POLICY ARTICLE

Mapping the online manipulation economy

A market perspective on digital manipulation may help improve online trust and safety

A. Dek[1*], Y. Kyrychenko[2], S. van der Linden[2], J. Roozenbeek[2*†]

The internet is awash with inauthentic activity. Although some of this activity is harmless, a substantial proportion can be classified as malicious, including bots designed to steal personal data, networks of inauthentic social media accounts promoting crypto scams, and coordinated political influence operations (*1*, *2*). To enhance security, online platforms have implemented measures to counter inauthentic activity, with most requiring users to confirm account authenticity via a one-time password (OTP) sent through short message service (SMS). However, there exists a thriving market for buying and selling on-demand SMS verifications, which forms a cornerstone of the wider online manipulation economy. Despite its centrality in the digital manipulation ecosystem, this SMS verifications market has not been systematically studied. To address this knowledge gap, we developed the Cambridge Online Trust and Safety Index (COTSI; cotsi.org), which tracks the daily price of SMS verifications for 197 countries and over 500 platforms.

Malicious inauthentic activity is common. For example, one study found that "social bots play a disproportionate role in spreading [news] articles from low-credibility sources" (*3*). Another investigation found that, in 2024, "bad bots" made up 24% of all internet traffic (*4*). However, the problem extends beyond bots: researchers distinguish between automation (to what extent online activity is governed by algorithms or code), inauthenticity (whether an account is truthful in its representation), and coordination (accounts working together to achieve certain goals) (*5*). Studies have shed light on the role of automated, inauthentic, and/or coordinated activity in the proliferation of disinformation campaigns, cryptocurrency scams on social media, and distributed denial of service attacks (*2*, *6*, *7*). With the rise of generative AI, this type of content has become easy to produce and extremely difficult to detect (*8*). Yet, rather than cracking down on inauthentic activity, many online platforms have cut measures to curb it, and have started paying users for creating engaging content, thus incentivizing a reliance on inauthentic engagement.

THE ONLINE MANIPULATION MARKETPLACE

To address these challenges, researchers have sought to understand how to detect digital manipulation efforts, for instance by developing bot detection algorithms (*8*) (see supplementary materials S1.1). However, with some exceptions (*9*), few studies have explicitly looked at inauthentic online activity through the lens of market economics. We argue that doing so is key for the study of financial crime and political influence operations, and for the design and implementation of interventions. The online manipulation marketplace is thriving: inauthentic likes, comments, views, and followers are all readily available for purchase. Offers range from low-end fake likes and views used for vanity metrics inflation, to mid-end manipulations used by search engine optimization specialists to promote businesses, and even high-end influence campaigns (*6*, *9*). As a starting point, manipulation service providers purchase physical or virtual SIM (subscriber identity module) cards in bulk. These SIM cards are linked to phone numbers, which are then used to verify accounts via SMS for the purpose of mass account registration.

We conceptualize this online manipulation economy as a transnational gray market (*10*) (see supplementary materials S1.2). Providers of manipulation services operate worldwide, but outside of authorized distribution channels. For example, using an SMS verification service to register a social media account goes against most online platforms' terms of service. However, whether doing so is illegal is not clear. Providers may be held liable in some jurisdictions, depending on how their services are used (e.g., for fraud), though legal precedent is lacking (see supplementary materials S1.2.1). There are also several benign use cases for such services, including by open-source investigators and people in countries with internet restrictions.

Buyers of online manipulation services take on risks, both formal (e.g., detection by platforms) and informal (e.g., account registrations being unsuccessful, their data being leaked or hacked). Because manipulation service providers often operate as informal businesses based in hard-to-reach jurisdictions, buyers have limited legal recourse to offset losses. Products are differentiated (not homogeneous), as there is substantial variation in quality, for example with respect to the success rate of account registrations.

It is often possible to register an account using an SMS verification from a country where it could be bought more cheaply, and later deploy or sell this account in a more expensive country—a phenomenon we refer to as arbitrage. However, most platforms have security measures in place to disincentivize such efforts. For example, a Facebook account registered in Vietnam pretending to be a U.S. voter may be flagged by Facebook's fingerprinting system—a process for gathering user information, e.g., through their browser or device settings—and banned. Buyers are therefore incentivized to register accounts using local SIM cards and phone numbers (see supplementary materials S1.2). However, such arbitrage is not always feasible. For messenger apps (e.g., WhatsApp), accounts' phone numbers are readily visible to all users, thus revealing their country of origin. Facebook public page administrators are required to disclose their country of registration, and LinkedIn's (voluntary) account verification system also shows this information.

A few studies are available that shed light on how the digital manipulation market operates (*11*, *12*). For example, Shi *et al.* (*9*) investigated the market for fake likes and comments in the European Union. Fredheim *et al.* (*6*) found that, in 2023, a mere €10 could buy tens of thousands of views, thousands of likes, and hundreds of followers on several high-profile online platforms. The authors point out that very little has been done since 2021 to counteract these flaws and protect platforms' integrity (see supplementary materials (S1.1)).

THE SMS VERIFICATIONS MARKET

Few studies have investigated the underlying SMS verifications infrastructure that serves as a cornerstone of the online manipulation economy. SMS verifications are an essential component for the registration of accounts on online platforms: without them, downstream online manipulation services (account creation, programming bots to post or engage with content, et cetera) become near-impossible. Studying the SMS verifications market is thus key to understanding how inauthentic content spreads, and to finding novel ways of countering digital manipulation.

By our count, there are at least 17 providers offering on-demand SMS verifications (see supplementary materials S2.1). These providers operate worldwide, but cater mainly to Russian- and, to a lesser extent, Chinese-speaking customers, as evidenced by their websites being

available in grammatically correct Russian (but incorrect English), and offering payment through Russian payment platforms. Visitors can select both the platform for which they wish to verify an account and the country of the phone number to which the SMS will be sent. The site then lists the price and the number of verifications available for that country. Providers bulk-purchase SIM cards to deliver their services. Because a SIM card can be used to register for a given platform only once, prices and stocks of SMS verifications vary considerably. The range of platforms on offer is broad: not only social media platforms, but also messenger apps (e.g., WhatsApp, Telegram), financial services (e.g., Revolut), crypto exchanges (e.g., Coinbase, Binance), and marketplaces (e.g., Amazon) (see supplementary materials S1.3).

We developed the COTSI (cotsi.org) to enable accurate assessments of price fluctuations over time, and monitor the implementation and evaluation of counter-measures. This free online platform tracks the daily price and availability of on-demand SMS verifications across 197 countries and over 500 online platforms. To create the index, we first identified 17 providers of SMS verification services and sorted them by the web traffic they receive (see Tables S1 and S20). These providers operate out in the open, with publicly accessible websites (some even keep blogs), and application programming interfaces (APIs) that allow visitors to extract data (e.g., on prices/stocks). From this list, we selected four providers to be included in the price index (SMSActivate, 5Sim, SMShub, and SMSPVA), based on (i) their verification prices and stocks being freely available via API, preferably without login or paywall; and (ii) the website being unique and not a clone of another provider (some websites look identical and have the same prices/stocks listed) (see supplementary materials, Table S20). We then developed scripts that automatically collect data through the providers' APIs. COTSI data consists of the price (averaged across providers) of SMS verifications, and the number of verifications available, for each combination of country and platform (see supplementary materials S2.1).

Predictors of SMS verifications prices

To understand what drives the cost of on-demand SMS verifications, we collected one year of COTSI data (25 July 2024 until 27 July 2025) and calculated the average price and stock of SMS verifications in each country per provider (see supplementary materials S3.1). We then linked this data with the Digital Society Project, a country-level dataset which consists of 37 variables divided into five categories: coordinated influence operations, digital media freedom, states' internet regulatory capacity and approach, online media polarization, and social cleavages (see supplementary materials S2.3 for details and Table S18 for an overview of item wordings and scale points used). Due to conceptual and statistical complications of aggregating the variables according to these categories (see supplementary materials S2.3), we recoded the variables into the following categories: the presence of (i) domestic and (ii) foreign social media disinformation, (iii) internet censorship, (iv) internet governance, (v) online social media use, (vi) online media diversity, (vii) the use of social media for political action, and (viii) political polarization (all Cronbach's $\alpha > 0.70$). We also collected country-level statistics: minimum SIM card price, mobile data cost, country population, the number of phones per capita, and GDP per capita (see supplementary materials S3.1 for further details).

We noticed substantial variation in the success rate of SMS verifications, depending on the provider. For instance, 5Sim lists its success rate per country on its website. In late July 2025, it reported its success rate for registering a Facebook account in the US as being at most 21.43%. This may be due to the fact that some providers rely primarily on virtual SIM cards, which are of lower quality than physical ones. This is likely because platforms can identify the carrier associated with a phone number, and their security algorithms may treat numbers from virtual providers with increased suspicion. After experimenting with registering our own verifications, we concluded that SMSPVA has the highest success rate (over 90%). In our analyses, we therefore look separately at the predictors of SMS verification prices for SMSPVA, in order to understand the differences between high-quality and lower quality services. Finally, many of the 500 platforms in the COTSI are of minimal use in influence campaigns. We therefore also look separately at the predictors of SMS verification prices for social media and messaging platforms: Google (incl. YouTube/Gmail), Facebook, Instagram, Twitter/X, WhatsApp, TikTok, LinkedIn, VKontakte, Discord, Telegram, Line, WeChat, Snapchat, Weibo, and QQ. Because of non-linearity in our data, we ran mixed-effects gamma regressions with log-transformed SMS verification price as the dependent variable, our aggregated Digital Society categories and country-level measures as independent variables, and providers and countries as random effects (see supplementary materials S3.2).

We find that account verifications via SMS are consistently most expensive in Japan ($4.93), Australia ($3.24), Turkey ($2.54), and Malta ($2.18). These prices are far higher than other countries including the US ($0.26), the UK ($0.10), and Russia ($0.08) (see the figure). This may be related to the price of (physical) SIM cards, which cost a minimum of $29.77 in Japan, $6.23 in Australia, and $10.37 in Malta. High SIM card prices may be partially a function of regulation: Japan requires proof of residency before buying a SIM card, and Australia requires photo identification (*13*). However, SIM card price does not tell the full story, as these cost ~$0.95 in Turkey. That said, registering a SIM card in Turkey is complex, as phone numbers are linked to passports.

We further find that the most important factors associated with the price of SMS verifications are the stock of available verifications in a given country for a given platform ($b = 0.92$ [0.92, 0.93], $p < .001$; lower stocks tend to increase the price), GDP per capita ($b = 1.20$ [1.08, 1.33], $p = .001$; higher GDP predicts higher price), and the number of phones per capita ($b = 0.89$ [0.81, 0.97], $p = .007$; fewer phones are linked to higher price). All other associations are non-significant at $p > .067$ (see Table S11). When zooming in on the provider with the highest verification success rate (SMSPVA), we find that (in addition to GDP and phones per capita) a higher minimum price for a SIM card is also associated with an increase in the price of SMS verifications ($b = 1.22$ [1.06, 1.40], $p = .004$). Curiously, the effect for stocks is also significant but in the opposite direction ($b = 1.05$ [1.02, 1.07], $p = .012$; higher stocks instead increase price). All other p-values > .055 (see Table S14). We find that, with minor variations, all these associations are robust when only looking at social media and messaging platforms (see Tables S13 and S15). Finally, we find no significant associations between our socio-political variables and SMS verification price. However, when looking at SMSPVA verifications only, we find marginally significant effects for (more) domestically produced disinformation on social media ($b = 1.26$ [0.99, 1.60], $p = .063$), and the (lower) use of social media for political action ($b = 0.79$ [0.63, 1.00], $p = .055$) (see Table S14; supplementary materials S3.3).

Influence operations during election campaigns

To understand how the online manipulation market might react to political events, we examined whether the price and availability of SMS verifications for eight social media platforms (Google/YouTube/Gmail, Facebook, Instagram, Twitter/X, WhatsApp, TikTok, LinkedIn, and Telegram) increases ahead of national elections. If confirmed, this may indicate increased demand for online manipulation services during contentious political moments, and a sign of influence operations. We

conducted a cumulative average abnormal returns analysis to see if SMS verifications prices and stocks yield abnormal returns in the 30 days leading up to 61 national elections held between July 2024 and June 2025, in countries with >1 million inhabitants. We find that for Telegram and WhatsApp (though with less statistical confidence for the latter), SMS verifications prices increase significantly ahead of national elections, indicating increased demand for such registrations during these periods. However, we do not find this to be the case for the other six platforms (see supplementary materials S3.4).

DISCUSSION

Substantial progress can be made in our understanding of online influence operations and financial crime by investigating the thriving market that buttresses online manipulation. We developed the COTSI to shed light on this underground market. The COTSI has many applications. For instance, it can be used to track changes in SMS verification prices and stocks ahead of elections, or to assess if interventions impact the price or availability of manipulation services.

Drawing on one year of COTSI data, we examined what predicts the price of SMS verifications. Based on these analyses, we offer several considerations for policymakers. First, the availability of cheap SIM cards in some countries may make it easier to mass-register accounts. Policymakers may therefore investigate malicious actors' reliance on local infrastructure, e.g., SIM farms. In April 2025, the United Kingdom was the first country in Europe to pass legislation making SIM farms illegal (*14*). As this law comes into effect, the COTSI will allow us to see if SMS verification prices in the UK indeed increase. As the price and availability of SIM cards are associated with high-quality SMS verification prices, policymakers may consider making the purchase and registration of a SIM card more complex (without unduly sacrificing users' privacy). Shi *et al*. (*9*) also suggest making manipulation service providers harder to find on search engines, thus reducing their visibility.

Regarding the finding that SMS verifications prices increase significantly ahead of elections for Telegram and WhatsApp, but not other social media platforms, we offer several explanations (see supplementary materials S3.4). Most importantly, arbitrage is more limited for messaging apps. It is easy to use a fake social media account registered in one country to post about events in another, since the account's country of registration is usually not visible to other users. Messaging apps, however, make it easy to see where an account is from, thus incentivizing the registration of local (inauthentic) accounts. Platforms may therefore consider making accounts' country of registration more transparent for all users.

Our study is not without limitations. We focus on the SMS verifications market, which we argue is a cornerstone of the wider online manipulation economy. However, while we posit a strong theoretical link between SMS verification prices and the cost of other services such as fake likes and comments (*9*), we are currently unable to trace these costs through the COTSI. Furthermore, there is a risk that the publication of this paper alerts SMS verification providers, who might then make access to their APIs more difficult, complicating future study. We also recognize that inauthentic behavior forms only one part of a wider set of challenges around digital information consumption and production. Nonetheless, we believe that our findings serve as a starting point for boosting trust and safety on online platforms.

REFERENCES AND NOTES

1. Z. Lin, J. Cui, X. Liao, X. Wang, Malla: Demystifying Real-world Large Language Model Integrated Malicious Services. *ArXiv Preprints*, doi: 10.48550/arXiv.2401.03315 (2024).
2. V. Chergarova, V. Arcanjo, M. Tomeo, J. Bezerra, L. M. Vera, A. Uloa, Cryptocurrency fraud: A study on the characteristics of criminals who are using fake profiles on a social media platform to persuade individuals to invest into cryptocurrency. *Issues in Information Systems* **23**, 242–252 (2022).
3. C. Shao, G. L. Ciampaglia, O. Varol, K.-C. Yang, A. Flammini, F. Menczer, The spread of low-credibility content by social bots. *Nat Commun* **9**, 4787 (2018).
4. T. Richabadas, "Threat Spotlight: Bad bots are evolving to become more 'human' " (2024); https://blog.barracuda.com/2024/11/19/threat-spotlight-bad-bots-evolving-more-human.
5. L. Mannocci, M. Mazza, A. Monreale, M. Tesconi, S. Cresci, Detection and Characterization of Coordinated Online Behavior: A Survey. *ArXiv Preprints* (2024).
6. R. Fredheim, S. Bay, T. Haiduchyk, A. Dek, M. Stolze, "Social Media Manipulation 2022/2023: Assessing the Ability of Social Media Companies to Combat Platform Manipulation" (2023); https://stratcomcoe.org/publications/social-media-manipulation-20222023-assessing-the-ability-of-social-media-companies-to-combat-platform-manipulation/272.
7. J. Pastor-Galindo, P. Nespoli, J. A. Ruipérez-Valiente, D. Camacho, Influence Operations in Social Networks. *ArXiv Preprints*, doi: 10.48550/arXiv.2502.11827 (2025).
8. K. Yang, F. Menczer, Anatomy of an AI-powered malicious social botnet. *Journal of Quantitative Description: Digital Media* **4** (2024).
9. L. Shi, G. Lenzini, B. Farrand, X. Shu, M. Hu, M. Kalameyets, A. Sergeeva, E. Jalilzade, H. Hammouchi, J. Wang, "FAMOUS: Fake Activity Market Observation System of Unethical Services" (Newcastle, 2025); https://drive.google.com/file/d/1ijSRgnvQmvAQN-UiILS1f-EGaUrUzJ0O/view.
10. L. P. Bucklin, Modeling the international gray market for public policy decisions. *International Journal of Research in Marketing* **10**, 387–405 (1993).
11. M. Mazza, G. Cola, M. Tesconi, Ready-to-(ab)use: From fake account trafficking to coordinated inauthentic behavior on Twitter. *Online Soc Netw Media* **31**, 100224 (2022).
12. T. Elmas, R. Overdorf, K. Aberer, Characterizing Retweet Bots: The Case of Black Market Accounts. *Proceedings of the International AAAI Conference on Web and Social Media* **16**, 171–182 (2022).
13. P. Bischoff, Which governments impose SIM-card registration laws to collect data on their citizens?, *Comparitech* (2025). https://www.comparitech.com/blog/vpn-privacy/sim-card-registration-laws/.
14. UK Home Office, Major step for fraud prevention with landmark ban on SIM farms , *gov.uk* (2025). https://www.gov.uk/government/news/major-step-for-fraud-prevention-with-landmark-ban-on-sim-farms.
15. A. Dek, Y. Kyrychenko, S. van der Linden, J. Roozenbeek, Data and code for "Mapping the online manipulation economy", *OSF* (2025). https://osf.io/t3xzg/.

ACKNOWLEDGMENTS

We thank Iryna Dek (Centre for Information Resilience) and Kyrylo Manakhov (Trementum Analytics) for their help with the COTSI. We are grateful to the UK Cabinet Office for funding. **Funding:** Cabinet Office of the United Kingdom (IRIS Academic, # SCH-00001-3391). The funding source had no involvement in the study design, data collection, analysis, writing, or decision to submit. All information necessary to replicate our findings (data, code, and supplementary analyses) can be found at (*15*).

Supplemental material

References

10.1126/science.adw8154

[1] Judge Business School, University of Cambridge; Cambridge UK. [2] Department of Psychology, University of Cambridge; Cambridge UK. * Equal contribution. † Email: jjr51@cam.ac.uk.