



Youth Internet Bans, Circumvention, and Substitution: Policy Briefing

Executive summary

- **The best existing evidence suggests that youth internet bans are unreliable tools for ensuring that under-16s do not access harmful online content.** The evidence that we have from a variety of countries and settings suggests that attempts to exclude children from online activities are frequently met with either circumvention or substitution into other online activities.
- The best direct evidence on circumvention comes from South Korean and Chinese internet and/or gaming bans, evidence from the United States, and preliminary evidence from the Australian social media ban. **Research consistently shows that circumvention of bans is common, with the most common methods of circumvention including VPN usage, borrowing unbanned accounts, adversarial methods to bypass age-assurance checks, and retaining use of accounts owned prior to the implementation of the ban.**
- The preliminary evidence from Australia is especially relevant given that it is the only example of a youth social media ban globally. While early evidence should not be taken to be definitive, **the research that does exist indicates that circumvention is relatively straightforward, and that many young people have retained social media accounts from prior to the implementation of the ban, although there is also some evidence suggesting a modest reduction of usage of banned platforms.**
- **There is also a concern that exclusion from online spaces may result in young people substituting into other online activities that do not comply with the ban.** Evidence suggests that this is highly likely to be the case; studies on online pornography bans show that searches for non-compliant rival sites increase dramatically after the implementation of pornography bans.
- **The central policy take-away from this rapid research synthesis is that youth internet bans should not be taken to be reliable policy instruments for preventing children from accessing harmful content.** This is true for two reasons: firstly, because circumvention is sufficiently common that many children simply will not be kept off the platforms that are trying to exclude them. Secondly, because many children will substitute their use of compliant platforms with similar platforms that do not comply with the ban.

Review question and scope

The policy briefing addresses three questions:

1. What happens when children and adolescents are banned or restricted from particular online services, content categories, or times of use?
2. Do young people circumvent these restrictions, and through what routes?
3. Among young people who do not circumvent restrictions, is there evidence of substitution onto other platforms?

The term “ban” is used broadly. Included evidence covers social-media age restrictions, gaming curfews and playtime mandates, pornography age-verification laws, platform age gates, home or ISP filters, parental controls, school phone bans, and related age-assurance systems. The reasoning here is that essentially no completed peer-reviewed studies evaluate broad social-media bans for under-16s, so other evidence is needed to give the broadest overview of the likely effects and efficacy of youth social media bans.

This report is a rapid synthesis of available evidence on youth internet bans, age assurance, substitution, and circumvention. Sources were identified through iterative database and web searching, backward and forward source chasing from relevant policy briefings, review of regulator and parliamentary materials, and targeted searches for Australian implementation evidence. Although we use a mixture of peer-reviewed evidence and preliminary evidence from non peer-reviewed sources (including working papers), we prioritised evidence that had strong causal identification and/or was peer-reviewed; we used non peer-reviewed evidence where stronger evidence was unavailable. In the case of the Australian under-16 ban, no peer-reviewed evidence with strong causal identification existed. However, we believe that because this is the *only* example of an under-16 social media ban globally, it would have been ill-advised to ignore the evidence that does exist.

Included sources were charted in a structured extraction table recording citation, year, jurisdiction, policy area, design and sample, circumvention measure, identification strategy, key findings, relevance to problematic use, and source URL or identifier. For causal interpretation, studies were also classified informally by identification strength. Randomised experiments, preregistered synthetic-control designs, difference-in-differences, interrupted time-series, and credible within-person longitudinal designs were treated as stronger evidence for effects of restrictions. Cross-sectional surveys, qualitative studies, regulatory audits, platform-compliance reports, and media investigations were treated as weaker for causal inference but often stronger for documenting implementation mechanisms, workarounds, and external validity.

Findings

Implemented restrictions show weak and incomplete exclusion effects, and may lead to substitution

Evidence from implemented restrictions is mixed but consistent on one point: legislation does not necessarily lead to reliable exclusion from restricted platforms or services. South Korea’s “shutdown” or “Cinderella” law offers useful evidence. The law aimed to block under-16s from playing online games between midnight and 6 a.m., primarily to reduce internet addiction and improve sleep, achieved by blocking accounts associated with children. Korean game companies were expected to verify users’ real name and age at sign-up, and the curfew then blocked accounts known to belong to under-16s. The strongest of the studies associated with this law is the difference-in-differences paper by Choi et al. (2018), using more than 240,000 observations from the Korea Youth Risk Behavior Web-based Survey. It estimated an initial reduction in daily internet use of approximately 3.6 minutes, with no detectable change in the policy’s aims of internet addiction or sleep hours. Looking at the same law, Lee et al. (2017) used regression-based policy evaluation on Korean youth survey data and reached a broadly compatible verdict by a different route: the policy did not achieve its aims and in fact appears to have increased

juvenile internet hours, alongside a reduction in the probability of internet addiction of approximately 0.7 percentage points and an increase in sleep duration of approximately 1.5 minutes. Neither study measured individual circumvention, and both are vulnerable to substitution into mobile gaming or other non-covered activities.

China's online-gaming restrictions provide further evidence on substitution and evasion. Zhou et al. (2024) surveyed 430 heavy adolescent gamers after the 2021 policy. While the majority of users reported complying with the policy, a minority used family IDs or rented accounts and the majority of users (59%) shifted from gaming to watching short-form videos. Zendle et al. (2023) look at the same Chinese law, and conduct before-and-after and within-participant analyses, finding that there is no evidence that the gaming restrictions actually reduced heavy play, and in fact their point estimate suggests that it may have *increased* heavy play, although this finding was below their pre-specified threshold for practical significance. Sensitivity analyses found a similar absence of reduction after the stricter 2021 adjustment. It should be noted that the study observes gamer profiles generally rather than users confirmed not to be adults, and the authors lacked age metadata, so this evidence should be understood as suggesting that mandates did not reduce heavy play among observed accounts rather than evidence on the activities of young players *specifically*.

Adult-content age-verification laws offer the clearest quasi-experimental evidence of circumvention and shifts to other non-compliant platforms. Lang et al. (2026) use a preregistered synthetic-control multiverse design exploiting staggered U.S. state age-verification laws. Three months after implementation, searches for a main compliant adult site fell 51%, searches for a large non-compliant rival rose 48.1%, and VPN searches rose 23.6%. Because the outcome is a normalised search-interest index rather than direct traffic, absolute changes in consumption cannot be inferred. Nonetheless, the study points towards the fact that age-verification measures can result in users simply switching to non-compliant sites rather than staying away from such sites entirely.

Other studies add to the concern that substitution is a real issue regarding these bans: a working paper by Agarwal et al. (2026) uses domain-state-week traffic data across fifteen U.S. states and a staggered difference-in-differences design to study adult-content age-verification laws. They find that although platforms that require age verification lose traffic, large ungated sites that did not require age verification increased their traffic by approximately 35 to 40%, suggesting cross-platform substitution rather than simple demand reduction. In this study, however, it is notable that fringe sites did not see comparable gains, and that substitution was concentrated in mainstream platforms that did not adopt age-verification requirements. The authors estimate that overall demand fell by 23%, indicating that there was some demand reduction alongside the substitution, rather than this being a pure case of cross-platform substitution.

Collis and Eggers (2022) provide further experimental evidence from social-media restrictions: in a randomised trial among university students, participants restricted to ten minutes per day on Facebook and other social platforms replaced their use of such platforms with instant messaging, mostly WhatsApp, and did not reduce total time spent on digital devices at all. These studies are not youth-ban evaluations, and they do not show that substitution necessarily leads to harm. They do, however, support the narrower claim that restricting named platforms or categories can simply result in users switching to similar services that have not been restricted.

A study by Madio et al. (2026) provides further quasi-experimental evidence on substitution. In December 2020, MindGeek (the owner of Pornhub) deleted approximately 80% of its video library overnight following pressure from both the media and payment processors. Using a difference-in-differences design across eight countries, the authors find that MindGeek lost 36 to 42% of visits within a month. The evidence suggests that users substituted into other platforms that had not deleted content: mainstream rivals like XVideos and XNXX gained 9 to 23% over six months, and fringe sites saw traffic increase by 56 to 58%. Search engines and aggregator websites became the main route to these fringe alternatives, overtaking direct navigation. Once again, the setting is not a youth ban, but it is evidence that restricting content on a better-moderated platform redirects users toward less-regulated venues.

The early Australian under-16 social-media-ban evidence is descriptive rather than causally identified, but it is the most directly relevant implementation evidence for broad social-media age bans. Before implementation, eSafety reported widespread underage use of social and messaging services: 84% of 8 to 12-year-olds had used at least one social-media or messaging service since the start of 2024 despite a limit to over 13s; 40% of users had their own account; 80% of account-holders had help setting up accounts; and only 13% had an account shut down for being underage (eSafety Commissioner, 2025). Qualitative research commissioned by eSafety found that children and parents anticipated workarounds, including false ages, use of a parent's face, and migration to similar platforms (eSafety Commissioner, 2026a). A large ABC Behind the News non-probability survey of 17,144 respondents aged 9 to 15 found that almost 80% used social media in some capacity, 75% of social-media users had their own account, 75% of social-media users did not plan to stop after the ban, and only 6% thought the ban would work (ABC Behind the News, 2025). These sources were not outcome evaluations, but they established that the policy was introduced into an environment where noncompliance and workarounds were predictable.

Early Australian data suggest the social media ban has had a noticeable but limited effect. A survey commissioned by the eSafety Commissioner in late January and early February 2026 asked around 900 parents and carers about their children's (aged 8 to 15) use of restricted platforms. The share of children reported to have their own accounts dropped by roughly 18 percentage points, from just under half to about a third. Yet the ban appears far from watertight: the majority of children who previously had accounts on Instagram, Snapchat, and TikTok still had them, as did around two-thirds of former Facebook users and nearly half of those who had been on YouTube. It is worth noting, however, that these figures rely on parental awareness rather than children's own accounts of their behaviour, meaning the true extent of continued access may be higher than the numbers suggest.

Of the parents who had a child who was able to keep access to their account, 66.8% said the platform had not yet asked for age verification, 23.7% said age verification got the child's age wrong, 23.3% reported another workaround such as makeup or someone else's ID, 10.3% said the child opened a new account after deactivation, 7.6% reported parent/carer help, and 6.7% reported VPN use (eSafety Commissioner, 2026c). The official compliance report update has evidence that points in the same direction: large numbers of children were able to keep access to their old accounts, create new accounts, or bypass the systems designed to keep them off the platforms. Platforms were also found to keep prompting children to attempt age-verification, which allowed them to repeat their attempts until allowed onto the platform (eSafety Commissioner, 2026b).

Bursztyjn et al. (2026) provide the strongest early teen-level evidence on this mechanism, although it should be noted that this is a working paper. In surveys of Australian teenagers four months after implementation, they estimate that only around 27% of banned 14 to 15-year-olds comply, and find that most banned teens regard circumvention as easy; their mechanism survey suggests compliance is unlikely to become self-sustaining because teens report needing roughly two-thirds of peers to stop using social media before they would stop themselves. Other Australian evidence suggests the same pattern of partial visible reduction and substantial leakage. Publicly reported YouthInsight polling, commissioned by the Molly Rose Foundation, suggests that 61% of 12 to 15-year-olds who previously had restricted accounts still had access to at least one account (Molly Rose Foundation and YouthInsight, 2026).

Age gates are leaky, and much evasion is ordinary or family-mediated

Across Europe, EU Kids Online found high levels of underage social-network profile ownership despite nominal age thresholds: 38% of 9 to 12-year-olds had a social-network profile, and approximately one in five had a Facebook profile (Livingstone et al., 2011, 2013). In the Net Children Go Mobile sample, 42% of 9 to 12-year-olds had a Facebook profile. Logistic models indicated that underage profile ownership was predicted by social motivations, communication skills, peer mediation, age, and online disinhibition rather than by a clear problematic-use profile (Barbovschi et al., 2015).

U.S. evidence shows that underage access was often family-assisted. Boyd et al. found that 36% of surveyed parents reported their child had joined Facebook before 13, and 68% of those parents helped create the account (Boyd et al., 2011). Pew found that 39% of online teens reported falsifying their age to gain access to a website or account, while a separate measure found that 26% of teen social-media users posted fake information, such as a fake name, age, or location, to help protect their privacy (Madden et al., 2013). UK evidence from the Advertising Standards Authority and Ofcom commentary similarly indicates that children use false ages and may be treated as adults by platform or advertising systems, although it should note that the sample size in the ASA study was extremely small (Advertising Standards Authority, 2013; Ofcom, 2022).

Discussion and Policy Implications

The evidence supports the conclusion that youth internet bans and age-gating methods are currently an ineffective system for excluding under-16s from accessing banned or regulated websites. In general, circumvention appears to be straightforward and common, and most of the empirical evidence suggests that under-16s find it relatively easy to discover and make use of publicised workarounds. Workarounds include (but are not limited to): using VPNs, using accounts from adult relatives or friends, using accounts retained from prior to the implementation of the ban, and other more technically sophisticated methods.

The strongest policy-relevant lesson is that bans should be assumed to come with a large degree of circumvention, and that any serious proposal for bans should accept that it is extremely unlikely to be wholly effective, and may well be mostly ineffective. This applies especially to particularly determined and technically sophisticated under-16s.

How platforms are run matters too. The Australian compliance evidence isn't just a story about the methods that teenagers use to avoid exclusion, it also shows how platform design choices can reduce

enforcement. eSafety reported practices like nudging underage users to try age checks, letting them keep trying until they get a 16+ result, doing little to stop them from making new accounts, and giving poor ways to report problems (eSafety Commissioner, 2026b). These are predictable problems with incentives: strictly keeping kids out means fewer users, more friction, and more complaints to handle, while leaky systems keep users on the platform and push the burden of proof onto regulators, parents, and children. Any honest assessment of these policies therefore has to treat platform compliance as something shaped by the regulation itself, rather than as a neutral matter of implementation.

For proposed under-16 social-media bans, the central policy implication is that the exclusion effect should be treated as uncertain but likely to be small for many adolescent age groups. The available evidence does not support treating a statutory age threshold as a reliable way to keep adolescents off popular social media platforms. Apparent reductions in visible, compliant-platform use are not reliable evidence of exclusion unless residual access, repeat account creation or other methods of bypassing age assurance, shared credentials, VPN use, and migration to other services are also reliably measured. The evidence points toward bypass or substitution with other platforms that are plausibly equally pernicious, rather than exclusion.

This matters because the headline policy claim behind under-16 bans is often stronger than the evidence can bear. If the policy objective is to remove under-16s from social media, the empirical record suggests that the instrument is likely to be weak: age gates are leaky, enforcement is lacklustre, family enabled access is common, and determined adolescents can often re-enter through new accounts, borrowed credentials, misclassification, platform non-enforcement, or substitute services. Australia's early evidence is consistent with this interpretation, but should still be read as early implementation evidence rather than a mature causal evaluation.

Reported account ownership fell, but large shares of people who previously held accounts were able to continue accessing their accounts, and multiple sources document platform errors, repeated attempts, new accounts, parent assistance, and VPN use (eSafety Commissioner, 2026c,b). The number of account takedowns, therefore, is not particularly strong evidence that the policy is working to keep young people away from banned platforms. Policymakers must account for the fact that to get a true sense of whether the restrictions are effective, there must be a clear account of the extent of residual access, the number of young people able to access to the platform through appeals, borrowed accounts, VPN usage, or other methods, and whether young people typically migrate to other platforms with algorithmic features or that do not offer protections for young people that are comparable to those offered by banned platforms.

It should also be noted that displacement to other platforms may actually *increase* risks for some young people. Lang et al. (2026), a quasi-experimental study using a preregistered synthetic-control multiverse design, found that age verification laws reduced the number of searches for a restricted platform, but resulted in large increases in searches for a non-compliant rival and VPNs. While this is only one study and not limited to under-16s, the core takeaway remains that even when people are excluded from restricted platforms, they may substitute their usage of those platforms with usage of less regulated, smaller platforms with weaker moderation and less willingness to comply with age-gating mechanisms. A full evaluation of any content restriction for under-16s should look seriously at the extent to which this redirection is likely to occur, rather than just assuming that users who are effectively restricted from banned platforms will replace those with less harmful alternatives.

The best current evidence suggests that youth internet bans and age restrictions are weak exclusion tools. Evidence of reduced access to compliant services should be interpreted alongside leakage, adaptation, and substitution. Early Australian evidence shows this pattern in a broad social-media-ban setting: account removals occurred, but substantial continuing access remained. The strongest available conclusion is that a proposed under-16 social-media ban should not be expected to keep most affected adolescents off social media; for many adolescents, the exclusion effect may be small or nonexistent. Future evaluations should measure not merely whether accounts are removed, but what young people do next, who is excluded, who circumvents, and whether substitute services or access routes are less safe than the banned platform.

References

- ABC Behind the News. (2025). *What 17,000-plus young people think about the social media ban*. <https://www.abc.net.au/news/2025-12-02/what-young-people-think-about-social-media-ban-btn-survey/106058976>
- Advertising Standards Authority. (2013). *ASA research shows children are registering on social media under false ages*. <https://www.asa.org.uk/news/asa-research-shows-children-are-registering-on-social-media-under-false-ages.html>
- Agarwal, S., Ananthakrishnan, U. M., Madio, L., Mitchell, M. F., Quinn, M., & Reggiani, C. (2026). *Are you 18? Age verification and adult-only consumption in the US* [SSRN working paper]. <https://doi.org/10.2139/ssrn.6619018>
- Barbovschi, M., Machackova, H., & Olafsson, K. (2015). Underage use of social network sites: It's about friends. *Cyberpsychology, Behavior, and Social Networking*, *18*(6), 328–332. <https://doi.org/10.1089/cyber.2014.0553>
- Boyd, D., Hargittai, E., Schultz, J., & Palfrey, J. (2011). Why parents help their children lie to Facebook about age: Unintended consequences of the Children's Online Privacy Protection Act. *First Monday*, *16*(11). <https://doi.org/10.5210/fm.v16i11.3850>
- Bursztyn, L., Duckworth, A., Jiménez-Durán, R., Leonard, A., Milojević, F., Roth, C., & Sunstein, C. R. (2026). *Why bans fail: Tipping points and Australia's social media ban* (Becker Friedman Institute for Economics Working Paper No. 2026-57). University of Chicago.
- Choi, J., Cho, H., Lee, S., Kim, J., & Park, E.-C. (2018). Effect of the online game shutdown policy on internet use, internet addiction, and sleeping hours in Korean adolescents. *Journal of Adolescent Health*, *62*(5), 548–555. <https://doi.org/10.1016/j.jadohealth.2017.11.291>
- Collis, A., & Eggers, F. (2022). Effects of restricting social media usage on wellbeing and performance: A randomized control trial among students. *PLOS ONE*, *17*(8), e0272416. <https://doi.org/10.1371/journal.pone.0272416>

eSafety Commissioner. (2025). *Children and social media*. Australian Government.

<https://www.esafety.gov.au/research/children-and-social-media>

eSafety Commissioner. (2026a). *Attitudes of children and parents to social media age restrictions ahead of implementation*. Australian Government.

<https://www.esafety.gov.au/sites/default/files/2026-03/Attitudes-of-children-and-parents-social-media-age-restrictions-Mar-2026.pdf>

eSafety Commissioner. (2026b). *Social media minimum age: Compliance update*. Australian Government.

<https://www.esafety.gov.au/sites/default/files/2026-03/SocialMediaMinimumAgeComplianceUpdateMarch2026.pdf?v=1774905032806>

eSafety Commissioner. (2026c). *Under the new age restrictions: Early insights from Australian parents*. Australian Government.

<https://www.esafety.gov.au/sites/default/files/2026-03/Under-the-new-age-restrictions-Early-insights-from-Australian-parents-March2026.pdf?v=1775024586194>

Lang, D., Listyg, B., Ross, B. V., Musquera, A. V., & Sanderson, Z. (2026). Age verification and public adaptation: A pre-registered synthetic control multiverse. *Journal of Law & Empirical Analysis*.

<https://doi.org/10.1177/2755323X251415424>

Lee, C., Kim, H., & Hong, A. (2017). Ex-post evaluation of illegalizing juvenile online game after midnight: A case of shutdown policy in South Korea. *Telematics and Informatics*, 34(8), 1597–1606.

<https://doi.org/10.1016/j.tele.2017.07.006>

Livingstone, S., Olafsson, K., & Staksrud, E. (2011). *Social networking, age and privacy*. EU Kids Online. <https://eprints.lse.ac.uk/35849/>

Livingstone, S., Olafsson, K., & Staksrud, E. (2013). Risky social networking practices among underage users: Lessons for evidence-based policy. *Journal of Computer-Mediated Communication*, 18(3), 303–320.

<https://doi.org/10.1111/jcc4.12012>

Madden, M., Lenhart, A., Cortesi, S., Gasser, U., Duggan, M., Smith, A., & Beaton, M. (2013). *Teens, social media, and privacy*. Pew Research Center.

<https://www.pewresearch.org/internet/2013/05/21/teens-social-media-and-privacy/>

Madio, L., Mitchell, M., Quinn, M., & Reggiani, C. (2026). Asymmetric content moderation in search markets: The case of adult websites. SSRN Working Paper.

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5106235

Molly Rose Foundation & YouthInsight. (2026). *Australia's social media ban: Is it working?* Molly Rose Foundation.

Ofcom. (2022). *How old is your child online? The risks of children using false ages on social media*. <https://www.ofcom.org.uk/online-safety/protecting-children/risks-of-children-using-false-ages-on-social-media>

Zendle, D., Flick, C., Gordon-Petrovskaya, E., Ballou, N., Xiao, L. Y., & Drachen, A. (2023). No evidence that Chinese playtime mandates reduced heavy gaming in one segment of the video games industry. *Nature Human Behaviour*, 7, 1753–1761. <https://doi.org/10.1038/s41562-023-01669-8>

Zhou, X., Liao, M., Gorowska, M., Chen, X., & Li, Y. (2024). Compliance and alternative behaviors of heavy gamers in adolescents to Chinese online gaming restriction policy. *Journal of Behavioral Addictions*, 13(2), 687–692. <https://doi.org/10.1556/2006.2024.00021>