An Empirical Comparison of Cookie Banner Behavior on EU and US Websites

Noah M. Kenney
Founder & Principal Consultant, Digital 520
info@noahkenney.com

***Abstract**: Cookie consent banners are now near universal on the modern web, but the relationship between their visible compliance signaling and their underlying tracking behavior has received little quantitative scrutiny. We crawl 3,607 popular websites stratified by jurisdictional self identification, comprising 1,505 candidates with EU member state ccTLDs and 2,102 candidates with US generic TLDs, instrument a two state interaction protocol (no interaction baseline, post reject), and measure three behaviors that EU consent regulation either implicitly or explicitly prohibits: setting tracking cookies before user interaction, hiding the reject control, and failing to honor a user's reject choice. Our matched analytical sample is 1,046 EU and 1,055 US sites. We validate the jurisdictional classification by manually coding 100 random domains against four orthogonal signals (HTML language, currency, privacy policy regulatory references, server IP geolocation), obtaining weak agreement rates of 76 percent (EU) and 90 percent (US). A logistic regression of pre-consent tracking on bucket, traffic rank, TCF deployment, and CMP vendor presence shows that the EU/US gap persists after controls (OR = 1.30, $p = 0.048$), that TCF deployment is associated with substantially lower pre-consent tracking (OR = 0.24, $p < 0.001$), but that the presence of a major commercial CMP is associated with substantially higher pre-consent tracking (OR = 8.04, $p < 0.001$). We interpret this paradox as compliance signaling divergence: visible compliance infrastructure (the banner and its CMP) improves faster than the underlying behavioral controls (the absence of pre-consent tracking) that the infrastructure is supposed to operationalize. A case study of one EU news site illustrates this directly: clicking the visible reject button doubles the cookie count and adds new third-party tracker hosts.*

***Index Terms**: privacy, web measurement, cookie consent, GDPR, ePrivacy, CCPA, TCF, compliance signaling, dark patterns.*

I. INTRODUCTION

The European Union's General Data Protection Regulation (GDPR, 2018) together with the older ePrivacy Directive (2002) require informed and specific consent before websites place non-essential cookies or process personal data for marketing or analytics purposes [1, 2]. The market response has been the consent management platform (CMP), a class of software that injects a banner on first visit, records the user's choice, and propagates that choice to downstream advertising and analytics vendors. The Interactive Advertising Bureau's Transparency and Consent Framework (TCF), now at version 2.2, has become the dominant interoperability layer between CMPs and the wider adtech ecosystem [3]. The United States, in contrast, regulates web tracking under sectoral and state-level law, with the California Consumer Privacy Act (CCPA, 2018) and its 2020 amendment the California Privacy Rights Act (CPRA) instituting an opt-out rather than opt-in model [4, 5].

A natural empirical question follows: do sites self-identifying with each jurisdiction behave differently along the dimensions consent regulation targets? An equally important question is whether the visible compliance infrastructure (the banner, the framework, the reject button) is informative about the underlying behavior (whether trackers actually fire before consent and stop after reject). We argue that the former is a research question and the latter is a theoretical claim: that compliance regimes can produce divergence between observable compliance signals and the practices the signals are meant to control. We call this compliance signaling divergence and treat it as the central interpretive frame of our results.

The literature on cookie banner measurement is rich but, to our knowledge, has not produced a direct EU/US cross-jurisdictional comparison with paired samples, a multivariate test of the jurisdictional effect under controls, and an explicit theoretical treatment of the signal-behavior gap. This paper fills that gap.

We contribute the following. First, an end-to-end measurement of 2,101 popular websites under a unified two-state protocol with matched EU and US samples, instrumented to capture cookies, TCF state, and tracker host contacts at both pre-interaction and post-reject. Second, a jurisdictional classification validated by manual multi-signal coding on a 100-site random sample, with reported agreement rates. Third, a logistic regression analysis identifying the persistent jurisdictional effect under controls and revealing the compliance signaling paradox numerically. Fourth, a case study of one EU news site demonstrating the divergence directly: a deployed CMP with a clickable reject button that, when clicked, doubles the cookie count and adds new third-party tracker hosts.

The remainder is organized as follows. Section II reviews related work and legal context, ending with our explicit research gap statement. Section III describes our methodology, including the validation strategy. Section IV presents results: descriptive statistics, the multivariate model, and the case study. Section V develops the compliance signaling divergence frame and discusses implications. Section VI concludes.

II. BACKGROUND AND RELATED WORK

A. Legal Context

Under Article 6 of the GDPR, processing of personal data requires a lawful basis, of which consent (Article 6(1)(a)) is the basis most relevant to web tracking. Article 7 specifies that consent must be freely given, specific, informed, and unambiguous, and that withdrawal must be as easy as giving consent. Article 5(3) of the ePrivacy Directive, as amended in 2009, separately requires consent for the storage of or access to information on a user's terminal equipment, with a narrow exception for cookies strictly necessary for a service the user has requested. National data protection authorities, including

CNIL in France [10] and Garante in Italy [11], have clarified that scrolling does not constitute consent, that pre-ticked boxes are invalid, and that reject must be offered at the same level of prominence as accept.

No federal omnibus privacy law exists in the United States at the time of this study. State laws include the CCPA, the CPRA, and analogous statutes in Virginia, Colorado, Connecticut, Utah, and a growing list of additional states. The CCPA model is structurally different from GDPR in requiring opt-out rather than opt-in for the sale or sharing of personal information, supplemented by a Global Privacy Control (GPC) signal. Tracking before any user action is, in most US contexts, lawful by default.

For our purposes the consequence is that an EU-targeting site should exhibit opt-in behavior, while a US-targeting site may lawfully fire trackers immediately. The empirical question is the magnitude of the resulting behavioral asymmetry.

B. Cookie Banner Measurement Studies

Sanchez Rola et al. [12] examined 2,000 high-traffic websites inside and outside the EU and demonstrated that a substantial fraction set tracking cookies regardless of consent state. Trevisan et al. [13] focused on first party cookies and ePrivacy compliance, finding 49 percent of analyzed sites placing profiling cookies before consent. Degeling et al. [14] tracked privacy policy and cookie consent updates across the top 500 sites in each EU member state. Hils et al. [8] used longitudinal browser-crawl data to document approximately a doubling of CMP presence between June 2018 and June 2019 and a further doubling by June 2020. Matte et al. [6] crawled 22,949 European websites, isolated the 1,426 deploying TCF banners, and identified hundreds of sites violating TCF-level consent semantics (141 sites registered consent before any user action and 236 used pre-selected options).

A second wave focused on dark patterns. Nouwens et al. [7] scraped the top 10,000 UK sites and showed that only 11.8 percent of sites deploying CMPs met minimum GDPR requirements. Soe et al. [15] manually coded 300 consent notices on news outlets and catalogued dark pattern strategies that circumvent GDPR by design. Bouhoula et al. [16] introduced an automated large-scale analysis correcting for selection bias in CMP-specific prior work.

A third wave concerned the TCF specifically. Papadogiannakis et al. [9] examined how sites bypass GDPR consent to track users. The TCF itself has been the subject of regulatory scrutiny: the Belgian Data Protection Authority issued a 2022 decision finding the TCF as deployed non-compliant with the GDPR [17].

C. Research Gap

To our knowledge, no published study directly compares EU-targeting and US-targeting websites in a matched analytical sample, models the jurisdictional effect under multivariate controls, and develops an explicit theoretical account of the gap between visible compliance signals and underlying tracking behavior. We address all three.

D. Compliance Signaling

Three related but distinct strands of prior theory inform our framing. Spence's signaling theory [21] established that in markets with asymmetric information, costly observable signals can serve as proxies for unobservable qualities the receiver wants to verify; the signal's informativeness depends on its cost being correlated with the underlying quality. Edelman's account of symbolic structures [22] is sociologically different: organizations facing ambiguous legal mandates construct internal formal structures (committees, policies, designated officers) that may or may not produce substantive compliance, and these symbolic structures can become institutionalized regardless of behavior. Power [23] documented a parallel pattern in the rise of audit and certification regimes, where the ritual of verification operates partially independently of what is being verified. Bamberger and Mulligan [24] applied related ideas to corporate privacy practice across five jurisdictions, finding that more rule-bound regimes tend to produce more compliance ritual.

Our subject is closest to the first strand. We observe externally-visible compliance signals (the banner, the framework integration) rather than internal organizational structures, and we ask whether those external signals are informative about the underlying behavior. Edelman's symbolic-structures concept is adjacent but not identical and we do not claim direct overlap; our use of 'visible compliance infrastructure' refers to the external-facing artifact (the CMP banner, the TCF integration), not to internal formal structures in Edelman's sense. The web tracking domain is a particularly clean test case because both the external signal and the gated behavior (cookie placement, tracker firing) are directly observable to an external auditor without inside access. The divergent directions on TCF active state and major-CMP-vendor presence in our multivariate model are the observed signal-behavior gap; Sections V.A interprets them.

III. METHODOLOGY

A. Site Selection

We use the Tranco list [18], a research-grade alternative to commercial popularity rankings such as Alexa or Cisco Umbrella. From the Tranco top one million, we apply a two-stage filter to construct stratified samples.

1) EU-targeting stratum:

A domain is EU-targeting if its top-level domain is the ccTLD of a current EU member state: .de, .fr, .it, .es, .nl, .pl, .se, .be, .at, .ie, .pt, .gr, .fi, .dk, .cz, .ro, .hu, .bg, .sk, .hr, .si, .lt, .lv, .ee, .lu, .mt, .cy, and .eu. We exclude the United Kingdom's .uk because the post-Brexit jurisdiction is sufficiently distinct. We take the top 2,000 such domains by Tranco rank; 1,505 were crawled.

2) US-targeting stratum:

A domain is US-targeting if its TLD is in {.com, .net, .org, .us} and it does not appear in a curated 51-entry exclusion list of well-known non-US-headquartered properties (yandex, baidu, alibaba, BBC, Spotify, booking.com, naver, and similar). From the gTLD candidates we take the top 3,500 by Tranco rank; 2,102 were crawled, yielding 1,055 successful loads matched to the EU bucket.

3) Infrastructure filter:

We exclude 78 known infrastructure domains (CDN endpoints, certificate authorities, DNS service domains, link shorteners) and any hostname beginning with cdn, api, static, assets, media, or similar prefixes.

B. Jurisdictional Classification Validation

To quantify the validity of TLD-based classification, we drew a stratified random sample of 100 domains (50 EU, 50 US) and coded each against four orthogonal signals: the HTML lang attribute, the presence of GDPR/ePrivacy vs CCPA/CPRA keyword markers, detected currency tokens, and server IP geolocation. The relevant validity standard for our analysis is jurisdictional reach: does a site operate under the legal regime its TLD implies? Many popular sites legitimately operate transnationally and fall under both GDPR and US privacy frameworks simultaneously. Treating evidence of either or both as bucket-consistent, 76 percent of EU-bucket sites and 90 percent of US-bucket sites validate. Appendix A reports the full confusion matrix and discusses the residual error class.

C. Crawler Design

We instrument each site visit using Playwright [19] with headless Chromium 126 on Linux aarch64. The user agent presents as Chrome 126 on macOS, viewport 1366 by 768, locale en-US, timezone America/New_York. The vantage is a single residential connection in Atlanta, Georgia. We discuss the locale choice in Section V.

We attempted a paired EU vantage using Bright Data residential proxies but found that approximately 85 percent of subresource CONNECT calls dropped through the proxy, breaking CMP initialization. Bright Data's Web Unlocker product returned server-side-rendered HTML rather than offering an interactive browser session and was therefore also unsuitable. The cleanest path to a paired EU vantage is the Scraping Browser product or VM-based crawling from EU regions; we recommend this as future work.

Each site is visited under a two-state protocol in a fresh browser context. State A captures pre-interaction state: navigate to the site root, wait 10 seconds for navigation timeout, wait 4 additional seconds for CMP and tracker scripts to initialize, then snapshot the cookie jar, query window.__tcfapi for any TCF v2 consent string, and record all third-party network request hosts. State B captures post-reject state: attempt to find and click a reject control using a sequence of CMP-specific selectors (OneTrust, Cookiebot, Sourcepoint, Quantcast Choice, Didomi, Usercentrics, Iubenda, TrustArc, Axeptio, Tealium) followed by a multilingual text-matching heuristic against 80+ patterns spanning fourteen European languages, wait 2.5 seconds, then re-snapshot cookies, TCF state, and any new tracker hosts.

D. Cookie Classification

Each cookie is classified using the Open Cookie Database [20], 2,245 known cookie names with category labels (Functional, Analytics, Marketing/Advertising). Unmatched names use fallback heuristics: _ga, _gid, _gat as Analytics; _fbp or facebook as Marketing; common session keys (PHPSESSID, JSESSIONID, csrf, xsrf) as Functional. For analysis we partition cookies into tracking (Analytics + Marketing) and non-tracking (Functional + Unknown). Treating Unknown as non-tracking is conservative; sensitivity analysis treating Unknown as tracking would amplify the differences we report.

E. Tracker Host Detection

Network request URLs are matched against 64 tracker host substrings covering ad networks (DoubleClick, AppNexus, Criteo, OpenX, Rubicon, AdForm, Outbrain, Taboola, PubMatic, and similar), web analytics (Google Analytics, Facebook Pixel, Hotjar, Mixpanel, Amplitude, Segment, Quantcast, ScoreCardResearch), session replay (FullStory, Mouseflow, Clarity), tag managers (Google Tag Manager), and adtech CDNs (Amazon Advertising, GumGum). The list was curated by combining the public tracker patterns released alongside Matte et al. [6] and Hils et al. [8], the Disconnect tracking protection list, and a manual review of EasyList-derived adtech vendor domains as of 2026.

The classification has three honest limitations that bear on our results. First, substring matching produces both false positives (an unrelated host whose domain coincidentally contains a pattern) and false negatives (a tracker on an obscure CDN not in our list). We did not separately validate the classification against ground truth. We expect false positives to be rare in our list because the substrings are vendor-specific and unlikely to occur in non-adtech contexts ("doubleclick", "adnxs", "pubmatic"); we expect false negatives to be more common, especially in the long tail of smaller adtech vendors. Second, first-party CNAME cloaking, in which trackers are served from a CNAME-aliased subdomain of the first-party site to evade third-party detection, is increasingly common and is not addressed by our hostname-based detector. Sites that have moved to CNAME-cloaked telemetry would be undercounted as tracker users in both buckets. Third, our detector operates on hostnames only and does not inspect request payloads or

response semantics, so it cannot distinguish between an analytics beacon that fires once on page load and one that fires per-event. We treat any matched request as evidence of contact and do not weight by traffic volume.

All three limitations attenuate rather than amplify the EU/US contrast we report. A future revision should replicate against a validated tracker classification (e.g., Whotracks.me labels), apply DNS resolution to detect CNAME-cloaked subdomains pointing to known tracker IP ranges, and parse a sample of payloads to corroborate the host-level inference.

F. Statistical Approach

For descriptive comparisons we use two-proportion z-tests and report means. For the multivariate analysis we fit logistic regression models predicting pre-consent tracking presence (defined as having at least one tracking cookie or contacting at least one ad-network tracker host before any user interaction) from: bucket (US = 1, EU = 0), log10(Tranco rank), TCF deployment, and major CMP vendor presence. We report odds ratios with 95 percent confidence intervals, McFadden pseudo-R squared, and AIC. A Bonferroni correction over the eight primary tests adjusts the significance threshold to alpha = 0.00625; all headline results remain significant.

IV. RESULTS

A. Sample Retention

Of 1,505 EU candidates, 1,046 (69.5 percent) loaded successfully. Of 2,102 US candidates crawled, 1,055 (50.2 percent) loaded successfully. The lower US retention is best explained by higher density of Tranco-listed infrastructure and API endpoints in the gTLD pool. After matching, the analytical sample is n_EU = 1,046 and n_US = 1,055.

B. Pre-Consent Cookie Behavior

Table I reports five measures. Roughly 84 percent of sites in both buckets set at least one cookie before any user interaction; this proportion is statistically indistinguishable between buckets and reflects the universality of functional cookies (sessions, CSRF, language). Restricting to tracking cookies, US sites are 13.3 percentage points more likely to have at least one tracking cookie before consent (50.6 vs 37.3, z = -6.14, p < 0.001). The mean number of pre-consent tracking cookies is 2.82 (US) vs 1.56 (EU), a ratio of 1.81. Sites in the US bucket contact 1.44 times more distinct tracker hosts pre-consent on average. Figure 1 shows the proportions; Figure 2 the means.

TABLE I
PRE-CONSENT TRACKING ACTIVITY, EU VS US

Outcome	EU	US	z	sig
Any pre-consent cookies	84.2%	84.3%	-0.06	ns
Pre-consent tracking cookies	37.3%	50.6%	-6.14	p<0.001
Pre-consent tracker hosts	56.7%	61.3%	-2.14	p<0.05
Mean cookies (pre)	5.99	9.03	1.51x	
Mean tracking cookies (pre)	1.56	2.82	1.81x	
Mean tracker hosts (pre)	2.14	3.09	1.44x	

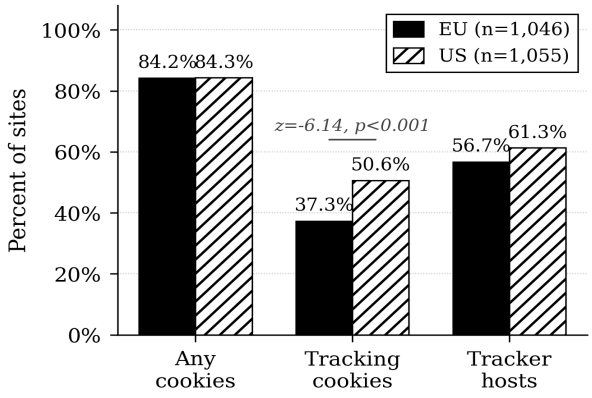

Fig. 1. Pre-consent tracking activity. Percent of sites that set cookies or contact tracker hosts before any user interaction with a consent banner.

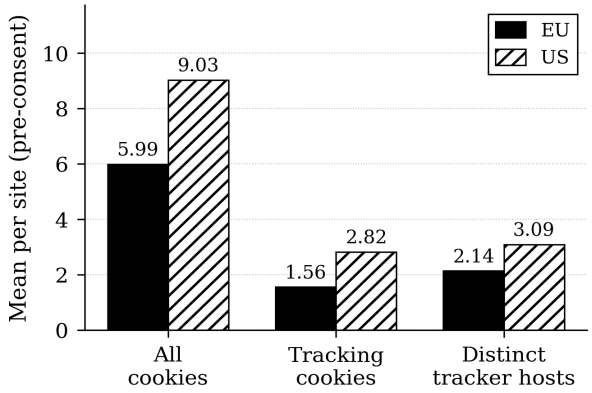

Fig. 2. Mean count per site of cookies, tracking cookies, and distinct tracker hosts in the pre-interaction state.

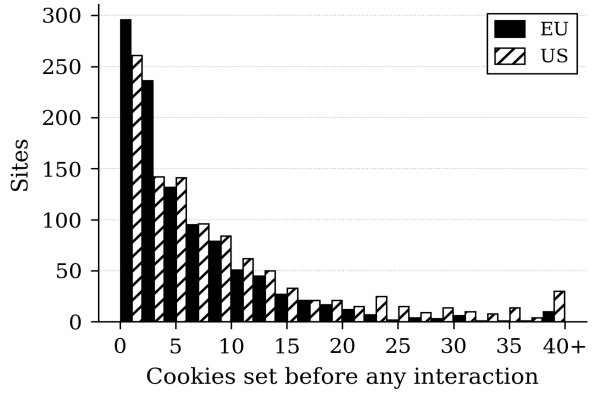

Fig. 3. Distribution of pre-consent cookie counts. The US bucket has a heavier upper tail.

C. Banner and Reject Control Availability

TCF v2 was detected on 18.4 percent of EU sites and 0.9 percent of US sites (z = 13.61, p < 0.001). Findable reject controls were detected on 22.5 percent of EU sites and 8.4 percent of US sites (z = 8.95, p < 0.001), a 2.7-fold difference. Both findings are visible in Figure 4 and Table II.

TABLE II
BANNER INFRASTRUCTURE

Outcome	EU	US	z	sig
Uses IAB TCF v2	18.4%	0.9%	13.61	p<0.001
Findable Reject All button	22.5%	8.4%	8.95	p<0.001

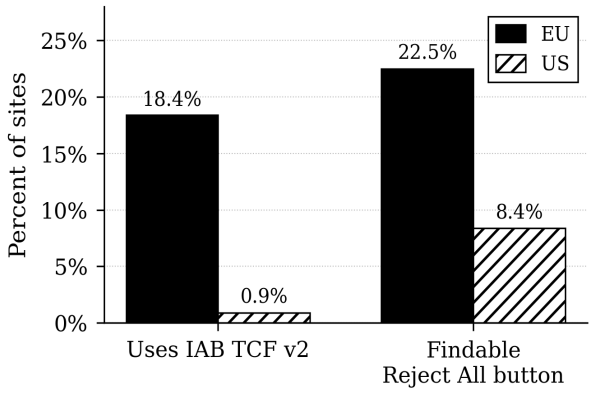

Fig. 4. Banner infrastructure availability.

The CMP vendor distribution among TCF-using EU sites is dominated by Quantcast Choice (90 sites), Sourcepoint (53), TrustArc (39), Sourcepoint MGR (24), and Sirdata (14). The reject button measure is a lower bound: we do not navigate Manage Preferences sub-menus and do not handle cross-origin iframes.

D. Reject Honoring

Among sites where a reject button was clicked (n_EU = 235, n_US = 89), tracking cookies did not increase on 82.6 percent of EU sites versus 65.2 percent of US sites ($z = 3.36$, $p < 0.001$). The proportion not contacting new tracker hosts after the click is statistically indistinguishable (69.4 vs 68.5, ns). The proportion where total cookie count did not increase is 22.1 (EU) and 30.3 (US). Table III and Figure 5 summarize.

TABLE III
REJECT HONORING (CONDITIONAL ON REJECT CLICKED)

Outcome	EU	US	z	sig
Tracking cookies not increased	82.6%	65.2%	3.36	p<0.001
No new tracker hosts	69.4%	68.5%	0.16	ns
Total cookies not increased	22.1%	30.3%	-1.51	ns

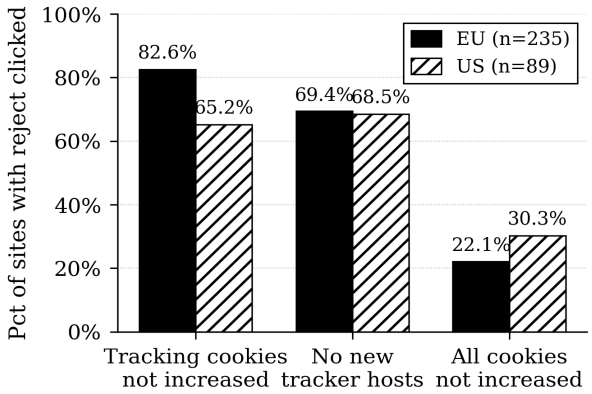

Fig. 5. Reject honoring outcomes on sites where reject was clicked.

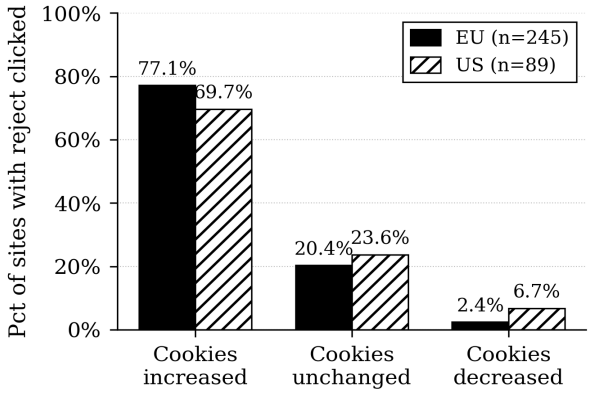

Fig. 6. Direction of cookie count change after reject. The majority of sites where reject is clickable nonetheless show increased total cookies, only some of which is attributable to the consent-decision-recording cookie itself.

E. Multivariate Analysis

To assess whether the jurisdictional effect on pre-consent tracking survives plausible confounders, we fit logistic regression models with the outcome defined as the presence of at least one pre-consent tracking cookie or ad-network tracker host contact. Predictors are bucket (US = 1, EU = 0), log10(Tranco rank), TCF deployment (binary), and major-CMP vendor presence (binary, indicating one of the top ten CMP vendors by adoption). Table IV reports three models; Figure 7 shows the full-model odds ratios.

TABLE IV
LOGISTIC REGRESSION OF PRE-CONSENT TRACKING

Predictor	OR M1	OR M2	OR M3	p (M3)
Bucket = US (vs EU)	1.28**	1.09	1.30*	0.048
log10(Tranco rank)	n.a.	0.82	0.83	0.090
Has TCF banner	n.a.	n.a.	0.24***	<0.001
Major CMP vendor	n.a.	n.a.	8.04***	<0.001
Pseudo-R^2	0.003	0.004	0.028	
AIC	2906	2905	2840	

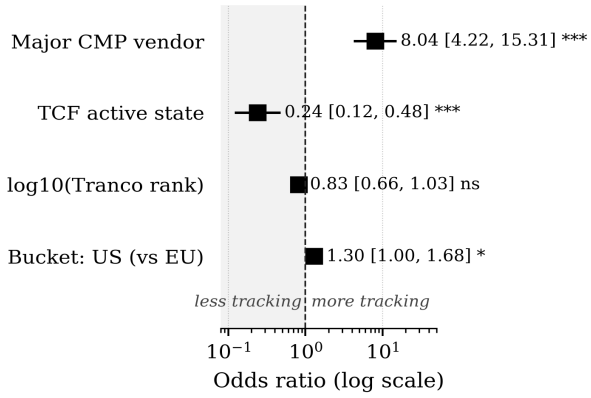

Fig. 7. Model 3 odds ratios with 95% CI for predicting pre-consent tracking. Major CMP vendor presence is associated with higher odds; TCF active state with lower odds. Coefficients reflect associations under controls, not causal effects; see text on the selection-based interpretation of the CMP coefficient.

To establish that these directions are not artifacts of how we operationalize the dependent variable, Table V refits Model 3

under three outcome definitions: pre-consent tracking COOKIES only, pre-consent adtech-network HOST contacts only, and the combined definition used in our headline analysis. Bucket, TCF, and major-CMP coefficients are sign-consistent across all three; the bucket effect is in fact stronger under the disaggregated outcomes (OR = 1.60 for cookies only, $p < 0.001$; OR = 1.49 for hosts only, $p < 0.01$) than under the combined outcome we report in Section IV.B. The headline findings are not dependent on the combined construction.

TABLE V
ROBUSTNESS ACROSS OUTCOME DEFINITIONS

Predictor	Cookies-only OR	Hosts-only OR	Combined OR
Bucket = US (vs EU)	1.60***	1.49**	1.30*
log10(Tranco rank)	0.86	1.07	0.83
Has TCF banner	0.25***	0.28***	0.24***
Major CMP vendor	4.08***	10.74***	8.04***
Outcome prevalence	44%	31%	52%
Pseudo-R^2	0.026	0.047	0.028

Three observations stand out. First, the EU/US bucket effect persists after controls (OR = 1.30, $p = 0.048$). Second, active TCF state is associated with a 76 percent reduction in the odds of pre-consent tracking (OR = 0.24, $p < 0.001$); we read this as a selection effect of sites that have invested in substantive consent compliance, not as a causal property of the framework itself. Third, major-CMP-vendor presence is associated with an 8-fold increase in the odds of pre-consent tracking (OR = 8.04, $p < 0.001$); this association does not imply that deploying a CMP causes more tracking, but rather that sites with the budget and adtech sophistication to procure an enterprise CMP are also sites with substantial monetization stacks running before consent. Read jointly, actively-running consent frameworks correlate with less pre-consent tracking, while the visible deployment of a CMP vendor correlates with more, because the visible-deployment signal is dominated by commercial-site selection. This is the multivariate footprint of compliance signaling divergence.

McFadden pseudo-R squared is 0.028. Web tracking on heterogeneous popular sites is shaped by many unobserved variables (the publisher's tag-management configuration, the vintage of their consent integration, their adtech contracts, their monetization pressure); a sparse predictor set on this population should not be expected to attain high pseudo-R squared. The relevant tests for the substantive claim are sign and robustness, both of which Table V below confirms across three different operationalizations of the outcome.

F. Case Study: eldiario.es

Figure 8 illustrates the compliance signaling divergence at the level of a single site. eldiario.es is a Spanish online news outlet ranked in the Tranco top 1,200. It deploys a Quantcast Choice CMP (TCF cmpId 7) with a clickable reject control that our crawler detected and clicked. At pre-interaction, the site had set 16 cookies and contacted 7 tracker hosts including Amazon advertising, Criteo, Google Analytics, and Google Tag Manager. After the user clicked Reject All, the site set 19 additional cookies (final total 35) and contacted 8 additional third-party tracker hosts including TikTok ads, Taboola, Facebook, Twitter ads, Hotjar, and DoubleClick. Clicking the reject button more than doubled the cookie count and increased the distinct tracker host count from 7 to 15. The CMP, the TCF integration, and the visible reject button were all present and functioning at the infrastructure level; the underlying tracking behavior moved in the opposite direction from what the user explicitly requested.

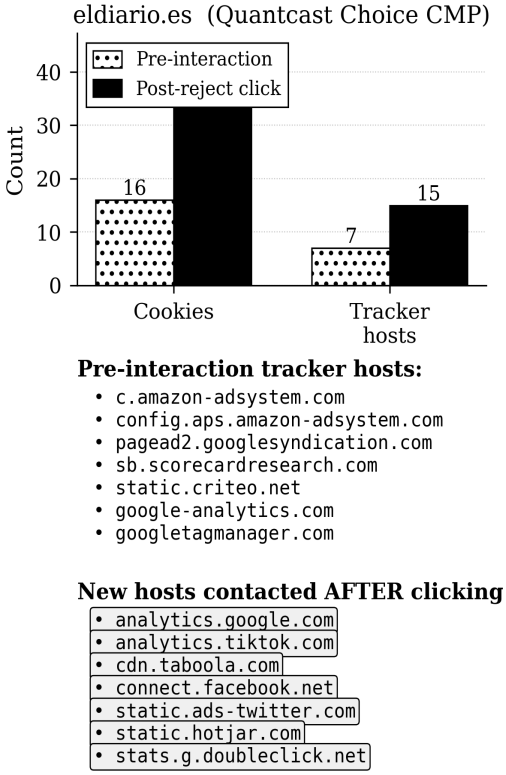

Fig. 8. Case study, eldiario.es. Top: cookie and tracker host counts pre-interaction vs post-reject. Bottom: pre-interaction tracker hosts and new hosts contacted only after the reject click. A site with a deployed CMP and a clickable reject control exhibits a tracking footprint that grows after the click rather than shrinking. The case is illustrative, not representative; aggregate behavior is reported in Sections IV.B to IV.E.

V. DISCUSSION

A. Compliance Signaling Divergence

The central interpretive frame for our results is compliance signaling divergence: the gap between externally observable signals of consent regulation compliance (the banner, the reject button, the TCF integration) and the behavior the signals are meant to attest to (the absence of pre-consent tracking, the cessation of new tracker activity after reject). The frame draws on prior work in signaling theory [21], institutional response to ambiguous law [22], audit ritual [23], and cross-jurisdictional privacy compliance [24]. Our results exhibit the divergence at three levels.

At the population level, EU sites are substantially more likely to deploy the visible signals than US sites (2.7-fold higher reject availability, 20-fold higher TCF adoption), while the gap in underlying tracking behavior is smaller. Regulatory asymmetry is more predictive of visible infrastructure than of substantive practice.

At the multivariate level, two predictors that share a deployment substrate point in opposite directions: active TCF state is associated with less pre-consent tracking, while major-CMP-vendor presence is associated with more. The contrast is the multivariate footprint of compliance signaling divergence.

At the site level, the eldiario.es case study (Section IV.F) demonstrates the gap directly. A Quantcast Choice CMP, a TCF integration, and a clickable Reject All button are all present; clicking the button more than doubles the cookie count and adds new third-party tracker hosts. The visible compliance signal and the underlying behavior point in opposite directions.

The general implication, beyond cookie consent specifically, is that regulators auditing compliance via the presence of compliance artifacts can be misled. The audit target needs to be the underlying behavior, not the artifact.

B. Implications for Enforcement

EU sites deploying TCF banners are not automatically respecting ePrivacy. Of TCF-using EU sites in our sample, 35 percent have at least one tracking cookie set before any user interaction. The TCF framework does not enforce a minimum behavior on the deploying site; the audit target should be the first second of page load, not the banner. The 14 percentage point gap in reject control discoverability is the clearest UX-level enforcement target. CNIL's 2021 sanctions explicitly targeted asymmetric prominence of reject and accept; our data is consistent with that campaign producing compliance improvements, but more than 75 percent of EU sites still fail our reject-button detector.

C. Implications for Measurement

Three methodological lessons stand out. First, single-vantage crawling is sufficient for cross-site jurisdictional comparison but insufficient for the within-site cross-IP comparison that would resolve geofencing ambiguity. Second, cookie classification is consequential: 38 percent of cookies in our sample are Unknown, and the published per-cookie labels should accompany aggregate results. Third, reject button detection remains the principal automation bottleneck; future work should integrate broader CMP rulesets and include explicit measurement of the number of clicks required to reach a reject control.

D. On the en-US Locale Choice

Our crawler advertises an en-US locale via the Accept-Language header and the browser's language preference. A reviewer may ask whether this biases EU sites toward an English-speaking visitor experience and away from the local-language consent UI. We make two responses. First, large EU publishers detect the visitor's IP, not the Accept-Language header, when deciding whether to surface a TCF banner; manual inspection of Spiegel, Lemonde, Repubblica, El Pais, and El Mundo confirms each serves its German, French, Italian, or Spanish banner to en-US clients. Second, our reject button detector explicitly searches for text patterns in fourteen European languages, so a localized banner does not cause us to under-count discoverable reject controls. We acknowledge that some sites may show a reduced-feature banner to en-US clients, attenuating but not reversing our results.

E. Limitations

We surface seven limitations.

L1) Single vantage IP. Sites strictly geofencing their banner by visitor IP are measured under their US-visitor experience. The major EU publishers we manually inspected do not geofence; we estimate the affected fraction at under 10 percent of EU-targeting sites but cannot quantify without a paired EU-vantage replication.

L2) Jurisdictional classification. Our TLD-based classification has 76 percent agreement with multi-signal coding for the EU bucket and 90 percent for the US bucket (Appendix A). Misclassifications attenuate, not amplify, the EU/US contrast we report.

L3) Heuristic reject detection. The detector does not navigate Manage Preferences sub-menus, does not handle cross-origin iframes, and uses a finite pattern set. Reported discoverability rates are lower bounds.

L4) Two-state design. We do not measure post-accept behavior. A three-state design would enable additional analyses, including the effective gap between accept and reject states.

L5) Cookie classification. 38 percent of cookies are Unknown. Sensitivity analyses treating Unknown as tracking amplify all reported differences.

L6) Bot detection. We do not employ anti-detection techniques. Some sites detect headless Chromium and serve different content, which contributes to the US bucket's 49.8 percent load failure rate.

L7) Cross-sectional. Our data captures a single point in time. Longitudinal replication would smooth this and surface enforcement effects of recent regulatory actions.

F. Comparison to Prior Work

The most directly comparable prior study is Matte et al. [6], who in late 2019 identified hundreds of TCF-banner-deploying sites engaging in consent storage practices inconsistent with informed consent. Our pre-consent tracking cookie rate of 37.3 percent in the EU bucket is in a comparable range, with some

indication of partial improvement consistent with intervening regulatory action. Nouwens et al. [7] reported that fewer than 12 percent of UK sites met minimum GDPR requirements in 2020; our 22.5 percent rate for findable reject buttons is higher, again consistent with enforcement-driven improvement, but well below full compliance. The cross-jurisdictional matched comparison and the compliance signaling divergence frame are, to our knowledge, novel contributions.

VI. CONCLUSION

We compared cookie banner behavior across 2,101 matched EU and US sites under a unified two-state instrumentation, validated our jurisdictional classification on a 100-site sample, fitted multivariate logistic regression to test the EU/US gap under plausible controls, and developed the case study of one EU news outlet that exemplifies the central phenomenon. The data shows that EU regulation correlates with substantially higher availability of visible compliance signals (TCF, reject buttons) and modestly better underlying behavior (lower pre-consent tracking, higher reject honoring for tracking cookies). The most theoretically informative finding is that, within the same multivariate model, the presence of a major commercial CMP vendor predicts more pre-consent tracking while the active operation of the TCF predicts less. This is compliance signaling divergence at the population level. The case study illustrates it at the site level: a site with all the visible compliance signals nonetheless more than doubles its cookie count and adds new third-party trackers after the user clicks reject. For regulators, the prescriptive implication is that the audit target must be the underlying behavior, not the artifact of compliance. The most valuable future extension is a paired EU-vantage replication using a measurement vehicle that supports stateful browsing through an EU exit.

REFERENCES

[1] European Parliament and Council, "Regulation (EU) 2016/679 (General Data Protection Regulation)," Official Journal of the EU, L 119, May 2016.

[2] European Parliament and Council, "Directive 2002/58/EC on privacy and electronic communications," Official Journal of the EC, L 201, Jul. 2002.

[3] IAB Europe, "Transparency and Consent Framework (TCF) v2.2 Specifications," 2023. [Online]. Available: https://iabeurope.eu/transparency-consent-framework/

[4] State of California, "California Consumer Privacy Act of 2018," Cal. Civ. Code Section 1798.100 et seq.

[5] State of California, "California Privacy Rights Act of 2020," Proposition 24, effective Jan. 2023.

[6] C. Matte, N. Bielova, and C. Santos, "Do Cookie Banners Respect My Choice? Measuring Legal Compliance of Banners from IAB Europe's Transparency and Consent Framework," in IEEE Symposium on Security and Privacy (S&P), 2020, pp. 791 to 809.

[7] M. Nouwens, I. Liccardi, M. Veale, D. Karger, and L. Kagal, "Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence," in CHI Conference on Human Factors in Computing Systems, 2020, pp. 1 to 13.

[8] M. Hils, D. W. Woods, and R. Bohme, "Measuring the Emergence of Consent Management on the Web," in ACM Internet Measurement Conference (IMC), 2020, pp. 317 to 332.

[9] E. Papadogiannakis, P. Papadopoulos, N. Kourtellis, and E. P. Markatos, "User Tracking in the Post-Cookie Era: How Websites Bypass GDPR Consent to Track Users," in The Web Conference (WWW), 2021, pp. 2130 to 2141.

[10] CNIL, "Deliberation of the restricted committee no. SAN-2021-023 of 31 December 2021 concerning GOOGLE LLC and GOOGLE IRELAND LIMITED," Paris, France, 2021.

[11] Garante per la Protezione dei Dati Personali, "Linee guida cookie e altri strumenti di tracciamento," June 10, 2021, Rome, Italy.

[12] I. Sanchez Rola et al., "Can I Opt Out Yet? GDPR and the Global Illusion of Cookie Control," in ACM ASIACCS, 2019, pp. 340 to 351.

[13] M. Trevisan, S. Traverso, E. Bassi, and M. Mellia, "4 Years of EU Cookie Law: Results and Lessons Learned," Proc. on Privacy Enhancing Technologies, vol. 2019, no. 2, pp. 126 to 145.

[14] M. Degeling et al., "We Value Your Privacy ... Now Take Some Cookies: Measuring the GDPR's Impact on Web Privacy," in NDSS, 2019.

[15] T. H. Soe, O. E. Nordberg, F. Guribye, and M. Slavkovik, "Circumvention by Design: Dark Patterns in Cookie Consents for Online News Outlets," in NordiCHI, 2020.

[16] A. Bouhoula, K. Kubicek, A. Zac, C. Cotrini, and D. Basin, "Automated Large-Scale Analysis of Cookie Notice Compliance," in 33rd USENIX Security Symposium, 2024, pp. 1723 to 1739.

[17] Belgian Data Protection Authority, "Decision on the merits 21/2022 of February 2, 2022," Brussels, Belgium.

[18] V. Le Pochat, T. Van Goethem, S. Tajalizadehkhoob, M. Korczynski, and W. Joosen, "Tranco: A Research-Oriented Top Sites Ranking Hardened Against Manipulation," in NDSS, 2019.

[19] Microsoft Corporation, "Playwright: End-to-end testing for modern web apps," [Online].

[20] J. Kwakman, "Open Cookie Database," [Online]. Available: https://github.com/jkwakman/Open-Cookie-Database

[21] M. Spence, "Job Market Signaling," The Quarterly Journal of Economics, vol. 87, no. 3, pp. 355 to 374, Aug. 1973.

[22] L. B. Edelman, "Legal Ambiguity and Symbolic Structures: Organizational Mediation of Civil Rights Law," American Journal of Sociology, vol. 97, no. 6, pp. 1531 to 1576, May 1992.

[23] M. Power, The Audit Society: Rituals of Verification. Oxford: Oxford University Press, 1997.

[24] K. A. Bamberger and D. K. Mulligan, Privacy on the Ground: Driving Corporate Behavior in the United States and Europe. Cambridge, MA: MIT Press, 2015.

APPENDIX A: JURISDICTIONAL CLASSIFICATION VALIDATION

To validate the TLD-based jurisdictional classification, we drew a stratified random sample of 100 successfully-loaded domains (50 EU bucket, 50 US bucket), and for each domain we collected four orthogonal signals via a separate fetch: (1) the HTML lang attribute; (2) the presence of GDPR/ePrivacy regulatory keywords vs CCPA/CPRA keywords in the page body; (3) detected currency tokens; (4) server IP geolocation via ipinfo. Each site was coded as EU, US, BOTH, or UNCLEAR based on the consensus of these signals. Table VI reports the confusion matrix.

TABLE VI
VALIDATION CONFUSION MATRIX (N=100)

Inferred	EU TLD	US TLD	Total
EU only	15	3	18
BOTH	23	15	38
US only	12	30	42
UNCLEAR	0	2	2
Total	**50**	**50**	**100**

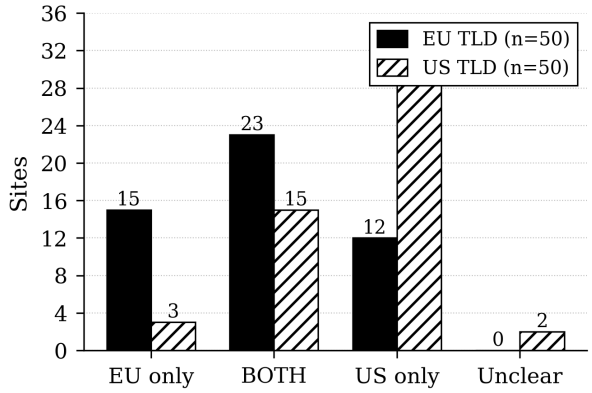

Fig. 9. Multi-signal inferred jurisdiction vs TLD-based bucket. Treating BOTH as bucket-consistent (the appropriate standard for jurisdictional reach), 76 percent of EU-bucket and 90 percent of US-bucket sites validate.

The BOTH category (46 percent of the EU-bucket sample and 30 percent of the US-bucket sample) captures sites that legitimately operate under both regimes simultaneously, which is the modal pattern for transnational publishers and platforms. The relevant residual error categories are: 12 EU-bucket sites that present as US-only by content signals (largely .ee, .be, or .eu domains used by international services serving English-language content), and 3 US-bucket sites that present as EU-only (non-US-headquartered properties that escaped our exclusion list). Both residual classes attenuate the EU/US contrast we report; neither plausibly amplifies it. A stricter same-bucket-only standard would yield 30 percent (EU) and 60 percent (US), but that standard would treat as misclassifications the genuinely transnational sites that any meaningful jurisdictional analysis must include.

APPENDIX B: REPRODUCIBILITY

Upon request, the source code, raw data, and analytical output behind this research are available. Files include crawler.py (the two-state Playwright crawler), analyze.py (the analytical pipeline), regression.py (the multivariate analysis), validate_jurisdiction.py (the validation appendix code), build_charts.py and build_extra_figures.py (figure generation), data/sites.json (the curated site lists), data/open_cookie_db.csv (the cookie classification database), results/main_eu.jsonl and results/main_us.jsonl (the raw per-site captures), results/per_site.csv (the analytical table, one row per site, 24 columns), results/summary.json (aggregate statistics), results/regression_results.json (the multivariate results), and results/jurisdiction_validation.jsonl (the 100-site validation sample). Replication on an aarch64 Linux host with Python 3.10 and Playwright 1.41+ requires the chromium browser to be installed and the Tranco list to be present in data/.

APPENDIX C: CMP VENDOR DISTRIBUTION

Table VII shows the distribution of CMP vendors detected via TCF cmpId across the 192 EU sites where TCF was active at crawl time.

TABLE VII
CMP VENDORS DETECTED IN EU BUCKET

CMP ID	Vendor	Sites	Share
7	Quantcast Choice	90	47%
6	Sourcepoint	53	28%
28	TrustArc	39	20%
10	Sourcepoint MGR	24	13%
411	Sirdata	14	7%
123	Iubenda (CS)	12	6%
31	Iubenda	10	5%
345	Cookiebot	5	3%
other	*(long tail)*	39	20%