Reading the Room: Nationality-of-Sponsor Effects on Survey Responses - Evidence from China

Kenny Miao* Jay Kao†

July 17, 2026

Abstract

Public opinion surveys are essential tools for studying authoritarian politics. Standard practice makes sponsorship salient through consent forms and institutional banners. Yet we know little about whether sponsor cues affect response patterns. We evaluate nationality-of-sponsor effects through a survey experiment with Chinese adults randomly assigned to varying sponsor nationality conditions. We find that U.S. sponsorship, compared to Chinese sponsorship, increased reported VPN usage and overseas connections, decreased pro-regime behaviors, reduced nationalist sentiment, and increased favorability toward the U.S. and its allies. These effects extended beyond politically sensitive items, appearing even in American music preferences, while leaving trust in domestic institutions unaffected, and they persist under the list experiment. Our findings underscore the necessity of accounting for sponsor nationality when interpreting survey data from authoritarian contexts.

Keywords: survey methodology; sponsor effects; survey experiments; public opinion; authoritarian politics; China

*Department of Government, University of Texas at Austin. Email: kenny.miao@utexas.edu.
†Department of Political Science, Loyola University Chicago. Email: jkao1@luc.edu.

1 Introduction

Much of what political scientists know about public opinion in authoritarian regimes rests on surveys. China scholars, for example, draw on large-scale datasets, such as the Chinese General Social Survey, the World Values Survey, and the Asian Barometer, as well as launching original surveys, to characterize Chinese citizens' political trust (Li, 2016), nationalist sentiment (Johnston, 2017), support for democratic values (Chu et al., 2008), ideologies (Pan and Xu, 2017), regime legitimacy (Jee and Zhang, 2025), and even personality traits (Truex, 2022). These surveys are conducted by institutions of varying national origins, and that origin is rarely hidden: standard research ethics require disclosing researcher affiliation prominently, and online platforms, such as Qualtrics, display institutional logos on every survey page. A fundamental question has gone unanswered: do respondents in authoritarian regimes give systematically different answers depending on these sponsor-nationality cues? If so, a survey's results are confounded with sponsor identity, as they carry a sponsor-specific component of unknown direction, so that figures obtained under one sponsor are not interchangeable with those another would have produced. This contingency is invisible within any single study, where every respondent sees the same sponsor; it surfaces only when results are compared across surveys or against differently sponsored work.

Although recent studies on the United States detect no sponsorship-induced differences in survey responses (Crabtree, Kern and Pietryka, 2022; Leeper and Thorson, 2020), extant research in authoritarian regimes yields a more mixed picture. Lei and Lu (2017) find that revealing Chinese Communist Party affiliation does not affect survey responses in China, while Tannenberg (2022) shows that African respondents answer political questions differently when they believe the government rather than a research institute commissions the survey, though these differences vanish for apolitical questions. All these studies, however, examine variation across domestic institutions—comparing academic, government, or commercial sponsors within the same country. The closest precedent for cross-national sponsor variation comes from Corstange (2014), who shows that foreign government sponsorship increases survey refusal and attrition in Lebanon relative to local survey firms. But Corstange's study concerns participation decisions (whether respondents enroll in the survey), not

response content (how enrolled respondents answer). Whether foreign affiliation also shifts the substantive answers of respondents who do participate remains unknown.

Two reasons lead us to expect that the nationality of sponsors matters. First, in authoritarian regimes respondents have reason to be cautious about the political repercussions that honest answers might invite. Even academic institutions such as universities are often affiliated with or connected to the state, and respondents may worry that candid answers to politically sensitive questions could be passed to the government. Such concerns should weigh more heavily under a domestic sponsor, which respondents can more readily imagine sharing data with the authorities, than under a foreign one; respondents may therefore answer more guardedly when the sponsor is domestic and more candidly when it is foreign.

Second, even absent such fear, respondents might adjust their answers to present themselves favorably to the audience they perceive—and a survey's sponsor signals who that audience is. Facing an audience they take to hold particular preferences, respondents may shade their responses toward the self-image that the audience is thought to prefer rather than report candidly. A Western sponsor plausibly cues an audience of liberal, cosmopolitan, internationally minded researchers; toward it, respondents have incentives to present themselves as more open to the outside world, more sympathetic to liberal-democratic values, and less nationalist than they would before a domestic sponsor, for whom a patriotic, regime-consonant self-presentation is the favorable one. Unlike fear, this audience-tailoring need not be limited to politically sensitive items. Wherever respondents believe the perceived audience holds a preference, self-presentation can move the answer.

Empirically, we preregistered an online survey experiment with more than 3,500 Chinese adults.[1] We randomly assigned respondents to conditions varying the survey sponsor's nationality (Chinese vs. American) and institutional type (academic vs. governmental); among the academic sponsors, we additionally varied university prestige.[2] Respondents encountered the sponsor's affiliation in text on the consent form and as a visual banner displayed on every subsequent survey page. Upon consent,

[1]The anonymized preregistration is available at https://osf.io/32tm6/?view_only=0b0de815718c44798d31619a9d6f9b5b.

[2]Among the internationally visible survey studies of foreign-sponsored research on Chinese opinion, US-affiliated researchers and institutions predominate. A U.S. sponsor is therefore the most realistic and externally valid foreign-sponsor cue for Chinese respondents, which motivates our choice of U.S. sponsors as the treatment.

they answered questions across four domains: foreign exposure (e.g., overseas family connections) and symbolic pro-regime behaviors (e.g., displaying party/national flag), nationalist sentiment and views toward foreign scholars, affect toward foreign entities, and trust in domestic political institutions. At the survey's conclusion, respondents also completed a list experiment on VPN use, which allowed us to test whether indirect questioning removes the effects the sponsor cue induces.

We find three main results. First, sponsor nationality affected responses across various question types: compared to Chinese sponsorship, U.S. sponsorship increased reported VPN usage[3] and overseas family connections, decreased reported pro-regime behaviors, and reduced nationalist sentiment while increasing favorability toward the United States and its allies. Second, contrary to preregistration, these effects extended beyond obviously politically sensitive items, appearing even in favorability toward U.S. music, while leaving reported trust in domestic political institutions unaffected. This suggests response differences reflect not only political concerns but also impression management toward Western sponsors. Third, this effect persists even within a list experiment, the field's leading anonymity-based technique for eliciting candid answers, indicating that standard anonymity protections do not remove it.

This study makes two contributions. First, we extend research on survey-sponsorship effects to the cross-national case. Prior work has varied institutional type and university prestige while holding sponsor nationality fixed (Crabtree, Kern and Pietryka, 2022; Leeper and Thorson, 2020). By crossing nationality with institutional type and academic prestige in a symmetric design, we isolate the foreign-versus-domestic contrast from these dimensions and show that it shifts respondents' answers across a wide range of items.

Second, our results speak to the self-censorship and preference-falsification literature (Robinson and Tannenberg, 2019; Jiang and Yang, 2016; Nicholson and Huang, 2023; Li, 2016; Li et al., 2026; Tannenberg, 2022), which focuses on how respondents in authoritarian settings conceal their true views on politically sensitive items. We find that the gap between respondents who believe the sponsor is American and those who believe it is Chinese is not confined to sensitive political

[3]Chinese citizens routinely utilize VPNs to bypass the state's digital firewall and access politically sensitive or restricted content. Although such a behavior is widespread, it is illegal according to the law and might incur government punishment.

questions but appears on apolitical ones as well, including a question about musical taste. Because the foreign-domestic gap reaches content that carries no political risk while trust in domestic institutions is unchanged, an account based purely on fear of repercussion cannot be the whole story; respondents also appear to engage in broader, audience-oriented impression management.

Our findings carry implications for the study of authoritarian politics and for public-opinion research more broadly. Because sponsorship is a built-in feature of any survey rather than a manipulable item, its effect is difficult to detect within a single study. That the foreign-domestic gap persists even under a list experiment indicates that item-level anonymity does not eliminate it. Researchers should therefore weigh how sponsor identity might affect respondents' behaviors at the design stage, report how it is disclosed to respondents transparently, and discuss to what extent the central conclusion and inference can be affected by the sponsor effect.

2 How Sponsor Nationality Shapes Responses

A handful of studies have examined how survey sponsorship shapes respondents' answers, with mixed results. In the U.S. context, surveys run by universities of differing prestige or ideological orientation produce no difference in responses (Crabtree, Kern and Pietryka, 2022), and surveys sponsored by commercial firms versus universities likewise elicit similar responses (Leeper and Thorson, 2020). Other studies find that sponsorship does matter: Tannenberg (2022) finds that African respondents overreport support for the regime when they believe the survey is government-run, and Corstange (2014) finds that respondents in Lebanon are more willing to participate in surveys sponsored by academic institutions than in those sponsored by a foreign government.

These studies leave two questions open. First, they do not isolate national origin: the U.S. studies hold sponsor nationality fixed and vary only institutional type, and where nationality does vary, it moves together with type (academic versus foreign government), so the independent effect of national origin is unidentified. Second, where effects appear they are typically read as self-censorship: respondents withhold risky truths on politically sensitive items, so the effects are expected to be confined to those items. Tannenberg (2022), for instance, finds that government-versus-academic

differences vanish for apolitical questions. Whether sponsorship also shifts responses to apolitical items, rather than only sensitive ones, remains an open question.

We argue that sponsor nationality can shape responses through two analytically distinct mechanisms that differ in their predicted scope. The first is political fear. Citizens in repressive settings have incentives to conceal heterodox views and behaviors, what Kuran (1991) termed preference falsification, a dynamic well documented in Chinese survey research, where respondents understate regime-critical attitudes and behaviors and indirect techniques routinely recover higher prevalence than direct questions (Carter, Carter and Schick, 2024; Nicholson and Huang, 2023; Li et al., 2026; Ratigan and Rabin, 2020). Because such concealment is calibrated to who is thought to have access to the data, the sponsor's nationality becomes relevant: respondents may doubt confidentiality assurances from domestic institutions, given that universities under authoritarian countries operate under state supervision and data could in principle be requisitioned by authorities, whereas foreign institutions have no comparable channel to domestic authorities, which may signal lower domestic-surveillance risk and license more candid answers.[4] Crucially, political fear is a targeted mechanism: it operates on items whose honest answer carries political risk, such as heterodox behaviors, regime criticism, and gives no reason to expect movement on items that are not politically sensitive.

The second mechanism is impression management. Even absent any threat of repercussion, respondents edit their answers to present themselves favorably to the audience they perceive (Holbrook and Krosnick, 2010), and a growing literature shows that they tailor this self-presentation to the perceived identity of whoever is collecting the data, including in settings where no answer carries tangible consequences (Adida et al., 2016; Gengler et al., 2021; White et al., 2018). Respondents are likely to perceive a Western sponsor as valuing cosmopolitan, liberal-internationalist outlooks and a domestic, state-linked sponsor as prizing patriotic conformity. What matters for self-presentation is not whether these perceptions are accurate but that they are widely shared—national stereotypes operate as consistent, commonly held images largely independent of the traits they ascribe (Terracciano et al., 2005). In post-Mao China, the West has long been imagined as a liberal, modernizing "other"

[4]If respondents disbelieved the foreign affiliation outright—suspecting a government "trick"—no systematic differences between conditions should emerge; the differences we observe are inconsistent with such blanket suspicion.

(Chen, 1995), while decades of state-led patriotic education have made nationalist sentiment the socially expected, regime-rewarded posture (Wang, 2008). Respondents may therefore signal in-group alignment under a co-national sponsor and seek to appear cosmopolitan, distancing themselves from ultra-nationalist caricatures under a U.S. one.

The two mechanisms converge on politically relevant outcomes, where both predict that respondents present as more internationally oriented and less regime-aligned under U.S. than Chinese sponsorship. We preregistered four hypotheses:

H1 (Pro-regime behavior). Respondents report less symbolic regime support—displaying national or Party flags, reporting government-critical speech—under U.S. than Chinese sponsorship.

H2 (Foreign exposure). Respondents report more foreign experiences and international connections—overseas travel, relatives abroad, VPN use—under U.S. than Chinese sponsorship.

H3 (Political attitudes). Respondents express warmer sentiment toward the United States and its allies, and lower nationalist sentiment, under U.S. than Chinese sponsorship.

H4 (Domestic institutional trust). Respondents report lower trust in domestic political institutions under U.S. than Chinese sponsorship.

Because both mechanisms imply the same directional predictions on these politically relevant items, H1–H4 do not by themselves distinguish targeted self-censorship from broad self-presentation. That distinction turns on apolitical outcomes, where only the impression-management account anticipates an effect. Our pre-analysis plan specified hypotheses for politically relevant outcomes only; we therefore examine apolitical items on an exploratory basis. One pre-registered outcome variable is suitable for this task: preference for American music. Systematic effects on this outcome would be hard to reconcile with self-censorship alone and would instead indicate that sponsor nationality engages a general process of audience-oriented self-presentation.

H5 (Apolitical preference; exploratory). Respondents are more likely to report having an American singer or band they like under U.S. than under Chinese sponsorship. Because this outcome was preregistered as a measured variable but not tied to a directional hypothesis, we report it as exploratory rather than confirmatory.

3 Experimental Design and Data

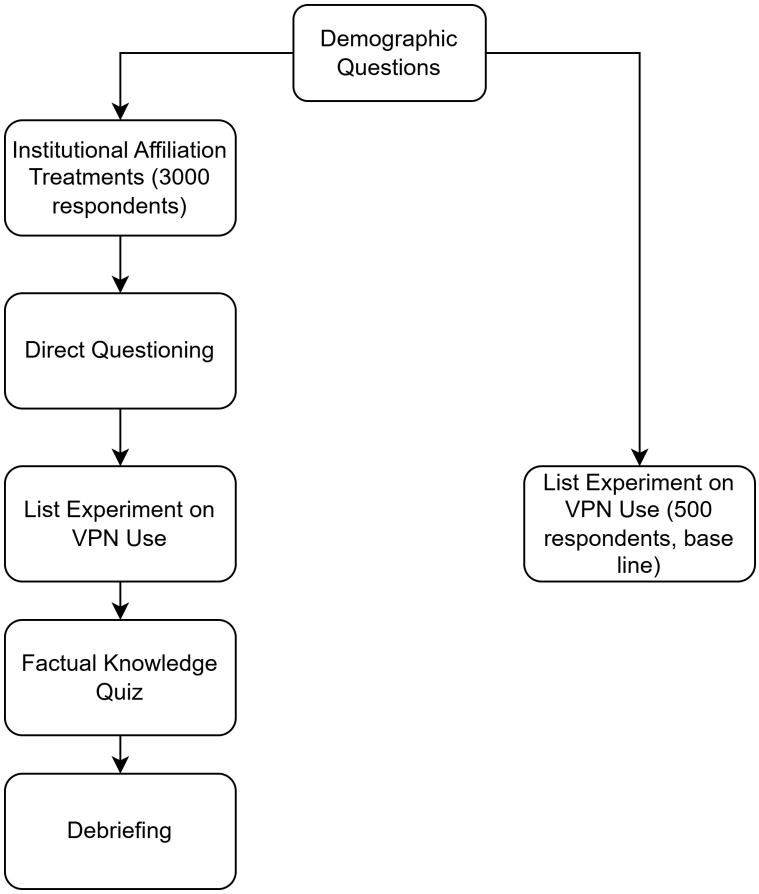
We fielded the survey experiment in June 2025 through Credamo, a major commercial Chinese online survey platform. A total of 3,520 Chinese adults entered the survey. As is typical for online samples (Li, Shi and Zhu, 2018), respondents skew younger and more educated than the general population, though the sample achieves broad regional coverage across all mainland provincial units (see Appendix A for the sample demographics and their comparison with China's population as a whole).

Our design, building on Crabtree, Kern and Pietryka (2022) and Leeper and Thorson (2020), randomly assigned respondents to conditions varying the stated institutional affiliation of the survey research team, manipulating both nationality (Chinese vs. American) and institutional type (academic vs. governmental). To account for potential effects of university prestige (Crabtree, Kern and Pietryka, 2022; Edwards, Dillman and Smyth, 2014), we included two universities per country varying in prestige: Harvard University and Loyola University Chicago for the United States; Tsinghua University and Hebei Normal University for China. To test government sponsorship effects (Corstange, 2014; Tannenberg, 2022), we included the U.S. Embassy in China and the Communist Youth League of China. Utilizing multiple institution types within each nationality, this design allows us to assess whether effects are driven primarily by sponsor nationality or vary by their institutional characteristics—a distinction that prior work with limited sponsor variation cannot address.

Upon entering the survey, and prior to any experimental manipulation, respondents answered demographic questions including age, gender, income, education, Chinese Communist Party (CCP) membership, local household registration status (hukou), and Internet usage. Measuring these items before the treatment ensures they are uncontaminated by the sponsor cue and can serve as valid pre-treatment covariates for assessing balance across conditions. The platform then randomly assigned respondents, with equal probability, to one of seven arms: six sponsor conditions ($n \approx 3{,}000$) and one list-experiment-only condition ($n \approx 504$). Respondents in the six sponsor conditions were shown a consent form stating: *"We are a research team affiliated with [INSTITUTION], conducting a study of Chinese Internet users' opinions on social issues and online behaviors."* The assigned institution's

banner appeared at the top of every subsequent survey page, making the sponsor cue salient (see Appendix B for banners and a survey page example). Respondents could decline to continue at the consent page. Upon consent, they completed four sets of outcome measures, including background and behaviors, nationalist sentiment and views toward foreign scholars, affect toward foreign entities, and trust in domestic institutions (full wording in Table 1). They were then asked to complete a short political-knowledge quiz to gauge engagement, and concluded with a list experiment on VPN use.[5] In contrast, respondents in the list-experiment-only condition received no sponsor cue, answered no direct questions of any kind, and completed only the list experiment.

Figure 1: Survey Flow



Per our preregistration, we excluded respondents whose geolocation coordinates fell outside mainland China and those with duplicate coordinates (12 respondents), leaving 3,508 respondents.

[5]We debriefed the respondents about our true institutional affiliation at the end of the survey. The study was approved by the Institutional Review Board (see Appendix P).

Table 1: Outcome Variables, Question Wording, Preregistration Status, and Response Scales

Question Label	Question Wording	Hypothesis / Status	Scale
Ever Abroad	Have you ever been abroad (study, travel, or work)?	H2	0–1
Relative Abroad	Do you have relatives living or working abroad?	H2	0–1
VPN Use	Do you regularly use technical tools to access foreign websites?	H2	0–1
Liking U.S. Singer/Band	Do you have an American singer or band you like?	Exploratory (H5)†	0–1
Flag Display	Do you display the national or Party flag at home or on your phone?	H1	0–1
Report Anti-China Speech	How would you respond to online comments critical of China? (1 = "report them")	H1	0–1
Family Works for Party/State	Do you have any immediate family members (parents/children/spouse) working in the party or government?	Exploratory†	0–1
Unconditional National Support	Even when their country is in the wrong, people should still support it.	H3	1–4 (SD→SA)
Foreign Scholars' Right to Criticize	Foreign scholars have the right to criticize problems in China.	H3	1–4 (SD→SA)
Foreign Scholars Ill-Intentioned	Foreign researchers are ill-intentioned.	H3	1–4 (SD→SA)
Most People Can Be Trusted	Most people can be trusted in our society.	Exploratory†	1–4 (SD→SA)
Feeling – U.S.	Please rate your overall feelings toward the U.S.	H3	0–100
Feeling – Taiwan	Please rate your overall feelings toward Taiwan.	H3	0–100
Feeling – EU	Please rate your overall feelings toward the EU.	H3	0–100
Feeling – Japan	Please rate your overall feelings toward Japan.	H3	0–100
Feeling – Russia	Please rate your overall feelings toward Russia.	Exploratory†	0–100
Trust – Neighborhood Committee	Please rate your trust in the Neighborhood Committee.	H4	0–100
Trust – Courts	Please rate your trust in the Courts.	H4	0–100
Trust – Local Officials	Please rate your trust in Local Government Officials.	H4	0–100
Trust – People's Daily	Please rate your trust in People's Daily.	H4	0–100

Note: H1 = Pro-Regime Behavior; H2 = Foreign Exposure; H3 = Political Attitudes; H4 = Trust in Domestic Institutions. H1–H4 are preregistered confirmatory hypotheses; † Preregistered as a measured variable but not tied to a hypothesis; analyzed as exploratory.

Four respondents declined after viewing the consent form, leaving 3,504 for analysis. Balance checks show that respondents are similar across treatment arms on nearly all pre-treatment covariates (Appendix C). The two exceptions are minor: income shows imbalance on the joint test across all seven arms (joint p=0.043) but is well balanced on our focal U.S.-versus-Chinese contrast, while education is balanced across arms but differs modestly on that focal contrast (about three percentage points). Both are small, of the magnitude expected by chance under multiple comparisons, and we report unadjusted estimates throughout (Appendix C).

4 Results

4.1 Main Results

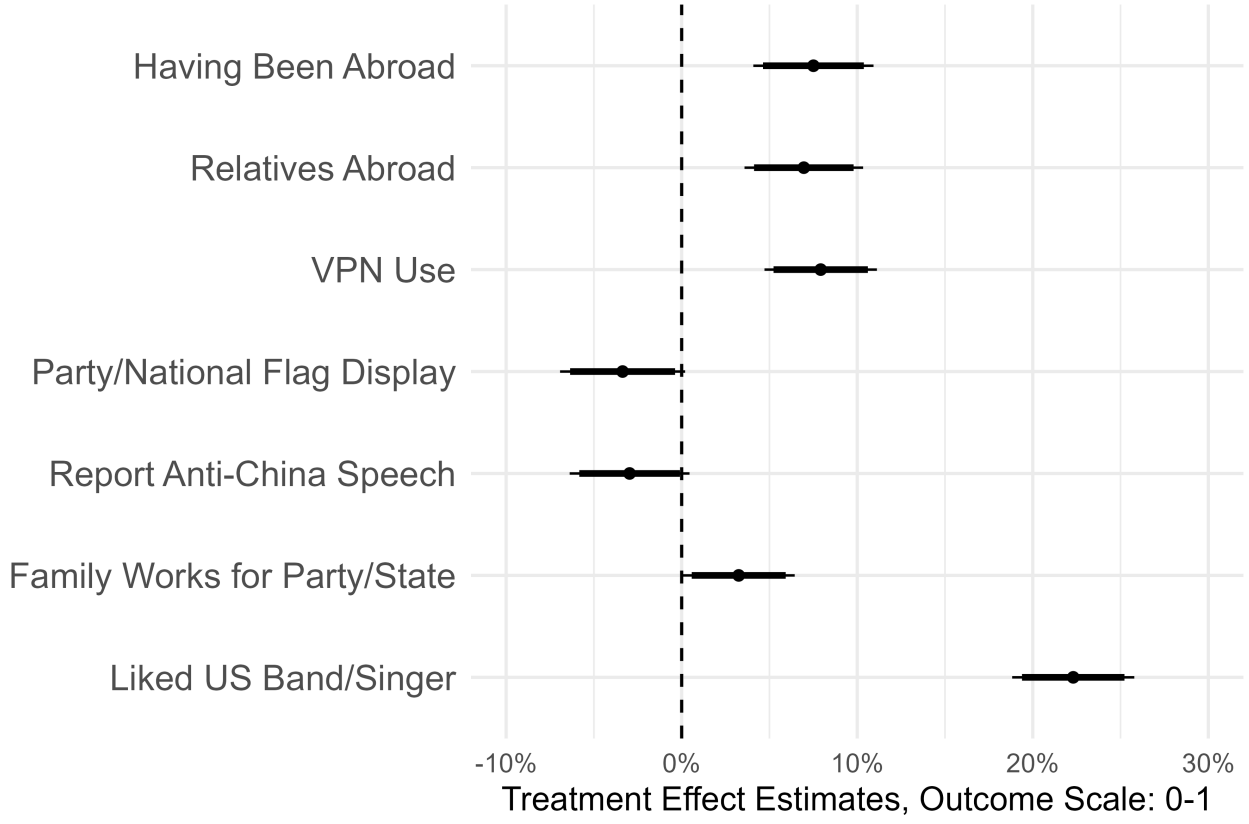
Our preregistered analysis compares respondents randomly assigned to U.S. versus Chinese sponsorship. Because treatment is randomized, our primary estimates are unadjusted: for each outcome, we estimate the U.S.-versus-Chinese difference in means—equivalently, an OLS regression of the

outcome on an indicator for U.S. sponsorship with no covariates.[6] The results reveal a consistent pattern: sponsor nationality systematically affects responses across diverse question types and sensitivity levels, from ostensibly objective background questions to subjective and sensitive political attitudes.

We begin with behavioral and background items. Figure 2 shows that, relative to Chinese sponsorship, U.S. sponsorship reduces reported symbolic pro-regime behavior (H1) and increases reported international exposure (H2). U.S. sponsorship reduces reported pro-regime behaviors: respondents are 3.3 percentage points less likely to report displaying Party or national flags at home (significant at the 10% level) and 2.9 percentage points less likely to report or intend to report anti-regime online speech to authorities (significant at the 10% level). Conversely, respondents assigned to U.S. conditions are 7.5 percentage points more likely to report having traveled abroad, 7.0 percentage points more likely to report having relatives living overseas, and 8.0 percentage points more likely to report using a VPN.

[6]We also report raw group means for all six treatment arms in Appendix E for transparency.

Figure 2: Effect of U.S. vs. Chinese Sponsorship on Behavioral and Background Outcomes



Note: Points represent OLS estimates of the U.S. sponsorship effect relative to Chinese sponsorship. Thick and thin lines indicate 90% and 95% confidence intervals, respectively. Positive values indicate higher reporting rates under U.S. sponsorship; negative values indicate lower rates. All outcomes are binary. See Table 1 for question wording.

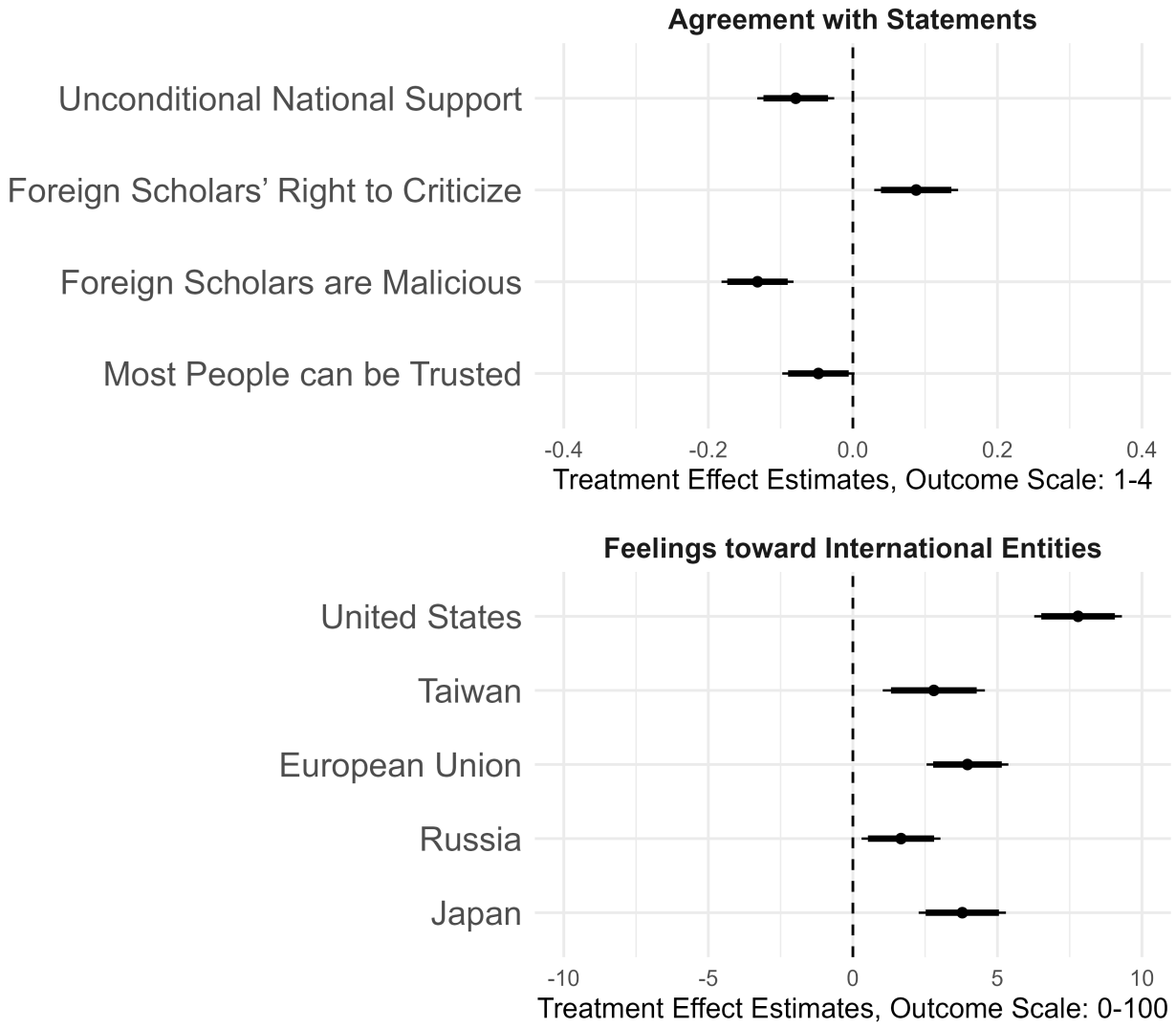
U.S. sponsorship also raises reported rates of immediate-family employment in Party or government agencies, an item we preregistered without a directional hypothesis. Such ties are not conventionally treated as sensitive, but the sustained anti-corruption campaign under Xi, which targets mainly governmental officials and public sector employees, may have inverted the usual disclosure calculus, leaving respondents more willing to reveal them to a foreign than to a domestic sponsor.

The largest effect in the study, however, concerns an apparently innocuous question: whether the respondent has an American singer or band they like. Respondents in the U.S. conditions are 22 percentage points more likely to report a favorite U.S. musician, an item we likewise preregistered without a directional hypothesis. Unlike the political items, where guarded answers could reflect caution about repercussions, a stated taste in music carries no realistic risk before Chinese researchers,

so fear cannot plausibly drive this gap. The pattern instead points to audience-oriented self-presentation: respondents present as more culturally cosmopolitan when they take the sponsor to be American.

We next examine political attitudes (H3). Figure 3 shows that sponsor nationality shapes expressed nationalist sentiment. Relative to Chinese sponsorship, U.S. sponsorship lowers unconditional national support, measured by agreement that "People should always support their country, whether or not it is right." Respondents in U.S. conditions also express greater acceptance of outside criticism: they are more likely to affirm that foreign scholars have a right to criticize China and less likely to view foreign scholars as malicious.

Figure 3: Political Attitudes



Note: Points represent OLS estimates of the U.S. sponsorship effect relative to Chinese sponsorship. Thick and thin lines indicate 90% and 95% confidence intervals, respectively. See Table 1 for question wording.

U.S. sponsorship also produces a modest reduction in general social trust, measured by agreement that "most people can be trusted" (significant at the 10% level), another outcome we preregistered without a directional hypothesis. One speculative interpretation is that the two responses carry different cultural valence: high social trust signals the collectivist values prized in Chinese settings, while skepticism aligns with the individualism associated with Western audiences.

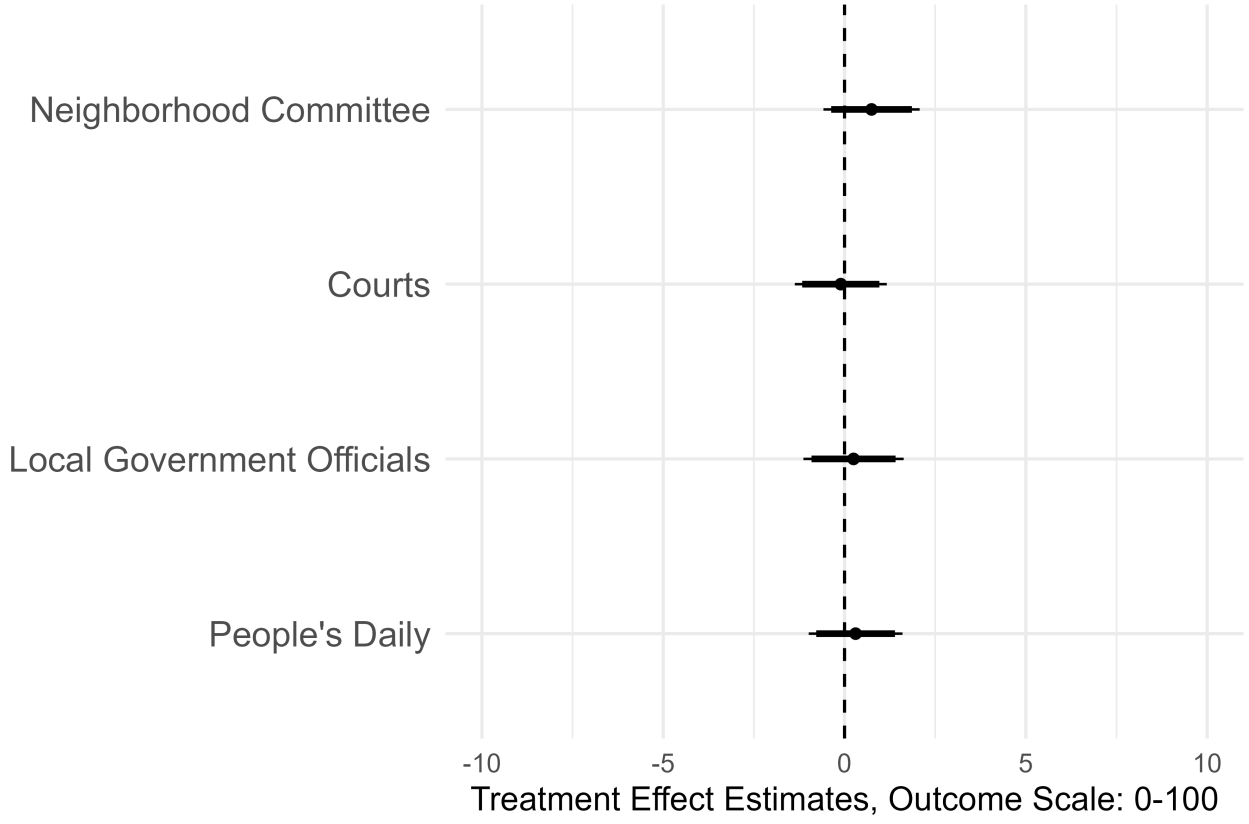
Turning to affect toward foreign entities measured by feeling thermometers (Figure 3, bottom

panel), we find that U.S. sponsorship significantly increases feeling thermometers toward U.S.-aligned countries and partners. Respondents rate the United States, the European Union, Taiwan, and Japan significantly more favorably under U.S. than Chinese sponsorship. Even Russia exhibits a modest positive shift despite its strategic partnership with China, though the effect is considerably smaller than for Western-aligned entities.

Finally, we examine trust in four domestic institutions (H4): neighborhood committees (grassroots state organs), courts, local government officials, and People's Daily (the Party's flagship newspaper).[7] Contrary to our preregistered hypothesis, Figure 4 shows no discernible effect of sponsor nationality on institutional trust. One plausible reading turns on the asymmetric sensitivity of targets within the regime. Nicholson and Huang (2023) find that Chinese citizens fear the central government far more than local authorities and feel relatively free to criticize grassroots and local officials. If respondents already regard criticism of local institutions as safe under a domestic sponsor, a foreign sponsor offers no additional protection, and the sponsor cue has little room to move these answers. We leave a direct test of this account to future work.

[7]We do not test sponsor effects on trust in the central government or top leadership, because platform rules prohibited questions on these most sensitive items.

Figure 4: Trust in Domestic Institutions



Note: Points represent OLS estimates of the U.S. sponsorship effect relative to Chinese sponsorship. Thick and thin lines indicate 90% and 95% confidence intervals, respectively. Positive values indicate higher reported trust under U.S. sponsorship; negative values indicate lower trust. See Table 1 for question wording.

4.2 Alternative Explanations and Robustness Checks

So far we have documented a large gap between the U.S. and Chinese sponsorship conditions across a wide range of items, from politically sensitive questions to apolitical ones. Could this gap reflect something other than respondents adjusting their answers to the sponsor's nationality? We consider several alternatives. The first is differential composition rather than differential response: that sponsor nationality shaped who entered the survey rather than how respondents answered. This concern is real in principle. Corstange (2014) finds that Lebanese respondents were more likely to refuse surveys sponsored by foreign governments than those sponsored by universities, so foreign sponsorship could have changed who participates. If the U.S. and Chinese conditions were composed of different kinds

of respondents, the gap we attribute to response adjustment could instead reflect that compositional difference.

Several features of our design and data weigh against this account. At the consent page, where respondents learned the sponsor's identity, we gave them the option to decline and recorded their choices. Refusal was rare: four respondents (0.27% of those assigned to U.S. conditions) declined under U.S. sponsorship, versus none under Chinese sponsorship. A recruitment effect this small cannot produce response differences of the magnitude we document. To confirm this, we compute the worst-case selection bounds of Lee (2009), which trim the higher-retention Chinese arm to match U.S. retention. Under this adjustment every estimate is essentially unchanged (Appendix D). Moreover, the realized samples are balanced on nearly all pre-treatment covariates (Appendix C), so respondents under the U.S. and Chinese conditions are compositionally similar.

Second, we consider whether the effects are driven mainly by the two governmental sponsors, the U.S. Embassy and the Communist Youth League, rather than by sponsor nationality more generally. Prior work finds that respondents adjust their behaviors most when they believe a survey is sponsored by a government, whether foreign or domestic (Corstange, 2014; Tannenberg, 2022). Were our effects confined to governmental sponsors, they would be less troubling for most survey practice, since most academic surveys are fielded by universities rather than states. We therefore re-estimate all effects using only the academic conditions—Harvard and Loyola versus Tsinghua and Hebei Normal—dropping the two governmental conditions (Appendix H). The core effects are essentially unchanged: foreign exposure, U.S. cultural affinity, and the international-thermometer differences remain as large and as precise as in the full sample. A few estimates attenuate, including reporting government-critical speech, flag display, and unconditional national support, indicating that these more overtly regime-signaling responses are partly specific to governmental sponsors. The broader pattern, however, is not a government-sponsor artifact: sponsor nationality shifts responses even among ordinary academic institutions.

Third, we examine whether the sponsor cue shapes responses not by changing their content but by changing respondents' effort and engagement, a data-quality channel that could in principle generate

spurious treatment effects if respondents disengage or satisfice differently depending on who appears to be asking. To assess this, we placed a set of factual-knowledge questions at the end of the survey and used accuracy on them as an unobtrusive measure of effort (Crabtree, Kern and Pietryka, 2022). Because online respondents can look up answers, we also recorded the time spent on these questions as a second indicator. Neither measure differs significantly across sponsor conditions (Appendix K). The effects we report therefore reflect what respondents report rather than how carefully they respond.

We also ask whether respondents react to the prestige or institutional type of the sponsor rather than to its nationality (Crabtree, Kern and Pietryka, 2022). Both are balanced across our nationality contrast by the symmetric design, so neither can confound the pooled effect. However, they could nonetheless be the active ingredient if elite university or governmental sponsors moved responses while others did not. Holding nationality fixed, within-nationality prestige contrasts are small and unsystematic in sign, as are the academic versus governmental type contrasts. Our equivalence tests (Lee, 2009) show that for most outcome variables affected by nationality, the prestige and institutional-type contrasts are bounded, at the 90% level, below the size of the nationality effect on the same outcome (Appendix I). We do not claim these contrasts are zero; the point is that neither prestige nor institution type reproduces the systematic, sizable shift that nationality produces.

We conduct two additional robustness checks. Because we test many outcomes, we recompute significance using the Benjamini–Hochberg and Holm–Bonferroni corrections for multiple comparisons (Appendix F). The central effects, including foreign exposure, nationalist sentiment, and warmth toward the United States and its allies, remain significant at $p < 0.001$ under even the stringent Holm correction; several smaller estimates, including reporting government-critical speech and flag display, do not survive correction, consistent with their marginal significance in the main analysis. We also re-estimate the main models with post-stratification weights as a robustness probe, not as preferred population estimates, given the sharp divergence between our online sample and the general-population benchmark (Appendix G). Standard errors inflate roughly twofold, so estimates lose precision, but most point estimates stay close to the unweighted results. The point estimates for three smaller contrasts, including feeling toward Russia, general social trust, and family employed by

party or state, fall to near zero, while the effect on VPN use increases. Full weight-adjusted results are in Appendix G.

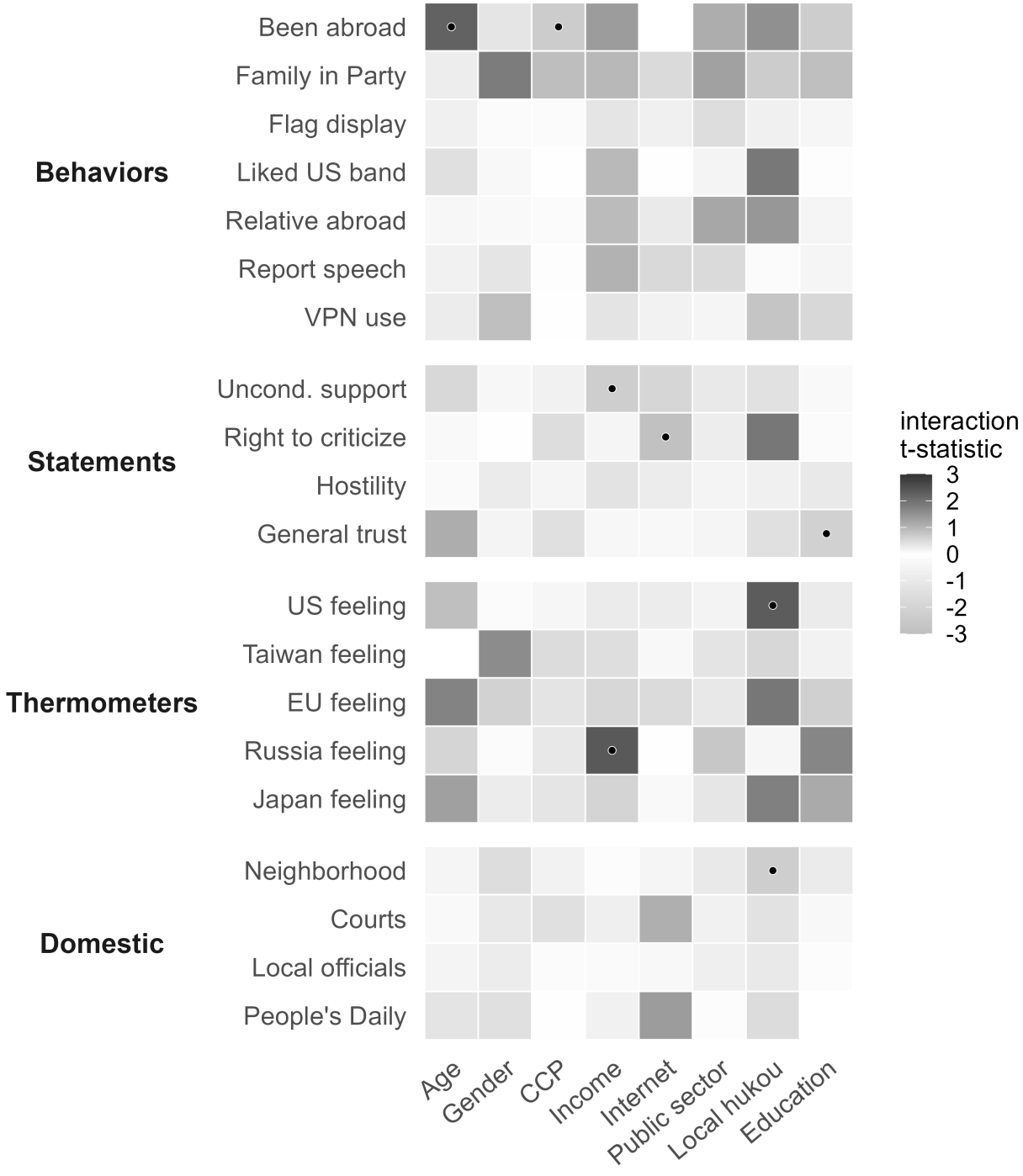
4.3 Effect Heterogeneity

Previous studies have examined which individual traits predict self-censorship and preference falsification, but the directions are mixed. In the Chinese context, for example, less-educated respondents, non-CCP members, and lower-income respondents are more likely to refuse to answer politically sensitive items (Shen and Truex, 2021; Ratigan and Rabin, 2020), whereas more-educated and higher-income respondents appear more prone to falsify their expressed preferences (Jiang and Yang, 2016; Robinson and Tannenberg, 2019).

Are the sponsorship effects we report driven by a particular subset of the population, the more marginalized or the more informed and well-off? To find out, we examine treatment-effect heterogeneity across all eight pre-treatment covariates: age, gender, CCP membership, income, internet-usage time, public-sector employment, local household registration (hukou), and education, each measured before the institutional treatment was revealed.[8]

[8]This analysis is not pre-registered and is purely exploratory; we report it because the patterns are substantively informative, and we treat the results as suggestive rather than confirmatory.

Figure 5: Treatment-Effect Heterogeneity Across Respondent Subgroups



Note: Each cell reports the t-statistic of the interaction between the moderator (column) and the U.S.-sponsorship indicator, estimated from a separate OLS regression of the outcome on U.S. sponsorship, the moderator, and their interaction, with the Chinese-institution conditions as the reference category and HC1 robust standard errors. Positive values (darker) indicate a larger U.S.-sponsorship effect among respondents scoring higher on the moderator; negative values (lighter) the reverse. Dots mark interactions significant at $p < .05$.

Figure 5 presents a heat map in which every cell reports the t-statistic of the interaction between a covariate (columns) and the U.S.-sponsor indicator for a given outcome (rows, grouped by domain), with cells significant at $p < .05$ marked by a dot. We find little evidence of effect heterogeneity: of the 160 moderator-by-treatment interactions, eight reach $p < .05$, exactly the number expected by chance. The sponsorship effect is thus general rather than concentrated in particular subgroups, which speaks against its being an artifact of any single type of respondent.

5 Can List Experiment Mitigate Sponsorship Effects?

Our main analysis demonstrates substantial sponsorship-induced response bias in a self-administered online survey. A natural question arises: can established techniques for reducing social desirability bias mitigate these effects? Prior research recommends indirect questioning methods, particularly list experiments, for eliciting truthful responses on sensitive topics (Di Maio and Fiala, 2020; Glynn, 2013; Blair and Imai, 2012). Studies in China have shown that list experiments successfully reduce response bias on politically sensitive items (Nicholson and Huang, 2023; Carter, Carter and Schick, 2024; Robinson and Tannenberg, 2019; Li et al., 2026). We therefore embedded a pre-registered list experiment on VPN use, a technique for bypassing government information controls and censorship (Chen and Yang, 2019), in the survey to test whether this technique can eliminate sponsorship effects.[9]

The list experiment was administered to all respondents at the survey's conclusion. Of the 3,000 respondents in the six sponsor conditions, 1,500 were assigned (with equal probability) to the treatment list and 1,500 to the control list; the no-sponsor baseline comprised 504 respondents (252 per list). Control-group respondents reported how many of the following four activities they typically engage in: (1) tipping livestreamers, (2) watching short videos, (3) using AI tools for study

[9]Our pre-analysis plan included a second list experiment, on whether foreign scholars have a right to criticize China, which we report descriptively in Appendix N rather than as a parallel test of the question in this section. Two features of its design make it unsuited to that purpose. First, its list arm was fielded only in the no-sponsor baseline, not under the sponsor conditions, so it yields no list-based estimate of the sponsor gap and cannot show whether sponsorship effects persist under anonymity. Second, the item-count technique requires a binary sensitive item, so this attitudinal item, measured on a four-point agree–disagree scale in direct questioning, had to be dichotomized for the list, leaving its list-experiment prevalence and its four-point direct measure on non-comparable scales. Both are design limitations we would avoid in hindsight, by fielding the list arm under all sponsor conditions and choosing a natively binary item. VPN use avoids both problems: it is intrinsically binary and was administered under every sponsor condition.

or work, and (4) donating to charities. Treatment-group respondents received the same list plus one sensitive item, VPN use.[10] Because the sponsor was introduced at the outset and the instrument was fielded under that single sponsor throughout, respondents reached the list experiment under the same perceived sponsor as the rest of the survey; baseline respondents encountered no sponsor at any point.[11]

If the anonymity the list experiment offers removes incentives to misreport under either nationality condition, list-experiment estimates should be invariant across the three conditions: Chinese sponsorship, U.S. sponsorship, and no sponsor cue. An invariant list estimate would then serve as an unbiased benchmark against the direct-questioning estimates in each condition, letting us determine the direction of the gap: whether respondents over-report VPN use under U.S. sponsorship, under-report under Chinese sponsorship, or both.

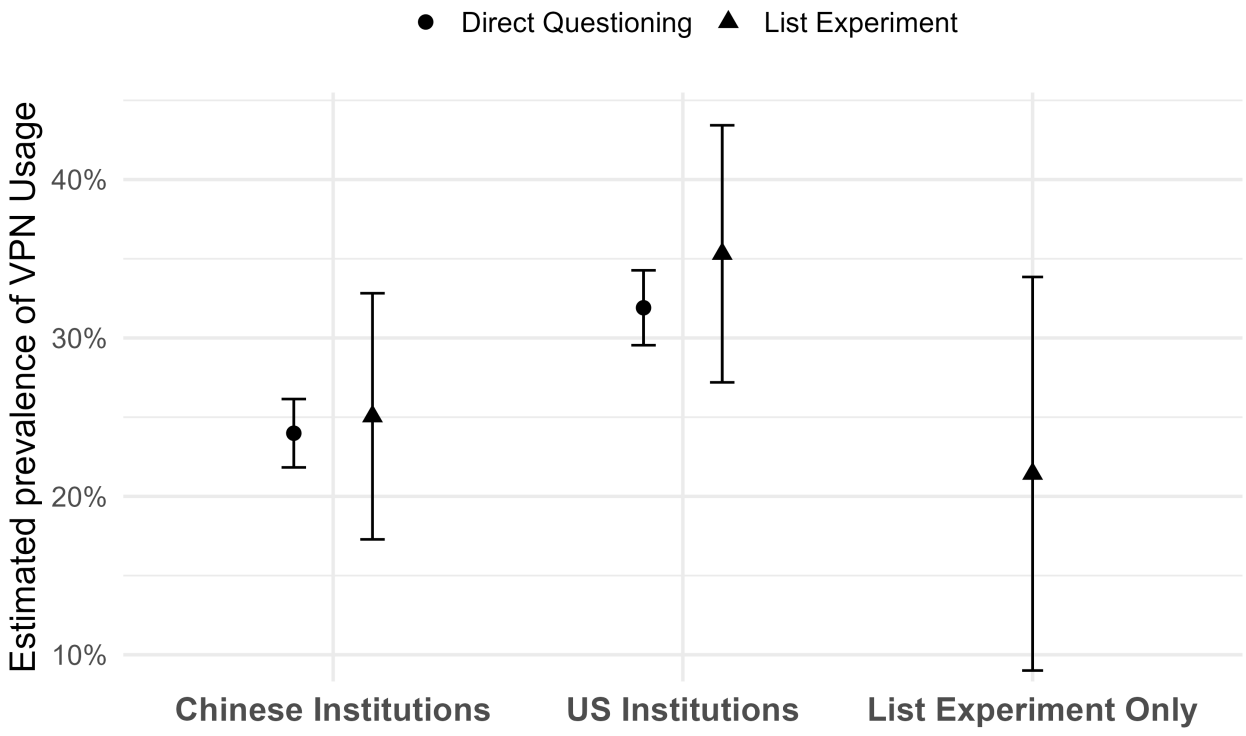
Figure 6 presents the results. For each treatment group, the list-experiment estimate aligns well with the sample mean from direct questioning. The list experiment recovers a U.S.–Chinese gap in VPN prevalence of 10.3 percentage points (0.353 versus 0.250; Table A.13), slightly larger than the 8-point gap under direct questioning. Anonymity, in short, does not pull the conditions toward a common value: respondents answer differently by perceived sponsor nationality even under indirect questioning. Because list-experiment estimates carry far larger sampling variance than direct questions, this gap is estimated less precisely, reaching significance at the 10% level (p= 0.07), but its magnitude is on par with the direct-questioning effect and shows no sign of the attenuation that genuine mitigation would produce. At least in this setting, list experiments do not appear to reduce sponsorship effects.[12]

[10]We employed a double-list design, which enhances statistical power by giving each respondent two lists and randomizing the placement of the sensitive item (Droitcour et al., 2004; Glynn, 2013). However, diagnostics indicate a carryover effect that depresses responses in List 2 (Diaz, 2024), so pooling both lists would yield downward-biased estimates. When a carryover effect is present, List 1 remains a valid single-list experiment. We therefore analyze List 1 and present complete results for both lists in Appendix L.

[11]We conduct two diagnostics on the list-experiment results. First, the Blair and Imai (2012) test finds no evidence of design effects (p = 1.000; Appendix L), indicating that adding the sensitive item does not alter responses to the control items. Second, we check for ceiling and floor effects, which arise when the control items are so common or so rare that respondents' counts pile up at the extremes; at those extremes a respondent's answer to the sensitive item becomes inferable, giving an incentive to misreport. The control-group count distribution clusters well away from both extremes: only 0.2% of List 1 respondents reported none of the four items and 10.5% reported all four (Appendix L), leaving little room for such concealment.

[12]Our finding parallels Karpowitz et al. (2023), who show that list-experiment estimates of attitudes toward immigrants

Figure 6: List Experiment by Country Conditions



Note: Estimates for the list experiment are derived from Ordinary Least Squares (OLS) regression. Results for direct questioning represent sample means. Error bars indicate 95% confidence intervals.

The baseline group, which saw no sponsor cue and completed only the list experiment, yields a VPN prevalence estimate (0.214) nearly identical to the Chinese-sponsorship list estimate (0.250) and far below the U.S. list estimate (0.353).[13] We offer two competing interpretations. The first treats the baseline as candid reporting: U.S. sponsorship then induces overreporting through impression management, as respondents present themselves as cosmopolitan and outward-looking, inflating the list estimate above the true rate. This account requires respondents to treat an unsponsored survey as neutral enough to answer candidly, a strong assumption given that most surveys they encounter are run by domestic institutions. The second holds that respondents default to assuming a domestic sponsor when none is named, so that the baseline, like the Chinese condition, underreports out of

still vary with interviewer ethnicity.

[13]The differences in the list-experiment estimates between the U.S. and Chinese conditions, and between the U.S. condition and the no-sponsor baseline, are statistically significant at the 0.1 level. Full results appear in Appendix M.

wariness about how the data might be used, and only an explicit foreign cue reassures them enough to answer candidly. For a behavior as politically exposed as VPN use, which is widespread yet legally restricted, this reading is plausible.

Our data cannot definitively adjudicate between these interpretations. Separating underreporting driven by fear from overreporting driven by impression management would require additional experimental variation, such as manipulating the perceived consequences of disclosure or varying sponsor nationality in ways that hold cosmopolitan signaling constant while varying political risk. We leave that test to future work. What is clear is that sponsorship effects persist even within the supposedly protective anonymity of a list experiment, an important methodological point: survey context shapes responses through channels that standard bias-reduction techniques may not fully address.

6 Discussion and Conclusion

This study examines how researchers' national institutional affiliation shapes survey responses in authoritarian contexts. Through a large-scale experiment in China, we find that sponsor nationality affects responses across diverse question types. In particular, sponsorship effects extend well beyond political questions to encompass cultural preferences, suggesting that respondents engage in self-presentation management rather than concealment of regime-critical views alone. Moreover, sponsorship effects persist within list experiments, revealing that perceived audience shapes responses through mechanisms that standard anonymity protections cannot fully address.

Our study has three main limitations. First, for any given item we cannot determine how much of the gap reflects underreporting driven by fear under Chinese sponsorship versus overreporting driven by impression management under U.S. sponsorship, or some combination of the two. Distinguishing these mechanisms at the item level remains an important avenue for future research. Second, platform restrictions prevented us from measuring trust in the most politically sensitive institutions, the central government and top leadership. Whether sponsorship effects intensify or attenuate for these high-stakes evaluations remains an open question.

Third, our findings are based on Chinese respondents, and we cannot directly establish whether sponsor nationality effects generalize to other authoritarian or semi-democratic contexts. We believe the results are substantively important on their own terms: China's 1.4 billion citizens constitute one of the world's largest survey populations, and documenting bias in this context alone has major implications for a large body of existing research. That said, the mechanisms we note, including political fear, impression management, and reflexive caution toward state-aligned institutions, are not unique to China. Analogous dynamics may operate wherever citizens face meaningful consequences for expressing heterodox views to perceived domestic authorities, or where foreign affiliation carries distinct social meaning. We encourage researchers working in other authoritarian contexts to replicate and extend this design, recognizing that the magnitude and direction of effects may vary with the local political environment.

Our findings carry important implications for public-opinion research in authoritarian regimes. For scholars conducting descriptive surveys that aim to characterize population attitudes or behaviors, disclosing sponsor nationality can introduce bias. Although Crabtree, Kern and Pietryka (2022) suggests that obscuring the sponsor may reduce such bias, we are skeptical that it eliminates response bias. As Tannenberg (2022) shows, respondents in authoritarian settings tend to assume the government is behind a survey and adjust their answers accordingly. Our own data point the same way: under the list experiment, the no-sponsor baseline estimate of VPN use matches the Chinese-sponsorship estimate rather than the U.S. one, suggesting that when no sponsor is named respondents default to assuming a domestic one and answer as guardedly as under explicit Chinese sponsorship.

We therefore recommend treating sponsorship effects as a design consideration rather than an afterthought. Researchers should anticipate them at the design stage, especially where the sponsor's nationality could plausibly shift answers to particular items through either mechanism we identify, impression management or fear. When reporting findings, they should acknowledge the potential for bias and discuss how it could affect their conclusions. Cross-national collaborations face this challenge most acutely, since they involve multiple institutional partners. We stop short of recommending

that all surveys randomize sponsorship, but transparency about institutional partnerships is essential: where feasible, pilot studies can test whether different partners trigger different response patterns, and at a minimum collaborative teams should discuss potential sponsor effects and document their disclosure decisions in published work.

For survey experiments the implication is different. Because a sponsor effect shifts the level of responses and the sponsor is held constant across treatment conditions, any such shift is common to treatment and control and differences out of the average treatment effect; randomization delivers this even when susceptibility to the sponsor varies across respondents. Experimental designs therefore remain internally valid. Their external validity, however, can break down when the sponsor's nationality interacts with the treatment, that is, when the treatment effect itself depends on who respondents believe is asking. The risk is greatest for treatments that carry national or cultural meaning. Consider an information experiment that presents a domestic policy and gauges its effect on support for the government, or a vignette featuring a foreign actor. In both cases, respondents may engage the content differently under a U.S. than a domestic sponsor, so the estimated effect is specific to that sponsor and need not reproduce under another.

Similar logic applies to quasi-experiments and natural experiments. We concur with Li et al. (2026) that a sound identification strategy can preserve internal validity, but this holds only if treatment assignment is orthogonal to sponsorship susceptibility. Randomization makes this automatic; when assignment is not random, it becomes an assumption that must be defended, since a sponsorship effect correlated with assignment would not difference out. This could arise, for instance, in a place-based natural experiment whose treated units are coastal, internationally connected regions, if their residents are also the most responsive to a foreign sponsor. Our design cannot test this directly, but the broad, demographically uniform character of the sponsor effect is at least consistent with its operating as a stable offset rather than a subgroup-specific phenomenon.

In conclusion, where sponsor cues are prominently displayed in surveys, either through consent forms or institutional banners, potential biases can be introduced without the realization of researchers. The indirect-questioning remedy we test does not eliminate the resulting bias, and our results give

reason to doubt that related anonymity-based fixes would fully succeed either. Progress therefore requires transparency about measurement limitations, explicit acknowledgment of assumptions, and careful consideration of how violations would affect substantive conclusions. Only through such vigilance can survey research continue to illuminate public opinion where citizens face genuine risks for speaking truthfully.

Competing interests

The author(s) declare none.

References

Adida, Claire L., Karen E. Ferree, Daniel N. Posner and Amanda Lea Robinson. 2016. "Who's Asking? Interviewer Coethnicity Effects in African Survey Data." *Comparative Political Studies* 49(12):1630–1660.

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

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

Chen, Xiaomei. 1995. *Occidentalism: A Theory of Counter-discourse in Post-Mao China*. Oxford University Press.

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

Chu, Yun-han, Larry Diamond, Andrew J. Nathan and Doh Chull Shin. 2008. *How East Asians View Democracy*. Columbia University Press.

Corstange, D. 2014. "Foreign-Sponsorship Effects in Developing-World Surveys: Evidence from a Field Experiment in Lebanon." *Public Opinion Quarterly* 78(2):474–484.

Crabtree, Charles, Holger L. Kern and Matthew T. Pietryka. 2022. "Sponsorship Effects in Online Surveys." *Political Behavior* 44(1):257–270.

Di Maio, Michele and Nathan Fiala. 2020. "Be Wary of Those Who Ask: A Randomized Experiment on the Size and Determinants of the Enumerator Effect." *The World Bank Economic Review* 34(3):654–669.

Diaz, Gustavo. 2024. "Assessing the Validity of Prevalence Estimates in Double List Experiments." *Journal of Experimental Political Science* 11(2):162–174.

Droitcour, Judith, Rachel A Caspar, Michael L Hubbard, Teresa L Parsley, Wendy Visscher and Trena M Ezzati. 2004. "The item count technique as a method of indirect questioning: A review of its development and a case study application." *Measurement errors in surveys* pp. 185–210.

Edwards, Michelle L., Don A. Dillman and Jolene D. Smyth. 2014. "An Experimental Test of the Effects of Survey Sponsorship on Internet and Mail Survey Response." *Public Opinion Quarterly* 78(3):734–750.

Gengler, Justin J., Mark Tessler, Russell Lucas and Jonathan Forney. 2021. "'Why Do You Ask?' The Nature and Impacts of Attitudes towards Public Opinion Surveys in the Arab World." *British Journal of Political Science* 51(1):115–136.

Glynn, Adam N. 2013. "What can we learn with statistical truth serum? Design and analysis of the list experiment." *Public Opinion Quarterly* 77(S1):159–172.

Holbrook, Allyson L. and Jon A. Krosnick. 2010. "Social Desirability Bias in Voter Turnout Reports: Tests Using the Item Count Technique." *Public Opinion Quarterly* 74(1):37–67.

Jee, Haemin and Tongtong Zhang. 2025. "Oppose Autocracy without Support for Democracy: A Study of Non-Democratic Critics in China." *Perspectives on Politics* 23(3):865–884.

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

Johnston, Alastair Iain. 2017. "Is Chinese Nationalism Rising? Evidence from Beijing." *International Security* 41(3):7–43.

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

Kuran, Timur. 1991. "Now out of Never: The Element of Surprise in the East European Revolution of 1989." *World Politics* 44(1):7–48.

Lee, David S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *The Review of Economic Studies* 76(3):1071–1102.

Leeper, Thomas J. and Emily A. Thorson. 2020. "Should We Worry About Sponsorship-Induced Bias in Online Political Science Surveys?" *Journal of Experimental Political Science* 7(3):209–217.

Lei, Xuchuan and Jie Lu. 2017. "Revisiting Political Wariness in China's Public Opinion Surveys: Experimental Evidence on Responses to Politically Sensitive Questions." *Journal of Contemporary China* .

Li, Ding, Xiaobo Lü, Shuang Ma and Wenhui Yang. 2026. "The Shadow of Social Desirability Bias: Evidence from Reassessing the Sources of Political Trust in China." *Political Science Research and Methods* .

Li, Lianjiang. 2016. "Reassessing Trust in the Central Government: Evidence from Five National Surveys." *The China Quarterly* 225:100–121.

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

Mutz, Diana C, Robin Pemantle and Philip Pham. 2019. "The perils of balance testing in experimental design: Messy analyses of clean data." *The American Statistician* 73(1):32–42.

Nicholson, Stephen P. and Haifeng Huang. 2023. "Making the List: Reevaluating Political Trust and Social Desirability in China." *American Political Science Review* 117(3):1158–1165.

Pan, Jennifer and Yiqing Xu. 2017. "China's Ideological Spectrum." *Journal of Politics* 80(1):254–273.

Rainey, Carlisle. 2014. "Arguing for a negligible effect." *American Journal of Political Science* 58(4):1083–1091.

Ratigan, Kerry and Leah Rabin. 2020. "Re-evaluating political trust: The impact of survey nonresponse in rural China." *The China Quarterly* 243:823–838.

Robinson, Darrel and Marcus Tannenberg. 2019. "Self-censorship of regime support in authoritarian states: Evidence from list experiments in China." *Research & Politics* 6(3):1–9.

Shen, Xiaoxiao and Rory Truex. 2021. "In search of self-censorship." *British Journal of Political Science* 51(4):1672–1684.

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

Terracciano, A., A. M. Abdel-Khalek, N. Ádám, L. Adamovová, C.-k. Ahn, H.-n. Ahn, B. M. Alansari, L. Alcalay, J. Allik, A. Angleitner, A. Avia, L. E. Ayearst, C. Barbaranelli, A. Beer, M. A. Borg-Cunen, D. Bratko, M. Brunner-Sciarra, L. Budzinski, N. Camart, D. Dahourou, F. De Fruyt, M. P. de Lima, G. E. H. del Pilar, E. Diener, R. Falzon, K. Fernando, E. Ficková, R. Fischer, C. Flores-Mendoza, M. A. Ghayur, S.

Gülgöz, B. Hagberg, J. Halberstadt, M. S. Halim, M. Hřebíčková, J. Humrichouse, H. H. Jensen, D. D. Jocic, F. H. Jónsson, B. Khoury, W. Klinkosz, G. Knežević, M. A. Lauri, N. Leibovich, T. A. Martin, I. Marušić, K. A. Mastor, D. Matsumoto, M. McRorie, B. Meshcheriakov, E. L. Mortensen, M. Munyae, J. Nagy, K. Nakazato, F. Nansubuga, S. Oishi, A. O. Ojedokun, F. Ostendorf, D. L. Paulhus, S. Pelevin, J.-M. Petot, N. Podobnik, J. L. Porrata, V. S. Pramila, G. Prentice, A. Realo, N. Reátegui, J.-P. Rolland, J. Rossier, W. Ruch, V. S. Rus, M. L. Sánchez-Bernardos, V. Schmidt, S. Sciculna-Calleja, A. Sekowski, J. Shakespeare-Finch, Y. Shimonaka, F. Simonetti, T. Sineshaw, J. Siuta, P. B. Smith, P. D. Trapnell, K. K. Trobst, L. Wang, M. Yik, A. Zupančič and R. R. McCrae. 2005. "National Character Does Not Reflect Mean Personality Trait Levels in 49 Cultures." *Science (New York, N.Y.)* 310(5745):96–100.

Truex, Rory. 2022. "Political Discontent in China Is Associated with Isolating Personality Traits." *The Journal of Politics* 84(4):2172–2187.

Wang, Zheng. 2008. "National Humiliation, History Education, and the Politics of Historical Memory: Patriotic Education Campaign in China." *International Studies Quarterly* 52(4):783–806.

White, Ariel, Anton Strezhnev, Christopher Lucas, Dominika Kruszewska and Connor Huff. 2018. "Investigator Characteristics and Respondent Behavior in Online Surveys." *Journal of Experimental Political Science* 5(1):56–67.

NOT FOR PUBLICATION

Supplementary Materials

"Reading the Room: Nationality-of-Sponsor Effects on Survey Responses - Evidence from China"

Appendix Table of Contents

Appendix A Respondent Demographics

Table A.1: Demographic Comparison

Category	Our Survey (%)	CFPS Internet (%)	CFPS All (%)
Gender			
Female	63.8	48.5	50.2
Male	36.2	51.5	49.8
Age Group			
18 - 30	58.2	35.9	26.9
30 - 45	34.8	32.7	24.5
46 - 55	5.2	16.5	16.4
Above 56	1.8	14.9	32.2
Education			
≤ Primary school	0.1	24.5	36.9
Junior high school	0.4	32.2	29.8
Senior high school	3.1	20.2	16.3
3-year college	8.2	11.6	8.6
≥ 4-year college	88.1	11.5	8.4
CCP Member			
No	77.3	89.2	90.3
Yes	22.7	10.8	9.7
Region			
Eastern	32.8	21.5	21.6
North	15.1	21.3	19.5
North Eastern	6.3	8.8	8.5
North Western	5.1	5.7	5.5
South Central	31.4	27.9	27.5
South Western	9.4	14.8	17.3

Note: We compared our respondent sample to CFPS (China Family Panel Studies), a national survey of China using multi-stage probability sampling. Following Nicholson and Huang (2023), we use its "resampled sample" (subsample=1), which is nationally representative by design (see CFPS User's Manual, 3rd Edition, http://isss.pku.edu.cn/cfps/docs/20200315092524928116.pdf). We define "Internet Users" as those who accessed the Internet using mobile devices or personal computers (qu201 = 1 or qu202 = 1).

Appendix B Treatment Example and Institution Banners

Appendix B.1 Treatment Example

Figure A.1: Treatment Example

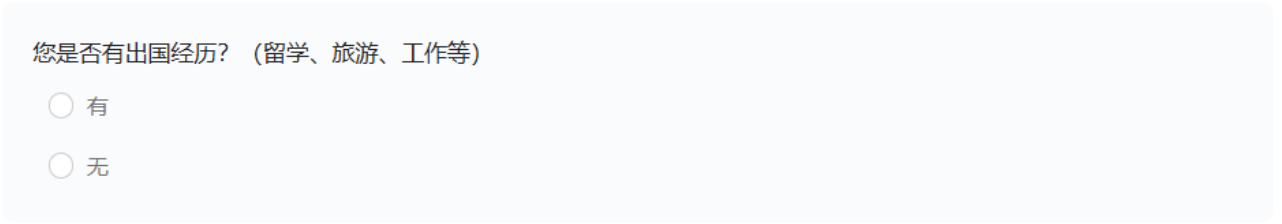
Translated Consent Script:

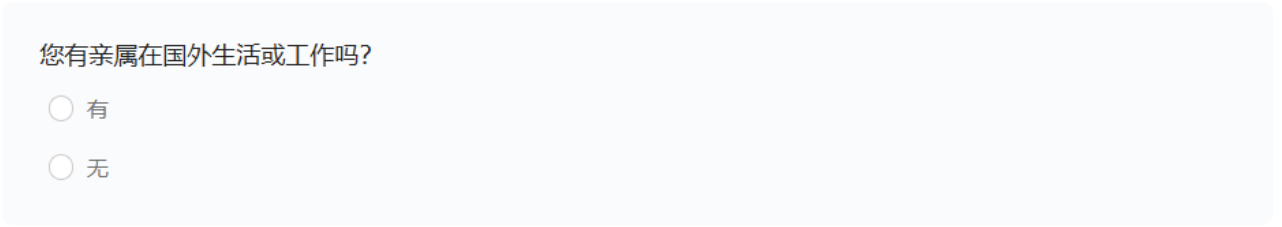
We are a research team affiliated with Harvard University in the United States, conducting a survey on Chinese internet users' social attitudes and online behavior. The following questionnaire will ask about your views on several social issues and your related behaviors. If you do not wish to participate in this survey, you may choose to exit now. If you are willing to continue, please click "Continue"; if you do not wish to participate, please click "Exit."

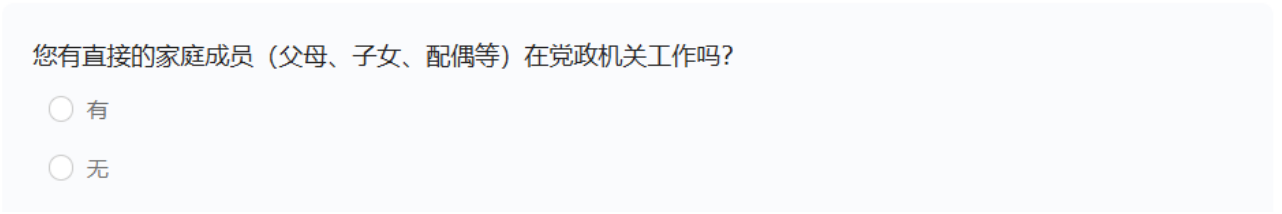
Continue

Exit

Figure A.2: Survey Page Example







Appendix B.2 Institution Logos

Figure A.3: US Embassy

Figure A.4: Tsinghua University

Figure A.5: Harvard University

Figure A.6: Loyola University Chicago

Figure A.7: Hebei Normal University

Figure A.8: CCP Youth League

Appendix C Balance Check

Because our central comparison is between U.S. and Chinese sponsorship, we assess balance with that contrast in focus while also certifying the full design. For each of eight pre-treatment covariates, Table A.2 reports the U.S. and Chinese group means, the U.S. minus China difference with its HC1 robust standard error and standardized difference, and a joint test of mean-independence across all seven arms. Seven of the eight covariates are balanced across the U.S. and Chinese conditions, with standardized differences at or below 0.06, and the joint tests across arms reveal no systematic imbalance.

The one difference on this focal contrast is education: respondents in the U.S. conditions are about three percentage points less likely to hold a four-year college degree (standardized difference -0.10). A compositional difference this small could bias our estimates only if education predicted the outcomes implausibly strongly, because any such bias is the product of a three-point difference in composition and the effect of education on the outcome. Its direction is moreover conservative for our main findings: the U.S. arm is, if anything, slightly less educated, which works against the pro-U.S. and cosmopolitan shifts we report. Because assignment is random, selecting covariates for adjustment on the basis of an observed imbalance has no statistical basis and would not improve the credibility of our estimates (Mutz, Pemantle and Pham, 2019), so we report unadjusted estimates throughout.

Table A.2: Covariate Balance: U.S. versus Chinese Sponsorship

Covariate	U.S.	China	Diff. (U.S. − China)	Joint p (all arms)
Age	30.659	30.518	0.141 (0.303)	0.434
			[0.017]	
Male	0.363	0.358	0.004 (0.018)	0.997
			[0.009]	
CCP Member	0.225	0.229	-0.003 (0.015)	0.522
			[-0.008]	
Income	2.218	2.179	0.039 (0.028)	0.043
			[0.052]	
Internet Use	2.908	2.866	0.042 (0.033)	0.748
			[0.047]	
Public Sector	0.262	0.281	-0.019 (0.016)	0.636
			[-0.042]	
Local Hukou	0.672	0.657	0.014 (0.017)	0.898
			[0.031]	
College Degree	0.867	0.898	-0.031** (0.012)	0.257
			[-0.098]	

Note: Entries in the U.S. and China columns are group means. The difference column reports the U.S. minus China difference with HC1 robust standard errors in parentheses; the standardized difference (difference divided by the pooled standard deviation) appears in brackets on the line below. "Joint p (all arms)" is from a Wald test that the covariate is mean-independent of assignment across all seven arms (HC1). "College Degree" denotes holding a four-year college degree. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Appendix D Selection into Treatment: Lee Bounds

Because each respondent saw the sponsoring institution on the consent screen before entering the survey, differential refusal at that stage could in principle select who appears in each condition and bias the comparison across sponsors. In practice this selection is minimal. Four respondents declined to proceed after seeing a U.S. sponsor (two under the U.S. Embassy, one under Harvard, and one under Loyola), and none declined after a Chinese sponsor, a difference in refusal of roughly 0.3 percentage points. We bound its consequences using the trimming bounds of Lee (2009), treating survey entrance as the selection outcome and U.S. sponsorship as the treatment. Because the higher-retention Chinese arm exceeds the U.S. arm by only about 0.27 percent of cases, the trimming proportion is correspondingly small, and the bounds reported in Table A.3 are essentially indistinguishable from the point estimates for every outcome. Selection into treatment therefore cannot account for the sponsorship effects we report.

Table A.3: Selection into Treatment: Lee (2009) Trimming Bounds

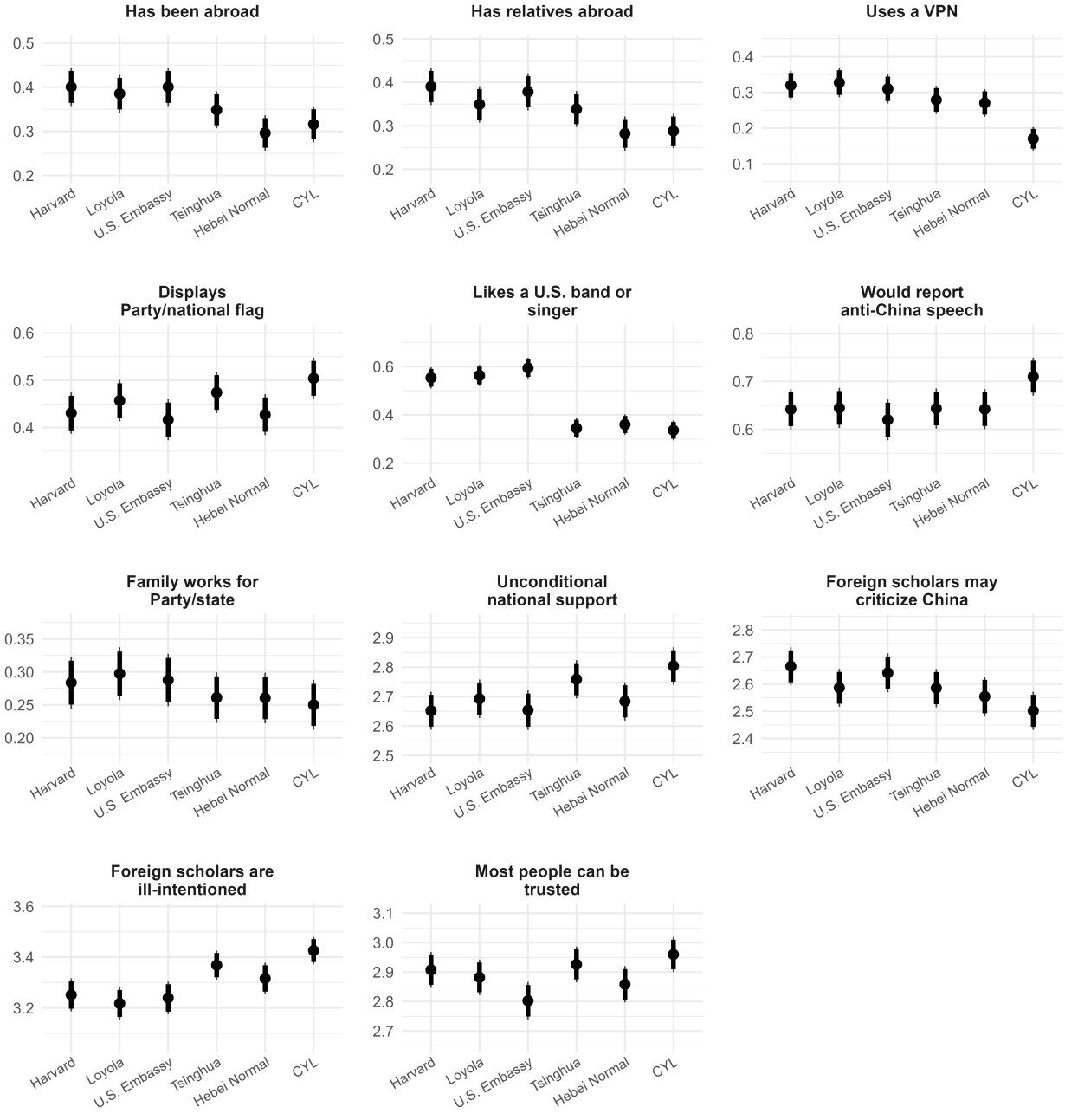
Outcome	Estimate	Lower Bound	Upper Bound
Behaviors			
Having Been Abroad	0.075	0.074	0.077
Relatives Abroad	0.070	0.069	0.071
VPN Use	0.079	0.079	0.081
Party/National Flag Display	-0.034	-0.035	-0.032
Report Anti-China Speech	-0.030	-0.031	-0.029
Family Works for Party/State	0.032	0.032	0.034
Liked US Band/Singer	0.223	0.222	0.225
Statement Agreement			
Unconditional National Support	-0.083	-0.087	-0.079
Foreign Scholars' Right to Criticize	0.084	0.080	0.088
Foreign Scholars are Malicious	-0.134	-0.140	-0.132
Most People can be Trusted	-0.051	-0.056	-0.048
International Entities			
United States	7.747	7.677	7.930
Taiwan	2.802	2.675	2.941
European Union	3.962	3.841	4.108
Russia	1.704	1.537	1.803
Japan	3.708	3.662	3.915
Domestic Institutions			
Neighborhood Committee	0.781	0.613	0.880
Courts	-0.031	-0.230	0.030
Local Government Officials	0.281	0.116	0.382
People's Daily	0.380	0.173	0.439

Note: Lee (2009) trimming bounds on the effect of U.S. versus Chinese sponsorship, addressing differential selection at consent. Four respondents declined after a U.S. sponsor (Embassy 2, Harvard 1, Loyola 1) and none after a Chinese sponsor, an implied trimming proportion of 0.27 percent (four cases trimmed from the higher-retention Chinese arm for each outcome). Bounds assume U.S. sponsorship only weakly affects participation. The point estimate is the unadjusted difference in means (U.S. minus Chinese).

Appendix E Main Experiment: Group Mean by Conditions

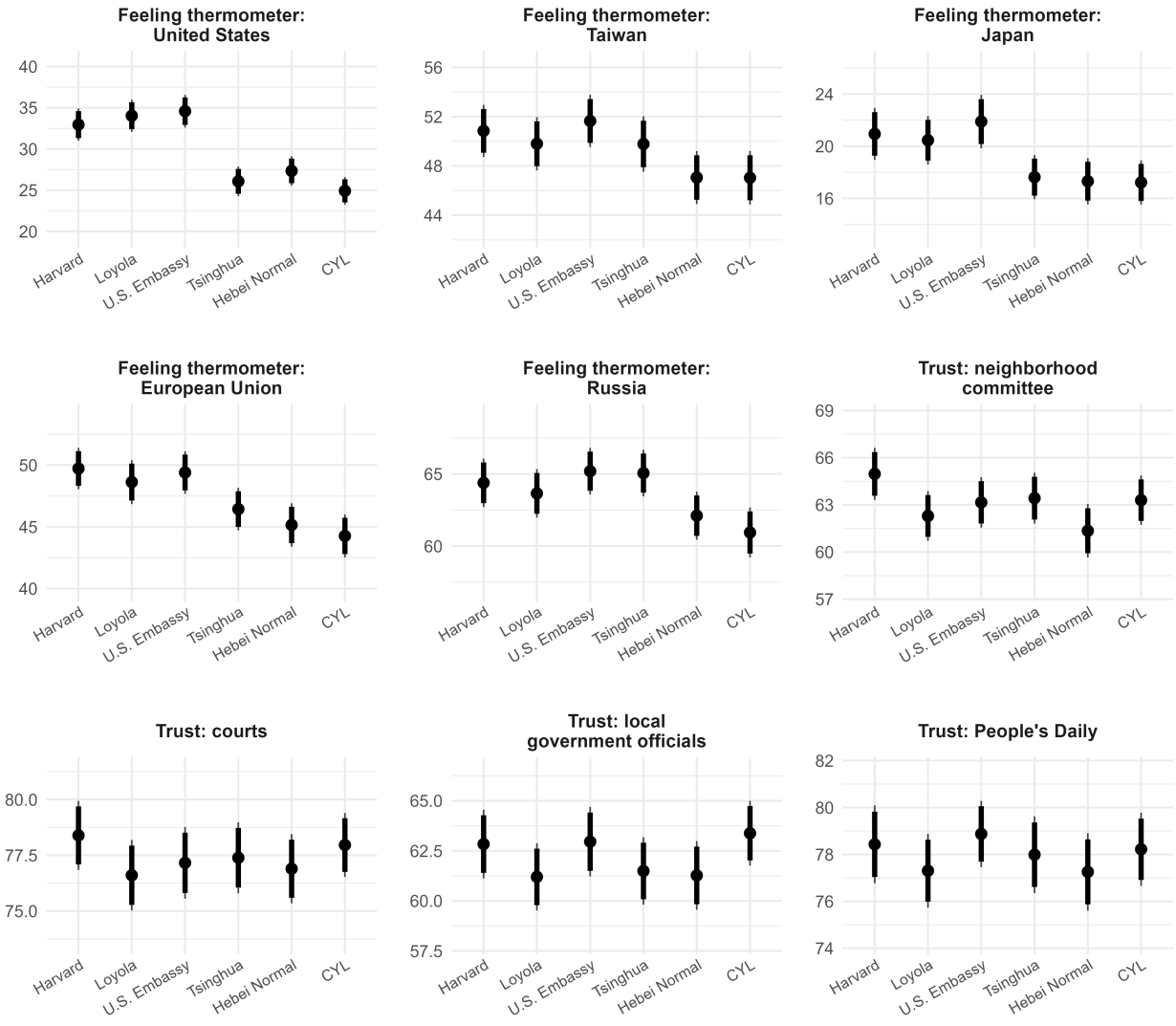
In the main text, we report the unadjusted U.S.-versus-Chinese difference for each outcome from an OLS regression with no covariates. For transparency, we present here the disaggregated group means for all six sponsor conditions.

Figure A.9: Group Means by Treatment Conditions 1



Note: Thick and thin lines indicate 90% and 95% confidence intervals, respectively.

Figure A.10: Group Means by Treatment Conditions 2



Note: Thick and thin lines indicate 90% and 95% confidence intervals, respectively.

Appendix F Robustness Check: Multiple Hypothesis p Value Correction

Because our survey experiment evaluates the effect of survey sponsorship nationality across 20 distinct outcome variables, there is a risk of Type I errors (false positives) driven by multiple comparisons. To ensure our results are not artifacts of multiple testing, we conduct multiple hypothesis testing (MHT) corrections across all outcome variables. In Table A.4, we report the original robust p-values alongside p-values adjusted using the Benjamini-Hochberg (BH) procedure, which controls the False Discovery Rate (FDR), and the Holm-Bonferroni method, which strictly controls the Family-Wise Error Rate (FWER). The results demonstrate that our core substantive findings remain robust. While a few outcomes attenuate, the effects of U.S. sponsorship on key behavioral indicators (such as reported VPN usage and having traveled abroad) and geopolitical attitudes (such as feeling thermometers toward the United States and its allies) remain statistically significant at the 0.05 level or tighter, even under the highly conservative Holm correction.

Table A.4: Treatment Effects with Multiple Hypothesis Testing Corrections

Outcome	Estimate (SE)	Original p-value	FDR (BH) p-value	Holm p-value
Behaviors				
Having Been Abroad	0.075*** (0.017)	< 0.001	< 0.001	< 0.001
Relatives Abroad	0.070*** (0.017)	< 0.001	< 0.001	< 0.001
VPN Use	0.079*** (0.016)	< 0.001	< 0.001	< 0.001
Party/National Flag Display	-0.034+ (0.018)	0.064	0.085	0.384
Report Anti-China Speech	-0.030+ (0.017)	0.089	0.111	0.443
Family Works for Party/State	0.032* (0.016)	0.046	0.067	0.367
Liked US Band/Singer	0.223*** (0.018)	< 0.001	< 0.001	< 0.001
Statement Agreement				
Unconditional National Support	-0.083** (0.027)	0.002	0.005	0.026
Foreign Scholars' Right to Criticize	0.084** (0.030)	0.005	0.008	0.046
Foreign Scholars are Malicious	-0.134*** (0.025)	< 0.001	< 0.001	< 0.001
Most People can be Trusted	-0.051* (0.025)	0.047	0.067	0.367
International Entities				
United States	7.747*** (0.774)	< 0.001	< 0.001	< 0.001
Taiwan	2.802** (0.902)	0.002	0.004	0.023
European Union	3.962*** (0.723)	< 0.001	< 0.001	< 0.001
Russia	1.704* (0.696)	0.014	0.024	0.130
Japan	3.708*** (0.770)	< 0.001	< 0.001	< 0.001
Domestic Institutions				
Neighborhood Committee	0.781 (0.677)	0.249	0.293	0.997
Courts	-0.031 (0.647)	0.962	0.962	1.000
Local Government Officials	0.281 (0.705)	0.690	0.726	1.000
People's Daily	0.380 (0.659)	0.564	0.627	1.000

Note: FDR (BH) controls the False Discovery Rate; Holm controls the Family-Wise Error Rate. Both are applied globally across all 20 outcomes. HC1 robust standard errors in parentheses. Stars correspond to the unadjusted p-values: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Appendix G Robustness Check: Weighted Results

To account for the demographic skew inherent in our online convenience sample, we re-estimated our main models using post-stratification weights. We applied an iterative proportional fitting (raking) algorithm to adjust our sample marginal distributions to match the nationally representative general adult population benchmark derived from the 2022 China Family Panel Studies (CFPS). The weighting procedure balanced the sample across three demographic variables: gender, age group (18 to 30, 30 to 45, 46 to 55, and above 56), and education level (holding a four-year college degree or no four-year college degree). To mitigate this variance inflation and prevent a small number of respondents from disproportionately driving the estimates, we trimmed the resulting weights at the 5th and 95th percentiles before calculating the robust standard errors.

Table A.5 presents the unweighted and weighted treatment effects. As expected, applying raking weights to the highly skewed online sample introduces variance inflation. Consequently, several marginally significant effects on secondary outcomes lose their statistical significance in the weighted models. However, the majority of the treatment effects maintain statistical significance despite this precision penalty.

Table A.5: Robustness Check: Unweighted vs. Weighted Treatment Effects

Outcome	Unweighted	Weighted (CFPS All)
Behaviors		
Having Been Abroad	0.075*** (0.017)	0.054+ (0.031)
Relatives Abroad	0.070*** (0.017)	0.063* (0.032)
VPN Use	0.079*** (0.016)	0.118*** (0.028)
Party/National Flag Display	-0.034+ (0.018)	-0.034 (0.034)
Report Anti-China Speech	-0.030+ (0.017)	-0.022 (0.034)
Family Works for Party/State	0.032* (0.016)	0.010 (0.029)
Liked US Band/Singer	0.223*** (0.018)	0.223*** (0.033)
Statement Agreement		
Unconditional National Support	-0.083** (0.027)	-0.067 (0.051)
Foreign Scholars' Right to Criticize	0.084** (0.030)	0.090 (0.056)
Foreign Scholars are Malicious	-0.134*** (0.025)	-0.093+ (0.049)
Most People can be Trusted	-0.051* (0.025)	0.034 (0.047)
International Entities		
United States	7.747*** (0.774)	7.264*** (1.522)
Taiwan	2.802** (0.902)	2.560 (1.703)
European Union	3.962*** (0.723)	3.077* (1.347)
Russia	1.704* (0.696)	-0.145 (1.332)
Japan	3.708*** (0.770)	2.294+ (1.371)
Domestic Institutions		
Neighborhood Committee	0.781 (0.677)	1.667 (1.221)
Courts	-0.031 (0.647)	0.259 (1.126)
Local Government Officials	0.281 (0.705)	0.164 (1.267)
People's Daily	0.380 (0.659)	0.325 (1.183)

Note: HC1 robust standard errors in parentheses. Weighted estimates use raking post-stratification to CFPS general-population marginals for gender, age, and education, with weights trimmed at the 5th and 95th percentiles. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Appendix H Robustness Check: Excluding Political Institutions

To assess whether the treatment effects are driven by the political-institution conditions (the U.S. Embassy and the Communist Youth League), we re-estimate each effect on the four academic conditions alone (Harvard and Loyola versus Tsinghua and Hebei Normal), dropping the two governmental conditions, using the same specification as in the main text and present the results in Table A.6. The core effects are robust: foreign exposure (overseas travel, relatives abroad, VPN use), the preference for U.S. music, warmth toward the United States, the European Union, and Japan, and the view that foreign scholars are not malicious all retain their sign and significance, with magnitudes close to the full sample. Several smaller and more overtly regime-signaling estimates attenuate once the governmental sponsors are removed: flag display, reporting government-critical speech, unconditional national support, and warmth toward Russia lose significance, and the Taiwan thermometer weakens to the 10% level. This pattern indicates that these particular responses are partly specific to governmental sponsorship. The broad foreign-versus-domestic gap, however, is not an artifact of the governmental conditions: sponsor nationality continues to shift responses among academic institutions.

Table A.6: Robustness Check: Excluding Political-Institution Conditions

Outcome	Original	Excluding Political Institutions
Behaviors		
Having Been Abroad	0.075*** (0.017)	0.070** (0.021)
Relatives Abroad	0.070*** (0.017)	0.059** (0.021)
VPN Use	0.079*** (0.016)	0.049* (0.020)
Party/National Flag Display	-0.034+ (0.018)	-0.007 (0.022)
Report Anti-China Speech	-0.030+ (0.017)	0.001 (0.021)
Family Works for Party/State	0.032* (0.016)	0.030 (0.020)
Liked US Band/Singer	0.223*** (0.018)	0.206*** (0.022)
Statement Agreement		
Unconditional National Support	-0.083** (0.027)	-0.049 (0.033)
Foreign Scholars' Right to Criticize	0.084** (0.030)	0.056 (0.036)
Foreign Scholars are Malicious	-0.134*** (0.025)	-0.108*** (0.032)
Most People can be Trusted	-0.051* (0.025)	0.002 (0.031)
International Entities		
United States	7.747*** (0.774)	6.791*** (0.956)
Taiwan	2.802** (0.902)	1.903+ (1.109)
European Union	3.962*** (0.723)	3.377*** (0.884)
Russia	1.704* (0.696)	0.436 (0.850)
Japan	3.708*** (0.770)	3.229*** (0.935)
Domestic Institutions		
Neighborhood Committee	0.781 (0.677)	1.239 (0.839)
Courts	-0.031 (0.647)	0.351 (0.799)
Local Government Officials	0.281 (0.705)	0.631 (0.867)
People's Daily	0.380 (0.659)	0.244 (0.831)

Note: Each cell is the OLS estimate of the U.S.-versus-Chinese sponsorship effect with HC1 robust standard errors in parentheses. The Original column uses all six sponsor conditions; the Excluding column restricts the sample to the four academic conditions (Harvard, Loyola, Tsinghua, Hebei Normal). + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Appendix I Test of Prestige and Type Effects

Because our design crosses nationality with institutional type and prestige in a symmetric structure, these two dimensions are balanced across the U.S.–Chinese contrast by construction and cannot confound the pooled nationality effect. We nonetheless estimate the type and prestige contrasts directly, to characterize whether either dimension moves responses in its own right. Table A.7 reports, for each outcome, the standardized nationality effect alongside the prestige contrast (elite versus non-elite, holding nationality fixed) and the type contrast (governmental versus academic, holding nationality fixed), each with an equivalence bound; following the negligible-effect framework of Rainey (2014), the bound is the 90% confidence-interval endpoint farthest from zero.

On the outcomes that nationality shifts most, the type and prestige contrasts are generally small and bounded beneath the nationality effect. We do not claim they are zero: a few prestige contrasts are non-trivial—most clearly warmth toward Russia and relatives abroad—and the governmental-type contrast lowers reported VPN use by about as much as nationality does. But these exceptions are scattered, sharing neither a common direction nor a common domain, and do not constitute an alternative account of the broad, systematic shift that sponsor nationality produces.

Table A.7: Prestige and Type Equivalence Relative to the Nationality Effect (Standardized)

		Prestige (elite − non-elite)		Type (gov. − academic)	
Outcome	Nationality $\hat{d}$	$\hat{\delta}$ (90% CI)	Excl. $\lvert\delta\rvert >$	$\hat{\delta}$ (90% CI)	Excl. $\lvert\delta\rvert >$
Behaviors					
Having Been Abroad	0.157*** (0.036)	0.07 [-0.00, 0.14]	0.14	0.00 [-0.06, 0.06]	0.06
Relatives Abroad	0.147*** (0.036)	0.10 [0.03, 0.18]	0.18	-0.01 [-0.08, 0.05]	0.08
VPN Use	0.176*** (0.036)	0.00 [-0.07, 0.08]	0.08	-0.13 [-0.19, -0.07]	0.19
Party/National Flag Display	-0.068+ (0.036)	0.02 [-0.05, 0.09]	0.09	0.03 [-0.04, 0.09]	0.09
Report Anti-China Speech	-0.062+ (0.037)	-0.00 [-0.08, 0.07]	0.08	0.05 [-0.02, 0.11]	0.11
Family Works for Party/State	0.073* (0.037)	-0.01 [-0.09, 0.06]	0.09	-0.02 [-0.08, 0.05]	0.08
Liked US Band/Singer	0.448*** (0.036)	-0.02 [-0.10, 0.05]	0.10	0.02 [-0.04, 0.08]	0.08
Statement Agreement					
Unconditional National Support	-0.111** (0.036)	0.02 [-0.05, 0.10]	0.10	0.04 [-0.02, 0.11]	0.11
Foreign Scholars' Right to Criticize	0.103** (0.036)	0.07 [-0.01, 0.14]	0.14	-0.03 [-0.10, 0.03]	0.10
Foreign Scholars are Malicious	-0.191*** (0.036)	0.06 [-0.01, 0.14]	0.14	0.06 [0.00, 0.13]	0.13
Most People can be Trusted	-0.073* (0.036)	0.07 [-0.01, 0.14]	0.14	-0.02 [-0.08, 0.05]	0.08
International Entities					
United States	0.359*** (0.036)	-0.05 [-0.13, 0.02]	0.13	-0.02 [-0.08, 0.05]	0.08
Taiwan	0.113** (0.036)	0.08 [0.00, 0.15]	0.15	-0.00 [-0.06, 0.06]	0.06
European Union	0.199*** (0.036)	0.06 [-0.01, 0.13]	0.13	-0.03 [-0.10, 0.03]	0.10
Russia	0.089* (0.036)	0.10 [0.02, 0.17]	0.17	-0.04 [-0.10, 0.03]	0.10
Japan	0.175*** (0.036)	0.02 [-0.05, 0.09]	0.09	0.02 [-0.04, 0.09]	0.09
Domestic Institutions					
Neighborhood Committee	0.042 (0.037)	0.13 [0.05, 0.20]	0.20	0.01 [-0.05, 0.08]	0.08
Courts	-0.002 (0.037)	0.06 [-0.01, 0.14]	0.14	0.01 [-0.05, 0.08]	0.08
Local Government Officials	0.015 (0.037)	0.05 [-0.03, 0.12]	0.12	0.08 [0.01, 0.14]	0.14
People's Daily	0.021 (0.037)	0.05 [-0.02, 0.13]	0.13	0.04 [-0.02, 0.11]	0.11

Note: All outcomes are standardized over the six-arm sponsor sample (units = SD). The nationality benchmark $\hat{d}$ is the U.S.-versus-Chinese contrast from a bivariate regression (six arms); the prestige contrast is the coefficient on an elite-institution indicator (Harvard, Tsinghua) holding nationality fixed, estimated on the four academic arms; the type contrast is the coefficient on a governmental-sponsor indicator (U.S. Embassy, Communist Youth League) holding nationality fixed, estimated on all six arms. HC1 robust standard errors in parentheses. Following Rainey (2014), "Excl. $\lvert\delta\rvert >$" is the 90% CI endpoint farthest from zero (the contrast the data are inconsistent with). No equivalence margin is imposed. Stars on $\hat{d}$: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Appendix J Treatment Heterogeneity

We conduct analysis on treatment heterogeneity using all eight pre-treatment moderators: age, gender, CCP membership, family income, Internet use, public-sector employment, local residence registration (Hukou), and education. For each moderator, we include an interaction term between the moderator and the U.S. treatment indicator. The reference categories are: below-median age, female, non-CCP members, below-median income, below-median Internet use, non-public-sector employees, respondents without local hukou, and respondents without a four-year college degree. Across all moderators, we do not observe a consistent or systematic pattern of heterogeneous treatment effects.

Table A.8: Heterogeneity Analysis: All Outcomes

Outcome	Age	Gender	CCP	Income	Internet use	Public sector	Local hukou	Education
Behaviors								
Been abroad	0.076* (0.034)	0.013 (0.036)	-0.102* (0.043)	0.027 (0.020)	-0.000 (0.020)	0.044 (0.040)	0.054 (0.035)	0.034 (0.051)
Family works for Party/State	0.008 (0.033)	0.063+ (0.034)	0.037 (0.042)	0.021 (0.021)	0.009 (0.018)	0.050 (0.039)	0.022 (0.033)	0.042 (0.048)
Flag display	-0.025 (0.036)	-0.006 (0.038)	-0.007 (0.043)	-0.029 (0.024)	-0.014 (0.020)	0.019 (0.041)	0.008 (0.038)	0.007 (0.057)
Liked US band/singer	0.015 (0.035)	-0.011 (0.037)	0.001 (0.042)	0.022 (0.023)	-0.001 (0.020)	-0.021 (0.040)	0.071+ (0.038)	-0.006 (0.054)
Relative abroad	0.004 (0.035)	-0.011 (0.036)	-0.008 (0.041)	0.021 (0.023)	-0.019 (0.019)	0.048 (0.039)	0.051 (0.036)	0.007 (0.053)
Report anti-China speech	-0.022 (0.035)	-0.043 (0.037)	-0.004 (0.041)	0.024 (0.023)	0.010 (0.020)	-0.068+ (0.039)	-0.007 (0.037)	-0.026 (0.056)
VPN use	-0.027 (0.032)	0.030 (0.035)	0.000 (0.040)	0.008 (0.022)	0.003 (0.018)	-0.015 (0.035)	0.028 (0.035)	-0.082+ (0.046)
Beliefs								
Uncond. Support	-0.097+ (0.054)	-0.021 (0.057)	-0.041 (0.065)	-0.079* (0.036)	0.018 (0.030)	-0.057 (0.062)	-0.080 (0.057)	-0.021 (0.085)
Right to Criticize	-0.017 (0.059)	-0.006 (0.063)	-0.112 (0.071)	0.008 (0.040)	-0.095** (0.033)	-0.051 (0.067)	0.114+ (0.061)	-0.012 (0.093)
Hostility	-0.013 (0.051)	-0.050 (0.053)	-0.026 (0.058)	-0.043 (0.034)	-0.023 (0.027)	0.011 (0.057)	-0.036 (0.052)	-0.082 (0.081)
General Trust	0.056 (0.050)	-0.029 (0.053)	-0.085 (0.060)	-0.009 (0.033)	0.004 (0.028)	0.011 (0.055)	0.021 (0.054)	-0.165* (0.079)
Thermometers								
US Feeling	1.344 (1.558)	-0.306 (1.634)	-0.742 (1.814)	0.311 (1.026)	-0.741 (0.871)	-0.924 (1.781)	3.733* (1.616)	0.724 (2.528)
Taiwan Feeling	0.007 (1.812)	3.029 (1.906)	-3.502 (2.132)	-1.799 (1.171)	-0.306 (1.013)	-2.529 (2.051)	1.025 (1.924)	0.486 (2.849)
EU Feeling	2.507+ (1.455)	0.917 (1.515)	-2.100 (1.688)	0.516 (0.956)	-1.375+ (0.823)	0.520 (1.664)	2.864+ (1.506)	1.430 (2.235)
Russia Feeling	0.795 (1.402)	-0.237 (1.493)	0.501 (1.620)	2.147* (0.919)	-0.050 (0.796)	1.189 (1.650)	-0.602 (1.466)	3.715+ (2.226)
Japan Feeling	1.967 (1.529)	0.371 (1.619)	-2.201 (1.798)	0.662 (1.030)	-0.176 (0.863)	-1.859 (1.733)	2.859+ (1.637)	2.762 (2.299)
Institutional Trust								
Neighborhood Comm.	-0.608 (1.340)	-2.199 (1.416)	-0.920 (1.577)	-0.155 (0.905)	-0.383 (0.750)	-1.527 (1.484)	-3.353* (1.484)	-1.951 (2.086)
Courts	-0.357 (1.295)	-1.435 (1.394)	-2.212 (1.543)	-0.654 (0.841)	0.763 (0.739)	-1.023 (1.465)	-1.885 (1.412)	-0.719 (1.889)
Local Officials	-0.700 (1.412)	-1.238 (1.471)	-0.314 (1.625)	-0.224 (0.933)	-0.285 (0.797)	-1.175 (1.584)	-1.531 (1.512)	0.039 (2.148)
People's Daily	-1.608 (1.295)	-1.935 (1.396)	-0.056 (1.516)	0.107 (0.870)	0.964 (0.742)	-0.332 (1.477)	-2.308 (1.434)	-0.080 (1.981)

Note: Each cell is the coefficient on the U.S. sponsorship-by-moderator interaction from a separate OLS regression of the outcome on treatment, the moderator, and their interaction, with HC1 robust standard errors in parentheses. + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Appendix K Respondent Engagement

Toward the end of the survey, respondents completed a four-item multiple-choice module on political knowledge: (i) the name of the United Nations Secretary-General, (ii) the year the Belt and Road Initiative was announced, (iii) the composition of the UN Security Council, and (iv) the member states of the Shanghai Cooperation Organization. We operationalize engagement in two ways: (1) the total number of correct answers, and (2) the total time spent on these four items, on the premise that longer durations indicate greater effort, for example looking up information online or consulting others. We regress the two variables on country treatment conditions, and we detect no significant differences in engagement between respondents under Chinese and U.S. sponsorship conditions.

Table A.9: Respondent Engagement

	Correct answers	Time (sec)
US Institutions	0.04	−1.85
	(0.03)	(5.01)
Num.Obs.	3000	3000

Note: HC1 robust standard errors in parentheses. Reference category: Chinese Institutions. + p <0.1, * p <0.05, ** p <0.01, *** p <0.001

Appendix L List Experiment Diagnostics

We pre-registered a double-list experiment to estimate the prevalence of VPN use. In a standard list experiment, respondents are randomly assigned to a control or a treatment group. The control group receives a list of innocuous items and reports how many apply to them; the treatment group receives the same list plus one sensitive item. The difference in mean counts identifies the prevalence of the sensitive behavior while preserving respondent privacy, but it often does so at the cost of high variance.

The double-list design adds a second list of distinct innocuous items and crosses the roles of the two groups, so that the entire sample contributes to the estimate and the standard error falls (Droitcour et al., 2004; Glynn, 2013). Respondents were randomly assigned to Group A or Group B. List 1 contained tipping livestreamers, watching short videos, using AI tools for study or work, and donating to charity; List 2 contained answering questions on an online platform (e.g., Zhihu), watching livestreams, maintaining a personal website, and using an e-reader or reading application. The sensitive item, VPN use, was added to List 1 for Group A and to List 2 for Group B. Group A is therefore the treatment group for List 1 and the control group for List 2, while Group B is the control group for List 1 and the treatment group for List 2.

We conducted standard diagnostic checks to validate the identification assumptions of the list estimator. First, we examined the data for floor and ceiling effects: the response distribution in Table A.10 shows that a relatively small share of respondents selected the minimum count of 0 or the maximum count (4 in the control condition, 5 in the treatment condition) for either list, suggesting that respondent anonymity was effectively preserved. Second, we tested for design effects to ensure that the inclusion of the sensitive item did not alter responses to the control items (Blair and Imai, 2012), with results in Table A.11. The Bonferroni-adjusted p-values for both lists were 1.00, so we fail to reject the null hypothesis of no design effect. These diagnostics confirm that the standard assumptions for the list experiment hold.

Table A.10: Response Distribution by List and Condition

	List 1 (Control)		List 1 (Treatment)		List 2 (Control)		List 2 (Treatment)	
Items Agreed (k)	N	%	N	%	N	%	N	%
0	4	0.23	2	0.11	14	0.80	10	0.57
1	49	2.80	30	1.71	164	9.36	131	7.48
2	787	44.92	560	31.96	614	35.05	527	30.08
3	728	41.55	794	45.32	890	50.80	810	46.23
4	184	10.50	312	17.81	70	4.00	251	14.33
5	0	0.00	54	3.08	0	0.00	23	1.31

Note: Response-count distributions by list and condition (number of items endorsed). Control groups endorse 0 to 4 control items; treatment groups additionally see the sensitive item and endorse 0 to 5. A count of 5 in a treatment group, or 0 in any group, would indicate a ceiling (floor) effect.

Table A.11: Design Effects Tests (Blair & Imai 2012)

	List 1		List 2	
Parameter	Est.	S.E.	Est.	S.E.
$\pi(Y_i(0) = 0, Z_i = 1)$	0.0011	0.0014	0.0023	0.0028
$\pi(Y_i(0) = 1, Z_i = 1)$	0.0120	0.0052	0.0211	0.0097
$\pi(Y_i(0) = 2, Z_i = 1)$	0.1416	0.0164	0.0708	0.0166
$\pi(Y_i(0) = 3, Z_i = 1)$	0.1039	0.0122	0.1164	0.0099
$\pi(Y_i(0) = 4, Z_i = 1)$	0.0308	0.0041	0.0131	0.0027
$\pi(Y_i(0) = 0, Z_i = 0)$	0.0011	0.0008	0.0057	0.0018
$\pi(Y_i(0) = 1, Z_i = 0)$	0.0160	0.0034	0.0725	0.0068
$\pi(Y_i(0) = 2, Z_i = 0)$	0.3076	0.0120	0.2797	0.0137
$\pi(Y_i(0) = 3, Z_i = 0)$	0.3116	0.0154	0.3916	0.0147
$\pi(Y_i(0) = 4, Z_i = 0)$	0.0742	0.0084	0.0268	0.0054

Note: Estimated population proportions for latent control-item counts (J = 4). Joint tests of no design effect give $p = 1.000$ (List 1) and $p = 1.000$ (List 2); $p > 0.05$ indicates no design effect.

However, these tests do not by themselves establish that the double-list design is valid. Double-list experiments additionally require a no-carryover assumption: the sensitive item in one list must not affect responses to the other (Diaz, 2024). This can fail if encountering the sensitive item in the first list alerts respondents to the study's purpose and changes how they answer the second. We test the

assumption with the diagnostic of Diaz (2024), regressing the item count on list order, treatment status, and their interaction, where the interaction equals the gap between the List 1 and List 2 single-list estimates. The gap is concentrated under U.S. sponsorship, where the List 1 estimate exceeds the List 2 estimate by about 13 percentage points ($p < 0.10$), while the two lists coincide under Chinese sponsorship and the pooled test is not significant (Table A.12).

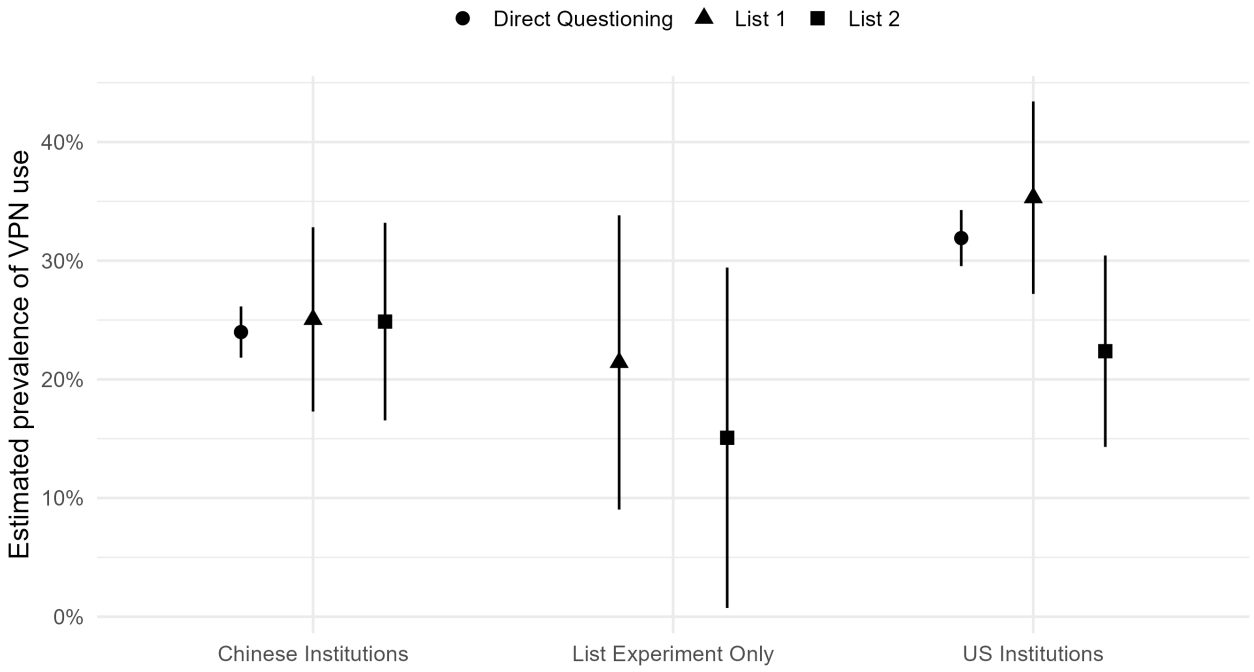
This divergence makes the pooled double-list estimator unreliable in our setting. When carryover is present, the second list no longer recovers the sensitive-item prevalence, whereas the first list remains a valid single-list estimate because its control group has not yet seen the sensitive item (Diaz, 2024). We therefore restrict the main analysis to the List 1 estimate, treating the design as a classical single-list experiment, and report both estimates alongside direct questioning in Figure A.11.

Table A.12: DiD Analysis on Double List

	Pooled	List Only	Chinese Sponsor	U.S. Sponsor
List 2 (vs List 1)	−0.115***	−0.115+	−0.135***	−0.095*
	(0.025)	(0.063)	(0.039)	(0.038)
Sensitive Item	0.289***	0.214***	0.251***	0.353***
	(0.026)	(0.063)	(0.040)	(0.041)
List 2 x Sensitive Item	−0.066	−0.063	−0.002	−0.129+
	(0.044)	(0.112)	(0.067)	(0.068)
Num.Obs.	7008	1008	3010	2990
R2	0.034	0.023	0.031	0.043
Std.Errors	by: respondent_id	by: respondent_id	by: respondent_id	by: respondent_id

Note: Cluster-robust standard errors (clustered by respondent) in parentheses; each respondent contributes one row per list. + p <0.1, * p <0.05, ** p <0.01, *** p <0.001

Figure A.11: Comparison between Disaggregated List Experiment Results and Direct Questioning



Appendix M List Experiment Results by Treatment Group

Table A.13 presents regression results testing whether list experiment estimates of VPN usage differ by sponsorship condition. In both models, the U.S. group serves as the baseline reference. The models interact the treatment indicator (assignment to the list with the sensitive item) with the sponsorship group. A negative value for the interaction term indicates that respondents are less likely to report VPN usage under the specified sponsorship condition compared to the U.S. baseline. In the comparison between U.S. and Chinese institutions, the interaction term for the China group is negative and statistically significant at the 10 percent level. This demonstrates that the estimated VPN usage drops by 10.3 percentage points when the survey is presented as sponsored by Chinese institutions rather than U.S. institutions, and by 13.9 percentage points when presented without a sponsor in the list-only condition. Overall, these results indicate that the sponsorship gap persists under the list experiment, though the list estimate is less precise ($p \approx 0.07$).

Table A.13: List Experiment Estimates by Sponsor Condition (List 1)

	US vs List Only	US vs China
Intercept	2.632***	2.632***
	(0.027)	(0.027)
Treated (US Sponsor Baseline)	0.353***	0.353***
	(0.041)	(0.041)
List Only Group	−0.120*	
	(0.050)	
China Group		−0.051
		(0.038)
Treated x List Only Group	−0.139+	
	(0.076)	
Treated x China Group		−0.103+
		(0.057)
Num.Obs.	1999	3000

Note: HC1 robust standard errors in parentheses. Reference group: US Institutions. + p <0.1, * p <0.05, ** p <0.01, *** p <0.001

Appendix N List Experiment 2: Foreign Scholars' Right to Criticize

Our pre-analysis plan included a second double list experiment, measuring agreement that foreign scholars have a right to criticize China. The two lists paired the sensitive item, "Foreign scholars have the right to criticize problems in China," with four non-sensitive control statements each, rotating the sensitive item across lists. List 1 comprised: (i) there is too little trust between people in our society; (ii) our society has already achieved gender equality; (iii) gender inequality still exists in our society; and (iv) replacing one's phone every year is a reasonable consumption habit. List 2 comprised: (i) garbage sorting is not that necessary in cities; (ii) China's environmental pollution has improved a lot compared with ten years ago; (iii) China's environmental pollution has not improved much compared with ten years ago; and (iv) it is reasonable for primary and middle school students to have eight hours of class each day. Each list contains a pair of logically opposed statements (gender equality achieved vs. still absent; pollution much improved vs. not), a standard device for negatively correlating control items and compressing the count distribution (Glynn, 2013); a consistent respondent therefore agrees with at most three of the four, so the observed control counts top out at three.

We administered it only in the no-sponsor baseline, using a crossover double-list (Lists 1 and 2). Unlike the VPN experiment, the two lists agree closely (0.508 and 0.456) and the Diaz (2024) carryover diagnostic finds no between-list difference (List 2 minus List 1 = -0.052, $p = 0.57$, clustered by respondent), so we pool them; the pooled double-list estimate places agreement at about 48 percent in the baseline (Table A.14).

This experiment cannot serve as a parallel test of whether anonymity removes the sponsorship effect, for two reasons. First, its list arm was fielded only in the no-sponsor baseline, so it yields no list-based estimate of the sponsor gap and cannot show whether the effect persists under anonymity. Second, the item-count technique requires a binary sensitive item, so this attitudinal item, measured on a four-point agree to disagree scale in direct questioning, had to be dichotomized for the list, leaving the list prevalence and the direct measure on non-comparable scales (Figure A.12). VPN use,

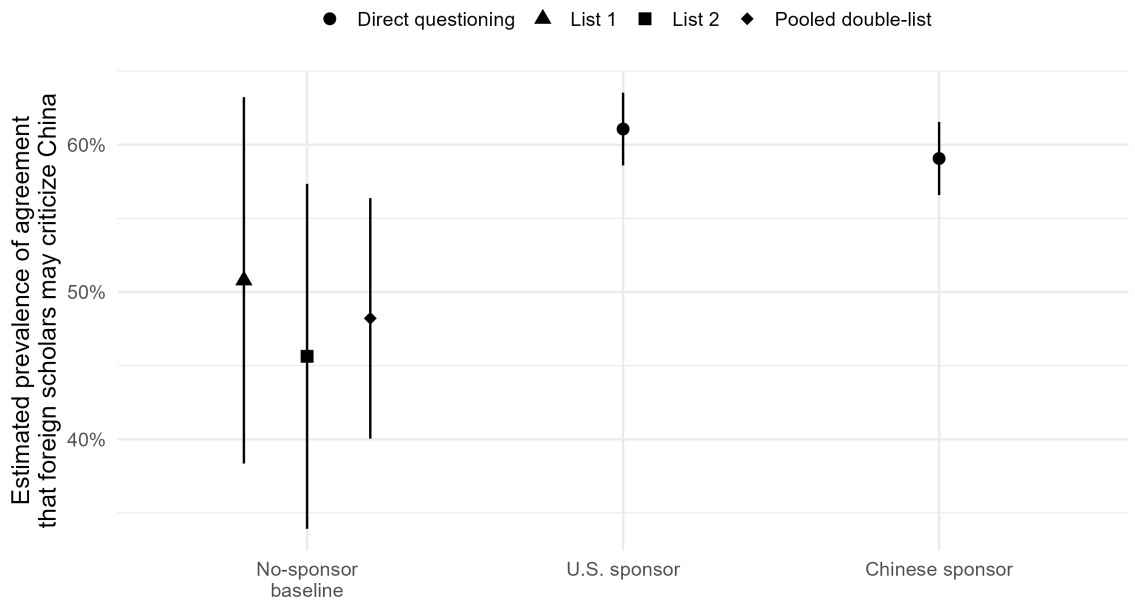
analyzed in the main text, avoids both problems: it is intrinsically binary and was administered under every sponsor condition.

Table A.14: List Experiment 2 (Foreign Scholars' Right to Criticize), No-Sponsor Baseline

Estimator	Prevalence (SE)
List 1 (single list)	0.508 (0.063)
List 2 (single list)	0.456 (0.060)
Pooled double-list	0.482 (0.042)

Note: Estimated agreement that foreign scholars have a right to criticize China, from the baseline-only double-list experiment (n = 504). The Diaz (2024) carryover diagnostic gives a between-list difference of -0.052 (p = 0.57, clustered by respondent), so the two lists are pooled. Cluster-robust standard errors in parentheses.

Figure A.12: List Experiment 2 versus Direct Questioning: Foreign Scholars' Right to Criticize



Appendix O Deviations from the Pre-Analysis Plan

We preregistered the study prior to data collection.[14] Table A.15 records where the reported analysis departs from that plan and the rationales for these deviations.

Table A.15: Deviations from the Pre-Analysis Plan

Pre-analysis plan	**What we did**	**Rationale**
Estimate VPN prevalence and the U.S. versus Chinese sponsor gap from the double-list experiment, pooling both lists.	Report the List 1 single-list estimate as primary; report the List 2 and pooled estimates in Appendix L.	The Diaz (2024) carryover diagnostic shows the two lists diverge under U.S. sponsorship. List 1 remains valid as a single-list experiment under carryover, whereas the pooled estimator does not.
A second list experiment (foreign scholars' right to criticize China) as a further test of the sponsorship effect.	Reported descriptively in Appendix N; not used to estimate the sponsor gap.	Its list arm was fielded only in the no-sponsor baseline, so it yields no sponsor-gap estimate; the four-point attitudinal item required dichotomization for the list, leaving it on a scale non-comparable to the direct measure.
Several outcomes were preregistered as measured but not tied to a directional hypothesis: Liking U.S. Singer/Band, Family Works for Party/State, Most People Can Be Trusted, and Feeling–Russia.	We analyze these as exploratory, alongside the preregistered confirmatory outcomes (H1–H4).	Reported for completeness; not part of the confirmatory tests. Their exploratory status is flagged in Table 1 in the main text.

[14] https://osf.io/32tm6/?view_only=0b0de815718c44798d31619a9d6f9b5b

Appendix P IRB Approval

Figure A.13: IRB Approval

Printed on: Monday, June 9, 2025

Dear [redacted]

On Monday, June 9, 2025 [redacted] Institutional Review Board (IRB) reviewed and approved your Initial application for the project titled "**Does the nationality of researchers affect survey responses? Evidence from China**". Based on the information you provided, the IRB determined that:

- the risks to subjects are minimized through (i) the utilization of procedures consistent with sound research design and do not unnecessarily expose participants to risk, and (ii) whenever appropriate, the research utilizes procedures already being performed on the subjects for diagnostic or treatment purposes
- the risks to participants are reasonable in relation to anticipated benefits, if any, to participants, and the importance of the knowledge that may reasonably be expected to result
- the selection of subjects is equitable
- informed consent be sought from each prospective subject or the subject's legally authorized representative, in accordance with, and to the extent required by §46.116
- informed consent be appropriately documented, in accordance with, and to the extent required by §46.117
- when appropriate, the research plan makes adequate provisions for monitoring the data collected to ensure the safety of subjects
- when appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of data
- when some or all of the subjects are likely to be vulnerable to coercion or undue influence, such as children, prisoners, pregnant women, mentally disabled persons, or economically or educationally disadvantaged persons, additional safeguards have been included in the study to protect the rights and welfare of these subjects

In addition, the IRB determined that documented consent is not required for all participants.
The IRB approved a waiver of documentation of informed consent.

This review procedure, administered by the IRB, in no way absolves you, the researcher, from the obligation to adhere to all Federal, State, and local laws and the [redacted] policies. Immediately inform the IRB if you would like to change aspects of your approved project (please consult our website for specific instructions). You, the researcher, are respectfully reminded that the University's ability to support its researchers in litigation is dependent upon conformity with continuing approval for their work.

[redacted]

[redacted]

Best wishes for your research,