucial, even if it occasionally overshadows strict adherence to the rule of law.”
6. *Authoritarian leadership*: “Valuing authoritarian leadership models for efficiency and stability is important, even if it means deviating from democratic principles and processes.”

Six specific statements ground the same dimensions in everyday scenarios:

1. *Collectivism*: “Supporting community projects and initiatives should be a priority, even if it means less time and resources for personal hobbies or interests.”
2. *Hierarchy*: “Showing respect and obedience to elders and authority figures is crucial, even if their views or decisions are different from one’s own beliefs.”
3. *Social harmony*: “Avoiding controversial topics in conversations to maintain social peace is important, even if it means not expressing one’s true opinions on important issues.”
4. *Family obligation*: “Choosing a career that aligns with family expectations is essential, even if it differs from one’s true passion or professional aspirations.”
5. *Traditional governance*: “Adhering to long-standing cultural practices and customs should be respected, even when they are in conflict with new legal standards or practices.”
6. *Authoritarian leadership*: “Supporting policies that ensure national stability and order is important, even if they may infringe upon certain human rights or personal liberties.”

The overall index averages the twelve items. Dimension subindices average each dimension’s general and specific item.

Nationalism

Nationalism averages two items. The first is identical in all four countries: “My nationality is at the core of my identity.” The second is country-specific: in Indonesia, “I support wearing batik every Friday to celebrate Indonesia’s cultural heritage”; in Malaysia, “The interests of Malaysia come before those of other countries”; in the Philippines, “If I have the opportunity to change citizenship, I would still choose to be a Filipino citizen”; in Vietnam, “I feel proud when listening to the National anthem.” Because the second item varies in content across countries, cross-country comparisons of nationalism levels should be made with caution; our analyses interact nationalism with treatment within pooled models that hold country constant.

Other measures

Universal human rights: “Do you agree that human rights are universal rights?” *Government satisfaction*: “Overall, how satisfied or dissatisfied are you with the national government?” *Perceived military strength*: “How strong is [country] from a military perspective?” (in Indonesia, “from an economic and military perspective”). *News attention*: attention to “News about politics in your country” and “News about politics outside your country.” *Political knowledge*: two factual questions, “In 2016, a Southeast Asian country sued China over their ‘nine dash line’ in the South China Sea. Do you recall who it was?” (correct: the Philippines) and “Which country in Southeast Asia had a military coup d’état in February 2021 that deposed the democratically elected government?” (correct: Myanmar); knowledge counts correct answers (0–2).

A.2 Descriptive Statistics

Table A1: Sample descriptives by country

	Indonesia	Malaysia	Philippines	Vietnam
Age	34.71 (10.80)	34.50 (10.57)	37.15 (12.79)	34.53 (10.93)
Female	0.48 (0.50)	0.49 (0.50)	0.49 (0.50)	0.46 (0.50)
Nationalism (0–10)	8.55 (1.64)	8.46 (1.80)	8.44 (1.89)	8.89 (1.60)
Perceived own-country military strength (0–10)	7.18 (1.85)	6.78 (2.30)	5.60 (2.21)	7.57 (1.95)
Human-rights advocacy (0–10)	4.88 (3.59)	8.56 (1.92)	8.73 (1.95)	9.21 (1.42)
Asian-values index (0–10)	6.40 (1.79)	6.71 (1.87)	6.45 (2.07)	6.91 (2.08)
N (full survey)	1,507	1,408	1,267	1,567
N (US experiment)	679	634	570	699
N (China experiment)	681	636	569	700

Note: Cell entries are means with standard deviations in parentheses. All attitudinal measures run 0–10.

Table A2: Asian-values index: means, subindices, and reliability

	Indonesia	Malaysia	Philippines	Vietnam	Pooled
Overall index	6.40	6.71	6.45	6.91	6.62
Collectivism	7.10	6.80	6.66	7.56	7.05
Hierarchy	6.57	6.85	6.79	6.91	6.78
Social harmony	6.58	6.89	6.31	7.17	6.76
Family obligations	6.24	6.45	6.07	6.59	6.35
Traditional governance	6.46	6.64	6.36	6.77	6.57
Authoritarian leadership	5.44	6.55	6.50	6.34	6.19
Cronbach’s α (12 items)	0.90	0.93	0.93	0.94	0.93

Note: Cell entries are means of the overall index and the six dimension subindices, by country and pooled. The final row reports Cronbach’s alpha over the twelve items.

A.3 Balance Tests

Within each experiment, we test whether pre-treatment covariates are balanced across the four arms. For each covariate we report arm means and the p -value from an F-test of joint equality; each table ends with an omnibus multinomial-logit test of all covariates jointly predicting assignment. Across the forty-eight covariate-by-arm tests in the two tables below, four reject equality at the five percent level, slightly more than the 2.4 that chance predicts. Per-country balance tables are available in the replication package.

Table A3: Balance, US experiment (pooled)

	Control	T1	T2	T3	p (joint)
Age	35.13	35.89	35.22	35.90	0.444
Female	0.47	0.47	0.46	0.46	0.956
Nationalism	8.56	8.54	8.60	8.66	0.640
Perceived military strength	6.74	6.74	6.90	6.93	0.256
Human-rights advocacy	7.84	7.66	7.80	7.83	0.679
Asian-values index	6.63	6.44	6.63	6.64	0.206
N	647	649	663	623	
Omnibus mlogit LR test: $\chi^2 = 12.46$, $p = 0.823$					

Note: Arm means with the p -value from an F-test of equal means across arms; final row reports the omnibus multinomial-logit test.

Table A4: Balance, China experiment (pooled)

	Control	T1	T2	T3	p (joint)
Age	35.09	35.84	35.61	36.09	0.418
Female	0.48	0.48	0.46	0.49	0.690
Nationalism	8.66	8.59	8.65	8.62	0.907
Perceived military strength	6.84	6.71	6.80	7.01	0.106
Human-rights advocacy	7.91	7.93	7.82	7.60	0.203
Asian-values index	6.64	6.54	6.73	6.64	0.430
N	648	628	677	633	
Omnibus mlogit LR test: $\chi^2 = 21.88$, $p = 0.237$					

Note: See note to Table A3.

A.4 Main Results: Full and Per-Country Estimates

This section reports the country-level counterparts of the paper's pooled estimates (Tables 4 and 5 of the paper), the full moderation results for the registered hypotheses H2 and H4, and the cross-country heterogeneity tests summarized in the paper's Figure 3. Per-country versions of the moderation estimates are available in the replication package.

Table A5: American criticism: Indonesia

	Opinion of US diplomat		Opinion of United States		Choose sides (1=China...7=US)	
	(1)	(2)	(3)	(4)	(5)	(6)
T1: Neutral policy	-0.083 (0.252)	-0.117 (0.263)	-0.125 (0.267)	-0.187 (0.271)	-0.346** (0.134)	-0.365** (0.139)
T2: Illiberal policy A	-2.543** (0.298)	-2.525** (0.308)	-1.693** (0.302)	-1.694** (0.311)	-0.255 (0.137)	-0.228 (0.146)
T3: Illiberal policy B	-0.973** (0.294)	-1.229** (0.307)	-0.860** (0.299)	-1.119** (0.306)	-0.287* (0.133)	-0.318* (0.139)
Controls	No	Yes	No	Yes	No	Yes
Control-group mean	6.69	6.69	6.27	6.27	4.82	4.82
p: T1=T2	0.000	0.000	0.000	0.000	0.500	0.331
p: T1=T3	0.002	0.000	0.012	0.002	0.655	0.723
p: T1=T2=T3	0.000	0.000	0.000	0.000	0.790	0.619
Observations	670	670	667	667	674	674
Adj. R ²	0.124	0.136	0.056	0.076	0.007	-0.005

Note: OLS, arm indicators against the within-country control group, robust standard errors. Outcomes as in Table 4 of the paper. * $p < 0.05$, ** $p < 0.01$.

Table A6: American criticism: Malaysia

	Opinion of US diplomat		Opinion of United States		Choose sides (1=China...7=US)	
	(1)	(2)	(3)	(4)	(5)	(6)
T1: Neutral policy	0.185 (0.294)	0.200 (0.308)	0.160 (0.297)	0.196 (0.310)	0.069 (0.150)	0.097 (0.156)
T2: Illiberal policy A	-0.484 (0.305)	-0.367 (0.315)	-0.264 (0.314)	-0.142 (0.323)	0.013 (0.156)	0.072 (0.162)
T3: Illiberal policy B	-0.796* (0.329)	-0.758* (0.332)	-0.314 (0.326)	-0.264 (0.329)	-0.062 (0.146)	-0.049 (0.150)
Controls	No	Yes	No	Yes	No	Yes
Control-group mean	5.83	5.83	5.22	5.22	4.35	4.35
p: T1=T2	0.031	0.074	0.189	0.306	0.720	0.877
p: T1=T3	0.003	0.004	0.156	0.169	0.369	0.329
p: T1=T2=T3	0.008	0.013	0.271	0.348	0.664	0.574
Observations	626	626	621	621	618	618
Adj. R ²	0.014	0.045	-0.000	0.028	-0.004	0.018

Note: See note to Table A5.

Table A7: American criticism: Philippines

	Opinion of US diplomat		Opinion of United States		Choose sides (1=China...7=US)	
	(1)	(2)	(3)	(4)	(5)	(6)
T1: Neutral policy	-0.308 (0.255)	-0.263 (0.263)	-0.030 (0.239)	-0.001 (0.242)	-0.024 (0.158)	-0.014 (0.165)
T2: Illiberal policy A	-0.716** (0.251)	-0.743** (0.259)	-0.636** (0.233)	-0.661** (0.238)	0.022 (0.151)	0.047 (0.154)
T3: Illiberal policy B	-1.062** (0.295)	-0.993** (0.297)	-0.255 (0.242)	-0.263 (0.249)	-0.027 (0.161)	0.020 (0.165)
Controls	No	Yes	No	Yes	No	Yes
Control-group mean	7.15	7.15	7.16	7.16	5.31	5.30
p: T1=T2	0.147	0.096	0.020	0.012	0.780	0.717
p: T1=T3	0.019	0.027	0.400	0.333	0.984	0.847
p: T1=T2=T3	0.058	0.067	0.061	0.039	0.945	0.936
Observations	564	563	560	560	522	521
Adj. R ²	0.022	0.027	0.010	0.039	-0.006	0.025

Note: See note to Table A5.

Table A8: American criticism: Vietnam

	Opinion of US diplomat		Opinion of United States		Choose sides (1=China...7=US)	
	(1)	(2)	(3)	(4)	(5)	(6)
T1: Neutral policy	-0.562* (0.251)	-0.599* (0.258)	-0.505* (0.251)	-0.543* (0.250)	-0.073 (0.139)	-0.041 (0.139)
T2: Illiberal policy A	-0.586* (0.254)	-0.640* (0.257)	-0.264 (0.245)	-0.275 (0.245)	0.082 (0.132)	0.105 (0.134)
T3: Illiberal policy B	-1.048** (0.286)	-1.057** (0.290)	-0.596* (0.270)	-0.594* (0.272)	0.057 (0.131)	0.058 (0.132)
Controls	No	Yes	No	Yes	No	Yes
Control-group mean	7.29	7.29	7.15	7.15	4.92	4.92
p: T1=T2	0.931	0.880	0.355	0.301	0.237	0.256
p: T1=T3	0.105	0.122	0.749	0.857	0.317	0.438
p: T1=T2=T3	0.211	0.256	0.448	0.438	0.457	0.515
Observations	691	690	681	680	646	645
Adj. R ²	0.016	0.039	0.004	0.025	-0.002	0.019

Note: See note to Table A5.

Table A9: Chinese endorsement: Indonesia

	Opinion of Chinese diplomat		Opinion of China		Choose sides (1=China...7=US)	
	(1)	(2)	(3)	(4)	(5)	(6)
T1: Neutral policy	0.435 (0.250)	0.622* (0.246)	0.425 (0.248)	0.631* (0.247)	0.108 (0.138)	-0.008 (0.134)
T2: Illiberal policy A	0.371 (0.255)	0.472 (0.259)	0.359 (0.240)	0.450 (0.247)	0.006 (0.139)	-0.010 (0.139)
T3: Illiberal policy B	0.644* (0.253)	0.722** (0.253)	0.311 (0.247)	0.385 (0.244)	0.052 (0.137)	-0.024 (0.133)
Controls	No	Yes	No	Yes	No	Yes
Control-group mean	6.36	6.36	6.13	6.13	4.50	4.50
p: T1=T2	0.795	0.547	0.777	0.449	0.475	0.992
p: T1=T3	0.386	0.681	0.637	0.316	0.694	0.914
p: T1=T2=T3	0.509	0.622	0.894	0.577	0.774	0.993
Observations	672	671	664	663	675	674
Adj. R ²	0.006	0.033	0.001	0.035	-0.003	0.050

Note: See note to Table A5.

Table A10: Chinese endorsement: Malaysia

	Opinion of Chinese diplomat		Opinion of China		Choose sides (1=China...7=US)	
	(1)	(2)	(3)	(4)	(5)	(6)
T1: Neutral policy	0.069 (0.258)	0.163 (0.262)	-0.289 (0.281)	-0.133 (0.277)	0.248 (0.140)	0.166 (0.139)
T2: Illiberal policy A	-0.037 (0.267)	-0.035 (0.265)	-0.126 (0.280)	-0.078 (0.278)	0.036 (0.148)	0.027 (0.147)
T3: Illiberal policy B	-0.065 (0.271)	0.088 (0.274)	0.120 (0.284)	0.316 (0.284)	0.062 (0.149)	0.033 (0.151)
Controls	No	Yes	No	Yes	No	Yes
Control-group mean	6.44	6.44	6.06	6.06	4.08	4.08
p: T1=T2	0.691	0.450	0.574	0.842	0.144	0.333
p: T1=T3	0.620	0.778	0.163	0.109	0.201	0.354
p: T1=T2=T3	0.868	0.750	0.374	0.229	0.268	0.536
Observations	629	629	625	625	619	619
Adj. R ²	-0.004	0.041	-0.001	0.069	0.001	0.028

Note: See note to Table A5.

Table A11: Chinese endorsement: Philippines

	Opinion of Chinese diplomat		Opinion of China		Choose sides (1=China...7=US)	
	(1)	(2)	(3)	(4)	(5)	(6)
T1: Neutral policy	0.254 (0.304)	0.299 (0.295)	-0.457 (0.365)	-0.424 (0.369)	0.143 (0.153)	0.146 (0.160)
T2: Illiberal policy A	-0.410 (0.296)	-0.346 (0.288)	-0.576 (0.341)	-0.467 (0.344)	-0.033 (0.152)	-0.098 (0.154)
T3: Illiberal policy B	-0.371 (0.303)	-0.245 (0.306)	-0.256 (0.363)	-0.115 (0.369)	0.017 (0.156)	-0.005 (0.160)
Controls	No	Yes	No	Yes	No	Yes
Control-group mean	6.04	6.04	5.26	5.26	5.37	5.37
p: T1=T2	0.022	0.018	0.723	0.901	0.243	0.113
p: T1=T3	0.035	0.066	0.573	0.401	0.417	0.338
p: T1=T2=T3	0.040	0.045	0.628	0.552	0.488	0.280
Observations	563	563	558	558	518	518
Adj. R ²	0.007	0.074	0.000	0.023	-0.003	0.007

Note: See note to Table A5.

Table A12: Chinese endorsement: Vietnam

	Opinion of Chinese diplomat		Opinion of China		Choose sides (1=China...7=US)	
	(1)	(2)	(3)	(4)	(5)	(6)
T1: Neutral policy	0.465 (0.301)	0.465 (0.292)	0.503 (0.306)	0.466 (0.299)	-0.026 (0.142)	-0.040 (0.141)
T2: Illiberal policy A	0.034 (0.307)	0.105 (0.309)	-0.029 (0.301)	-0.059 (0.303)	-0.038 (0.141)	-0.067 (0.140)
T3: Illiberal policy B	-0.213 (0.308)	-0.250 (0.299)	0.054 (0.303)	-0.051 (0.295)	0.156 (0.139)	0.109 (0.139)
Controls	No	Yes	No	Yes	No	Yes
Control-group mean	5.61	5.61	4.97	4.97	5.02	5.02
p: T1=T2	0.145	0.226	0.078	0.082	0.930	0.847
p: T1=T3	0.022	0.014	0.138	0.083	0.202	0.300
p: T1=T2=T3	0.068	0.050	0.168	0.135	0.304	0.414
Observations	693	692	679	678	634	633
Adj. R ²	0.003	0.049	0.002	0.039	-0.001	0.032

Note: See note to Table A5.

Table A13: Moderation by the Asian-values index, US experiment (pooled)

	(1)	(2)	(3)	(4)	(5)	(6)
T1: Neutral policy	0.404 (0.533)	0.456 (0.555)	0.120 (0.531)	0.119 (0.544)	0.108 (0.283)	0.181 (0.296)
T2: Illiberal policy A	-1.836** (0.570)	-1.836** (0.580)	-1.191* (0.541)	-1.192* (0.553)	0.280 (0.283)	0.314 (0.293)
T3: Illiberal policy B	-1.369* (0.588)	-1.498* (0.588)	-1.282* (0.572)	-1.357* (0.570)	0.020 (0.275)	0.022 (0.280)
Asian-values index	0.318** (0.054)	0.330** (0.056)	0.310** (0.055)	0.315** (0.057)	0.116** (0.030)	0.121** (0.032)
T1 × Asian values	-0.078 (0.078)	-0.086 (0.081)	-0.029 (0.078)	-0.029 (0.080)	-0.031 (0.043)	-0.039 (0.044)
T2 × Asian values	0.110 (0.082)	0.114 (0.083)	0.068 (0.079)	0.073 (0.081)	-0.049 (0.043)	-0.048 (0.044)
T3 × Asian values	0.060 (0.085)	0.072 (0.086)	0.111 (0.083)	0.115 (0.083)	-0.015 (0.042)	-0.013 (0.042)
Controls	No	Yes	No	Yes	No	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,478	2,477	2,463	2,463	2,394	2,393
Adj. R ²	0.125	0.146	0.155	0.173	0.088	0.093

Note: OLS estimates of Equation 2 with the continuous Asian-values index, pooled with country fixed effects. Marginal effects across the index are plotted in Figure 4 of the paper. * $p < 0.05$, ** $p < 0.01$.

Table A14: Moderation by the Asian-values index, China experiment (pooled)

	(1)	(2)	(3)	(4)	(5)	(6)
T1: Neutral policy	0.265 (0.541)	0.391 (0.536)	0.260 (0.542)	0.465 (0.535)	-0.035 (0.277)	-0.202 (0.270)
T2: Illiberal policy A	0.723 (0.584)	0.741 (0.582)	0.715 (0.578)	0.762 (0.582)	-0.192 (0.283)	-0.232 (0.286)
T3: Illiberal policy B	0.796 (0.568)	0.895 (0.563)	0.718 (0.563)	0.886 (0.568)	-0.263 (0.277)	-0.384 (0.275)
Asian-values index	0.447** (0.058)	0.431** (0.059)	0.515** (0.059)	0.514** (0.059)	0.062* (0.030)	0.053 (0.030)
T1 $\times$ Asian values	0.014 (0.078)	0.005 (0.077)	-0.021 (0.080)	-0.042 (0.079)	0.024 (0.043)	0.043 (0.041)
T2 $\times$ Asian values	-0.111 (0.084)	-0.105 (0.084)	-0.122 (0.084)	-0.121 (0.085)	0.027 (0.043)	0.031 (0.044)
T3 $\times$ Asian values	-0.119 (0.084)	-0.121 (0.083)	-0.098 (0.084)	-0.111 (0.085)	0.050 (0.042)	0.064 (0.042)
Controls	No	Yes	No	Yes	No	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,512	2,511	2,486	2,485	2,404	2,403
Adj. R ²	0.119	0.146	0.160	0.183	0.132	0.153

Note: See note to Table A13.

Table A15: Treatment effects interacted with country, US experiment

	(1) Diplomat	(2) Power
T1: Neutral policy	-0.083 (0.252)	-0.125 (0.267)
T2: Illiberal policy A	-2.543** (0.299)	-1.693** (0.302)
T3: Illiberal policy B	-0.973** (0.294)	-0.860** (0.299)
Control $\times$ Indonesia	0.000 (.)	0.000 (.)
Control $\times$ Malaysia	0.000 (.)	0.000 (.)
Control $\times$ Philippines	0.000 (.)	0.000 (.)
Control $\times$ Vietnam	0.000 (.)	0.000 (.)
T1: Neutral policy $\times$ Indonesia	0.000 (.)	0.000 (.)
T1: Neutral policy $\times$ Malaysia	0.267 (0.387)	0.285 (0.399)
T1: Neutral policy $\times$ Philippines	-0.226 (0.359)	0.095 (0.358)
T1: Neutral policy $\times$ Vietnam	-0.480 (0.356)	-0.380 (0.366)
T2: Illiberal policy A $\times$ Indonesia	0.000 (.)	0.000 (.)
T2: Illiberal policy A $\times$ Malaysia	2.059** (0.427)	1.430** (0.435)
T2: Illiberal policy A $\times$ Philippines	1.827** (0.390)	1.057** (0.381)
T2: Illiberal policy A $\times$ Vietnam	1.957** (0.392)	1.429** (0.389)
T3: Illiberal policy B $\times$ Indonesia	0.000 (.)	0.000 (.)
T3: Illiberal policy B $\times$ Malaysia	0.177 (0.441)	0.546 (0.442)
T3: Illiberal policy B $\times$ Philippines	-0.089 (0.417)	0.605 (0.385)
T3: Illiberal policy B $\times$ Vietnam	-0.074 (0.410)	0.265 (0.403)
p: effects equal across countries	0.000	0.001
Observations	2,551	2,529
Adj. R ²	0.081	0.098

Note: Arm effects (Indonesia is the reference country) and arm-by-country interactions. The homogeneity row reports the joint test that all interactions are zero. * $p < 0.05$, ** $p < 0.01$.

Table A16: Treatment effects interacted with country, China experiment

	(1) Diplomat	(2) Power
T1: Neutral policy	0.435 (0.250)	0.425 (0.248)
T2: Illiberal policy A	0.371 (0.255)	0.359 (0.240)
T3: Illiberal policy B	0.644* (0.253)	0.311 (0.247)
Control $\times$ Indonesia	0.000 (.)	0.000 (.)
Control $\times$ Malaysia	0.000 (.)	0.000 (.)
Control $\times$ Philippines	0.000 (.)	0.000 (.)
Control $\times$ Vietnam	0.000 (.)	0.000 (.)
T1: Neutral policy $\times$ Indonesia	0.000 (.)	0.000 (.)
T1: Neutral policy $\times$ Malaysia	-0.366 (0.359)	-0.715 (0.375)
T1: Neutral policy $\times$ Philippines	-0.181 (0.393)	-0.883* (0.441)
T1: Neutral policy $\times$ Vietnam	0.030 (0.391)	0.078 (0.394)
T2: Illiberal policy A $\times$ Indonesia	0.000 (.)	0.000 (.)
T2: Illiberal policy A $\times$ Malaysia	-0.409 (0.369)	-0.485 (0.369)
T2: Illiberal policy A $\times$ Philippines	-0.781* (0.391)	-0.934* (0.417)
T2: Illiberal policy A $\times$ Vietnam	-0.337 (0.399)	-0.387 (0.385)
T3: Illiberal policy B $\times$ Indonesia	0.000 (.)	0.000 (.)
T3: Illiberal policy B $\times$ Malaysia	-0.709 (0.370)	-0.191 (0.376)
T3: Illiberal policy B $\times$ Philippines	-1.014* (0.395)	-0.566 (0.439)
T3: Illiberal policy B $\times$ Vietnam	-0.856* (0.398)	-0.257 (0.391)
p: effects equal across countries	0.153	0.241
Observations	2,557	2,526
Adj. R ²	0.029	0.051

Note: See note to Table A15.

A.5 Robustness

A.5.1 Alternative codings of the Asian-values index

The twelve Asian-values items are all positively keyed, so the raw index mixes substantive affinity with acquiescent response style. We re-estimate the H2 and H4 moderation under four alternative constructions: the index net of the respondent's human-rights commitment; an ipsatized index that subtracts the respondent's mean across all 0–10 agree-keyed rating items; a political subindex (authoritarian leadership and traditional governance only); and a within-country standardized index. No coding produces the registered negative (H2) or positive (H4) gradients; under the style-corrected codings the wrongly signed interactions in the US experiment grow slightly.

Table A17: Alternative codings of the moderator, US experiment

	(1) Raw	(2) Net of HR	(3) Ipsatized	(4) Political	(5) Within-z	(6) Raw	(7) Net of HR	(8) Ipsatized	(9) Political	(10) Within-z
T1: Neutral policy	0.404 (0.533)	-0.185 (0.143)	-0.212 (0.155)	0.143 (0.451)	-0.109 (0.131)	0.120 (0.531)	-0.144 (0.147)	-0.128 (0.159)	-0.029 (0.453)	-0.069 (0.131)
T2: Illiberal policy A	-1.836** (0.570)	-1.176** (0.156)	-1.036** (0.163)	-2.013** (0.497)	-1.106** (0.139)	-1.191* (0.541)	-0.833** (0.155)	-0.746** (0.164)	-1.230** (0.469)	-0.741** (0.135)
T3: Illiberal policy B	-1.369* (0.588)	-0.948** (0.164)	-0.783** (0.175)	-1.747** (0.502)	-0.969** (0.148)	-1.282* (0.572)	-0.527** (0.158)	-0.332* (0.168)	-1.570** (0.478)	-0.544** (0.139)
T1: Neutral policy × Asian values (raw)	-0.078 (0.078)					-0.029 (0.078)				
T2: Illiberal policy A × Asian values (raw)	0.110 (0.082)					0.068 (0.079)				
T3: Illiberal policy B × Asian values (raw)	0.060 (0.085)					0.111 (0.083)				
T1: Neutral policy × Net of human rights		-0.039 (0.043)					-0.030 (0.044)			
T2: Illiberal policy A × Net of human rights		-0.062 (0.047)					-0.079 (0.047)			
T3: Illiberal policy B × Net of human rights		0.022 (0.050)					0.015 (0.049)			
T1: Neutral policy × Ipsatized			-0.322 (0.377)					-0.089 (0.377)		
T2: Illiberal policy A × Ipsatized			0.234 (0.385)					-0.076 (0.371)		
T3: Illiberal policy B × Ipsatized			0.779 (0.424)					0.883* (0.412)		
T1: Neutral policy × Political subindex				-0.043 (0.067)					-0.012 (0.068)	
T2: Illiberal policy A × Political subindex				0.138 (0.073)					0.071 (0.070)	
T3: Illiberal policy B × Political subindex				0.121 (0.075)					0.158* (0.071)	
T1: Neutral policy × Within-country z					-0.133 (0.154)					-0.037 (0.155)
T2: Illiberal policy A × Within-country z					0.162 (0.165)					0.086 (0.159)
T3: Illiberal policy B × Within-country z					0.127 (0.169)					0.213 (0.166)
Observations	2,478	2,471	2,478	2,499	2,478	2,463	2,456	2,463	2,484	2,463
Adj. R ²	0.125	0.077	0.090	0.125	0.123	0.155	0.105	0.116	0.154	0.154

Note: Columns 1–5: diplomat rating; columns 6–10: country rating. Each column interacts the arms with one coding of the Asian-values measure. * $p < 0.05$, ** $p < 0.01$.

A.5.2 Construct validity of the Asian-values battery

The searches in the previous subsection ask whether the moderation null survives alternative specifications. A more basic objection is possible: that the battery measures nothing in particular, in which case a null involving it would be uninformative. This subsection addresses that objection directly by treating the battery as an outcome rather than a moderator, and asking whether it behaves the way a valid measure of the construct should (Adcock and Collier, 2001). Table A19

Table A18: Alternative codings of the moderator, China experiment

	(1) Raw	(2) Net of HR	(3) Ipsatized	(4) Political	(5) Within-z	(6) Raw	(7) Net of HR	(8) Ipsatized	(9) Political	(10) Within-z
T1: Neutral policy	0.265 (0.541)	0.282* (0.144)	0.343* (0.159)	0.032 (0.470)	0.365** (0.131)	0.260 (0.542)	0.018 (0.155)	0.041 (0.175)	-0.034 (0.466)	0.123 (0.140)
T2: Illiberal policy A	0.723 (0.584)	0.026 (0.143)	-0.057 (0.159)	0.397 (0.510)	-0.012 (0.137)	0.715 (0.578)	-0.046 (0.145)	-0.033 (0.164)	0.369 (0.498)	-0.095 (0.138)
T3: Illiberal policy B	0.796 (0.568)	-0.038 (0.145)	-0.051 (0.163)	0.475 (0.497)	0.011 (0.138)	0.718 (0.563)	-0.026 (0.150)	0.007 (0.171)	0.234 (0.474)	0.069 (0.142)
T1: Neutral policy $\times$ Asian values (raw)	0.014 (0.078)					-0.021 (0.080)				
T2: Illiberal policy A $\times$ Asian values (raw)	-0.111 (0.084)					-0.122 (0.084)				
T3: Illiberal policy B $\times$ Asian values (raw)	-0.119 (0.084)					-0.098 (0.084)				
T1: Neutral policy $\times$ Net of human rights		-0.020 (0.044)					-0.032 (0.045)			
T2: Illiberal policy A $\times$ Net of human rights		0.022 (0.045)					0.045 (0.043)			
T3: Illiberal policy B $\times$ Net of human rights		0.013 (0.042)					-0.017 (0.042)			
T1: Neutral policy $\times$ Ipsatized			0.134 (0.378)					-0.113 (0.400)		
T2: Illiberal policy A $\times$ Ipsatized			-0.254 (0.406)					0.269 (0.410)		
T3: Illiberal policy B $\times$ Ipsatized			-0.168 (0.380)					-0.146 (0.392)		
T1: Neutral policy $\times$ Political subindex				0.046 (0.069)					0.017 (0.071)	
T2: Illiberal policy A $\times$ Political subindex				-0.070 (0.075)					-0.081 (0.074)	
T3: Illiberal policy B $\times$ Political subindex				-0.077 (0.075)					-0.032 (0.073)	
T1: Neutral policy $\times$ Within-country z					0.043 (0.150)					-0.033 (0.155)
T2: Illiberal policy A $\times$ Within-country z					-0.213 (0.163)					-0.229 (0.164)
T3: Illiberal policy B $\times$ Within-country z					-0.205 (0.162)					-0.185 (0.163)
Observations	2,512	2,502	2,512	2,527	2,512	2,486	2,479	2,486	2,501	2,486
Adj. R ²	0.119	0.066	0.050	0.105	0.118	0.160	0.102	0.090	0.148	0.159

Note: See note to Table A17.

collects the results.

The items are unidimensional. The first eigenvalue is 6.70 and the second is 0.78, so a single factor accounts for 56 percent of the variance in twelve items; item-total correlations run from 0.68 to 0.79, with no item detached from the rest; and reliability is high in every country (Table A2). Whatever the battery is measuring, respondents answer its items as though they belong together.

What it is measuring separates the groups that the Asian-values debate is itself about. The sharpest available contrast is between Malay and Chinese Malaysians, the two communities whose claims animated the original argument, and the gap is 3.4 points on the eleven-point index, larger than one and a half standard deviations. Religious affiliation moves the index in the predicted direction in all four countries, significantly in three: Muslims score higher than non-Muslims in Malaysia, religiously affiliated Vietnamese higher than the unaffiliated, and Catholics higher than other Filipinos. In Vietnam, the one country with substantial variation in affiliation, the gap appears on all six dimensions and is largest on family obligation, traditional governance, and authoritarian leadership, which is where a religious tradition should bite hardest.

The socio-demographic correlates also run the way the literature on value formation expects. Net of country, education is negatively associated with the index, and the association is strongest for family obligation, the dimension most directly about deference to social expectation. Age shows only a small negative association (-0.07 per within-country standard deviation), which departs from the positive cohort pattern reported in cross-national work (Welzel, 2013); we attribute this to our sampling frame, since older respondents reachable through online panels in these countries are a comparatively urban and educated group, and we do not read this weak association as informative about cohort change.

Finally, the index correlates *differentially* with other agree-keyed items, which is what a substantive construct produces and what an acquiescent response style does not. If respondents were simply agreeing with whatever they were shown, the index would correlate uniformly with every other zero-to-ten agreement item in the survey. Instead, it correlates at 0.40 with perceived state strength and 0.39 with satisfaction with government, both conceptually adjacent, at 0.32 with nationalism, but at only 0.14 with the proposition that human rights are universal, and negatively with factual political knowledge. Some acquiescence is surely present, which is why we also report the style-corrected codings above. But a measure that distinguishes among agree-keyed items in a theoretically ordered way is measuring something, and the null it produces in the moderation tests is therefore informative.

A.5.3 Searching for heterogeneity by Asian values

The registered moderation tests find no gradient in the Asian-values index, and a reasonable objection is that the index, or the linear-interaction specification, is simply too blunt to detect resonance that is really present. To take that objection seriously, we searched for the missing heterogeneity in every place the resonance account says it should live. Table A20 collects the results. Throughout, illiberal rhetoric pools the T2 and T3 arms against the control group, and the outcome is the diplomat rating.

The first search is the most demanding version of the resonance hypothesis. Rather than interacting the treatment with a whole-sample index, we pair each vignette with the value dimension it directly implicates and interact the arm effect with that dimension alone: Indonesia's criticism of LGBTQ+ discrimination with family obligation, its blasphemy-law criticism and the three speech-restricting policies with social harmony, Malaysia's Sharia criticism with traditional governance, and the Philippine drug war and Vietnamese persecution of dissidents with authoritarian leadership. If citizens punish criticism because it collides with a value they hold, this is where the collision should register. It does not.

Table A19: Construct validity of the Asian-values battery

Check	Quantity	Value
Structure	first eigenvalue	6.70
	second eigenvalue	0.78
	variance explained, first factor	56%
	item-total correlations	0.68–0.79
	Cronbach’s α (pooled)	0.93
Known groups	Malaysia: Malay vs. Chinese	+3.37**
	Malaysia: Muslim vs. non-Muslim	+0.69**
	Vietnam: religious vs. unaffiliated	+0.41**
	Philippines: Catholic vs. other	+0.29*
	Indonesia: Muslim vs. non-Muslim	+0.26
Value formation	education (within-country SD)	-0.08**
	age (within-country SD)	-0.07*
	education, family-obligation subindex	-0.11**
Discriminant	correlation with perceived state strength	+0.40
	correlation with satisfaction with government	+0.39
	correlation with nationalism	+0.32
	correlation with universal human rights	+0.14
	correlation with factual political knowledge	-0.16

Note: Structure rows are computed on the twelve items pooled across countries. Known-group rows are differences in the overall index from within-country bivariate regressions, in index points. Value-formation rows are coefficients from a pooled regression with country fixed effects, with education and age standardized within country. Discriminant rows are pairwise correlations. Robust standard errors throughout. * $p < 0.05$, ** $p < 0.01$.

The second search relaxes functional form, following Hainmueller, Mummolo and Xu (2019), who show that linear interaction models can conceal nonlinear heterogeneity. A tercile binning estimator and a quadratic interaction both leave the penalty essentially flat across the index, and the highest tercile is in fact the least punitive. The third search compares the extreme deciles, where any true gradient should be widest. The fourth estimates the interaction for each of the twelve items separately, in case averaging washes out a signal carried by one of them. The fifth restricts attention to respondents whose value responses are internally coherent, on the reasoning that only crystallized values can resonate. The sixth restricts attention to respondents who spent longer on the vignette, since resonance requires processing the message. The seventh estimates the interaction within each country.

None of these searches recovers the predicted gradient. Across roughly thirty specifications, the interaction is never negative and significant in the US experiment, nor positive and significant in the China experiment. What appears instead, in a handful of specifications, is the opposite sign: three of the twelve items in the US experiment show a significant positive interaction, meaning that the penalty is if anything smaller among respondents higher in Asian-values affinity. We do not build on those scattered results, and we note that a search of this size will produce some significant coefficients by chance (Simonsohn, Simmons and Nelson, 2020). But their direction matters for interpretation. A measure too blunt to detect a real effect produces null coefficients scattered around zero; it does not produce coefficients that lean consistently against the hypothesis. Taken together with the alternative codings above and the horse-race in Section A.6, in which Asian-values affinity moves to the wrong side of zero once nationalism is held constant, we read this as evidence that the moderation null reflects the world rather than our measurement of it (Rainey, 2014).

Table A20: Searching for heterogeneity by Asian values

Search	Specification	US experiment	China experiment
Policy-matched dimension	T2 vs. control	0.116	-0.011
	T3 vs. control	0.075	-0.048
Tercile of the index	lowest tercile	-0.948**	-0.021
	middle tercile	-1.434**	0.174
	highest tercile	-0.642**	-0.308
	quadratic interaction	0.009	-0.006
Extreme groups	bottom decile	-1.209**	0.321
	top decile	-0.788	-0.699
Item level (12 items)	significant, predicted sign	0 of 12	0 of 12
	significant, opposite sign	3 of 12	1 of 12
Value crystallization	coherent responses	-0.072	-0.134
	diffuse responses	0.021	-0.044
Attentive respondents	above median time	0.026	-0.119
	top tercile of time	0.083	-0.108
Within country	Indonesia	0.287	-0.135
	Malaysia	-0.103	-0.092
	Philippines	-0.181	-0.181
	Vietnam	0.218	-0.022

Note: Entries are interaction coefficients, except in the tercile and extreme-group rows, where they are the effect of illiberal rhetoric within the stated subgroup, and the item-level rows, which report counts. Illiberal rhetoric pools T2 and T3 against control; the outcome is the diplomat rating; country fixed effects are included except in the within-country rows. H2 predicts a negative interaction in the US experiment and H4 a positive one in the China experiment; “predicted sign” is scored against each hypothesis accordingly. * $p < 0.05$, ** $p < 0.01$.

Table A21: Ordered logit, US experiment (pooled)

	(1) Diplomat	(2) Power	(3) Choose sides
main			
T1: Neutral policy	-0.101 (0.090)	-0.061 (0.092)	-0.143 (0.105)
T2: Illiberal policy A	-0.654** (0.094)	-0.441** (0.095)	-0.028 (0.105)
T3: Illiberal policy B	-0.525** (0.101)	-0.272** (0.099)	-0.085 (0.101)
Observations	2,551	2,529	2,460

Note: Ordered-logit coefficients on arm indicators, country fixed effects included. * $p < 0.05$, ** $p < 0.01$.

Table A22: Ordered logit, China experiment (pooled)

	(1) Diplomat	(2) Power	(3) Choose sides
main			
T1: Neutral policy	0.199* (0.097)	0.036 (0.099)	0.142 (0.102)
T2: Illiberal policy A	-0.002 (0.099)	-0.057 (0.095)	0.002 (0.104)
T3: Illiberal policy B	0.008 (0.100)	0.036 (0.100)	0.094 (0.104)
Observations	2,557	2,526	2,446

Note: See note to Table A21.

Table A23: Main effects excluding the fastest decile of vignette readers

	(1) US diplomat	(2) US	(3) CH diplomat	(4) China
T1: Neutral policy	-0.195 (0.141)	-0.080 (0.143)		
T2: Illiberal policy A	-1.212** (0.152)	-0.804** (0.149)		
T3: Illiberal policy B	-1.092** (0.159)	-0.610** (0.152)		
T1: Neutral policy			0.336* (0.149)	0.038 (0.159)
T2: Illiberal policy A			-0.001 (0.150)	-0.095 (0.153)
T3: Illiberal policy B			0.018 (0.150)	0.069 (0.157)
Observations	2,296	2,277	2,298	2,274
Adj. R ²	0.072	0.098	0.035	0.058

Note: Respondents in the fastest decile of vignette page time within country are excluded. Columns 1–2: US experiment; columns 3–4: China experiment. * $p < 0.05$, ** $p < 0.01$.

Table A24: Benjamini-Hochberg q-values for the eighteen main treatment-effect tests

Experiment	Outcome	Arm	Estimate	p	q (BH)
US	Diplomat rating	T2	-1.095	0.000	0.000
US	Diplomat rating	T3	-0.970	0.000	0.000
US	Country rating	T2	-0.721	0.000	0.000
US	Country rating	T3	-0.519	0.000	0.001
China	Diplomat rating	T1	0.319	0.022	0.081
China	Choose sides	T1	0.116	0.104	0.313
US	Choose sides	T1	-0.102	0.158	0.403
US	Diplomat rating	T1	-0.178	0.179	0.403
US	Choose sides	T3	-0.082	0.245	0.490
China	Choose sides	T3	0.075	0.298	0.537
US	Country rating	T1	-0.122	0.359	0.588
China	Country rating	T1	0.082	0.586	0.772
US	Choose sides	T2	-0.038	0.594	0.772
China	Country rating	T3	0.075	0.614	0.772
China	Country rating	T2	-0.067	0.643	0.772
China	Diplomat rating	T3	0.016	0.913	0.951
China	Choose sides	T2	-0.005	0.940	0.951
China	Diplomat rating	T2	0.009	0.951	0.951

Note: All arm-by-outcome tests from the two pooled experiments, ordered by p -value, with the Benjamini-Hochberg false-discovery-rate adjustment.

Table A25: Randomization-inference p-values, diplomat ratings

	US: diplomat		China: diplomat	
	Estimate	RI p	Estimate	RI p
T1: Neutral policy	-0.178	0.252	0.319	0.028
T2: Illiberal policy A	-1.095	0.000	0.009	0.888
T3: Illiberal policy B	-0.970	0.000	0.016	0.882

Note: Assignment re-randomized 1,000 times within country; entries compare each estimate with its permutation distribution.

A.6 Mechanisms

Table A26: Treatment effects on the attribution items, US experiment

	(1) Criticism of lib-dem standards	(2) Security role	(3) Economic partner	(4) Telling how to run politics
T1: Neutral policy	0.010 (0.041)	-0.041 (0.042)	-0.060 (0.038)	-0.013 (0.043)
T2: Illiberal policy A	-0.032 (0.039)	-0.023 (0.041)	-0.063 (0.038)	-0.084* (0.042)
T3: Illiberal policy B	0.051 (0.042)	0.002 (0.041)	-0.055 (0.038)	-0.045 (0.043)
Control-group mean	-0.29	0.24	0.52	-0.25
p: T1=T2	0.293	0.667	0.923	0.090
p: T1=T3	0.316	0.306	0.898	0.448
Observations	2,541	2,539	2,542	2,540
Adj. R ²	0.003	0.092	0.072	0.056

Note: Outcomes are coded -1 (displeased), 0 (unaffected), +1 (pleased); pooled with country fixed effects. * $p < 0.05$, ** $p < 0.01$.

Table A27: Treatment effects on the attribution items, China experiment

	(1) Approval: Asian values	(2) Security role	(3) Economic partner	(4) Approval: domestic politics
T1: Neutral policy	-0.017 (0.041)	-0.085* (0.043)	-0.075 (0.042)	-0.054 (0.043)
T2: Illiberal policy A	0.029 (0.040)	-0.067 (0.041)	0.029 (0.041)	-0.007 (0.041)
T3: Illiberal policy B	-0.014 (0.041)	-0.051 (0.042)	-0.009 (0.042)	-0.002 (0.042)
Control-group mean	0.08	0.08	0.13	0.08
p: T1=T2	0.255	0.660	0.011	0.262
p: T1=T3	0.930	0.429	0.116	0.218
Observations	2,544	2,543	2,547	2,544
Adj. R ²	0.121	0.108	0.207	0.080

Note: See note to Table A26.

Table A28: Correlations among the moderators (pooled)

	Asian values	Nationalism	Human rights	Mil. strength	Govt satisf.	Knowledge	News: dom.	News: for.
Asian values	1.00							
Nationalism	0.31	1.00						
Human rights	0.13	0.18	1.00					
Mil. strength	0.40	0.43	0.07	1.00				
Govt satisfaction	0.39	0.38	0.13	0.50	1.00			
Knowledge	-0.16	0.02	0.05	-0.16	-0.06	1.00		
News: domestic	0.26	0.35	0.09	0.31	0.43	0.08	1.00	
News: foreign	0.26	0.22	0.08	0.24	0.35	0.10	0.68	1.00

Note: Correlations computed on the sample with complete data on all eight measures (listwise deletion, $N = 5,068$); they therefore differ slightly from the pairwise correlations in Table A19. Columns follow the same order as the rows.

Table A29: Moderator horse-race, China experiment (diplomat rating)

	(1) Nationalism	(2) Mil. strength	(3) Human rights	(4) Asian values	(5) Horse-race
T1: Neutral policy	0.714 (0.741)	0.309 (0.523)	-0.239 (0.376)	0.391 (0.536)	0.777 (0.784)
T2: Illiberal policy A	-0.340 (0.684)	0.051 (0.481)	0.378 (0.373)	0.741 (0.582)	0.510 (0.786)
T3: Illiberal policy B	0.381 (0.730)	-0.100 (0.496)	0.698 (0.367)	0.895 (0.563)	1.219 (0.802)
T1: Neutral policy $\times$ Nationalism	-0.037 (0.086)				-0.104 (0.098)
T2: Illiberal policy A $\times$ Nationalism	0.045 (0.080)				0.104 (0.097)
T3: Illiberal policy B $\times$ Nationalism	-0.033 (0.085)				0.010 (0.094)
T1: Neutral policy $\times$ Mil. strength		0.010 (0.074)			0.002 (0.084)
T2: Illiberal policy A $\times$ Mil. strength		-0.002 (0.069)			-0.028 (0.082)
T3: Illiberal policy B $\times$ Mil. strength		0.019 (0.070)			0.038 (0.082)
T1: Neutral policy $\times$ Human rights			0.076 (0.047)		0.058 (0.046)
T2: Illiberal policy A $\times$ Human rights			-0.042 (0.047)		-0.050 (0.046)
T3: Illiberal policy B $\times$ Human rights			-0.077 (0.046)		-0.073 (0.044)
T1: Neutral policy $\times$ Asian values				0.005 (0.077)	0.008 (0.085)
T2: Illiberal policy A $\times$ Asian values				-0.105 (0.084)	-0.119 (0.091)
T3: Illiberal policy B $\times$ Asian values				-0.121 (0.083)	-0.145 (0.090)
Observations	2,540	2,540	2,542	2,511	2,484
Adj. R ²	0.102	0.113	0.066	0.146	0.168

Note: Companion to Table 7 of the paper. Columns 1–4 interact the arms with each moderator separately; column 5 estimates all jointly. * $p < 0.05$, ** $p < 0.01$.

Table A30: Government satisfaction, knowledge, and news attention as moderators, US experiment

	(1) Govt satisfaction	(2) Knowledge	(3) News: domestic	(4) News: foreign
T1: Neutral policy	0.578 (0.427)	-0.370 (0.222)	0.527 (0.434)	0.437 (0.345)
T2: Illiberal policy A	-0.921* (0.448)	-0.861** (0.235)	-0.362 (0.461)	-0.949* (0.377)
T3: Illiberal policy B	-0.210 (0.474)	-0.796** (0.254)	0.705 (0.497)	-0.824* (0.408)
T1: Neutral policy $\times$ B2_govt.1	-0.107 (0.060)			
T2: Illiberal policy A $\times$ B2_govt.1	-0.031 (0.063)			
T3: Illiberal policy B $\times$ B2_govt.1	-0.119 (0.067)			
T1: Neutral policy $\times$ knowledge		0.193 (0.175)		
T2: Illiberal policy A $\times$ knowledge		-0.212 (0.184)		
T3: Illiberal policy B $\times$ knowledge		-0.172 (0.204)		
T1: Neutral policy $\times$ K1.1			-0.093 (0.060)	
T2: Illiberal policy A $\times$ K1.1			-0.104 (0.063)	
T3: Illiberal policy B $\times$ K1.1			-0.237** (0.069)	
T1: Neutral policy $\times$ K1.2				-0.098 (0.054)
T2: Illiberal policy A $\times$ K1.2				-0.025 (0.058)
T3: Illiberal policy B $\times$ K1.2				-0.032 (0.063)
Observations	2,509	2,531	2,527	2,525
Adj. R ²	0.098	0.070	0.091	0.097

Note: Each column interacts the arms with one pre-treatment moderator, pooled with country fixed effects. The China-experiment counterpart is available in the replication package. * $p < 0.05$, ** $p < 0.01$.

A.7 Replication

All analyses can be reproduced from the replication package, which contains the raw survey exports, a codebook, numbered Stata do-files that construct the analysis datasets and produce every table and figure in the paper and this SI, and a ReadMe describing the pipeline. Every number reported in the text was additionally verified against an independent Python implementation of the same pipeline. The package also contains the per-country balance, moderation, and value-subindex tables omitted here for length.

A.8 Ethics & IRB

Respondent anonymity was a priority throughout. Participants could skip any survey item except age verification and consent. [Institution A]’s and [Institution B]’s Institutional Review Boards classified the survey as minimal risk (#STUDY00002300; #804650), as no personally identifiable information is retained. A consent form informed participants of their right to discontinue at any point. Recognizing that expressing political views contrary to government positions can carry risk in parts of the region, particularly Vietnam, the study relies on randomized vignettes, which protects respondents because no individual answer reveals a respondent’s own political position, and compensates respondents through the panel provider’s standard arrangements.

A.9 Pre-registration and Deviations

A pre-analysis plan was lodged at OSF on November 26, 2023, before data collection. Hypotheses H1 through H4 in the paper reproduce the registered hypotheses verbatim. We document all deviations and additions here.

1. **Registered hypotheses that failed.** H2, H3, and H4 are not supported. The paper reports these failures as findings, and the framing of the paper, including the sovereignty account, was developed after observing them.
2. **Exploratory analyses.** The symmetry test, the attribution-item analyses, the moderator horse-race, the alternative index codings, the government-satisfaction, knowledge, and news-attention moderators, and the cross-country heterogeneity analyses were not registered. The paper labels all of them exploratory.
3. **Multiple-testing corrections and randomization inference** were not specified in the plan; we add them as robustness (Section A.5).
4. **Sample.** The plan’s sampling section references six countries in one passage; the study was designed for and fielded in the four countries analyzed here.
5. **Severity ordering of the illiberal-policy arms.** The plan labels the two illiberal-policy arms “example 1” and “example 2.” The intended ordering, in which policy A is the comparatively less severe and policy B the more severe within each country, guided vignette construction but is not stated in the plan. We therefore present it as a design-intent claim, and we do not rest any registered conclusion on it.

References

- Adcock, Robert and David Collier. 2001. “Measurement Validity: A Shared Standard for Qualitative and Quantitative Research.” *American Political Science Review* 95(3):529–546.
- Hainmueller, Jens, Jonathan Mummolo and Yiqing Xu. 2019. “How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice.” *Political Analysis* 27(2):163–192.
- Rainey, Carlisle. 2014. “Arguing for a Negligible Effect.” *American Journal of Political Science* 58(4):1083–1091.
- Simonsohn, Uri, Joseph P. Simmons and Leif D. Nelson. 2020. “Specification Curve Analysis.” *Nature Human Behaviour* 4(11):1208–1214.
- Welzel, Christian. 2013. *Freedom Rising: Human Empowerment and the Quest for Emancipation*. New York: Cambridge University Press.