Discussion Paper Series

IZA DP No. 19015

October 2026

Teen Social Media Use Is a Collective Action Problem and Requires Urgent, Concerted Effort Beyond Bans

Stefanie Schurer
The University of Sydney
and IZA@LISER

The IZA Discussion Paper Series (ISSN: 2365-9793) ("Series") is the primary platform for disseminating research produced within the framework of the IZA@LISER Network, an unincorporated international network of labour economists coordinated by the Luxembourg Institute of Socio-Economic Research (LISER). The Series is operated by LISER, a Luxembourg public establishment (établissement public) registered with the Luxembourg Business Registers under number J57, with its registered office at 11, Porte des Sciences, 4366 Esch-sur-Alzette, Grand Duchy of Luxembourg.

Any opinions expressed in this Series are solely those of the author(s). LISER accepts no responsibility or liability for the content of the contributions published herein. LISER adheres to the European Code of Conduct for Research Integrity. Contributions published in this Series present preliminary work intended to foster academic debate. They may be revised, are not definitive, and should be cited accordingly. Copyright remains with the author(s) unless otherwise indicated.

Teen Social Media Use Is a Collective Action Problem and Requires Urgent, Concerted Effort Beyond Bans*

Abstract

In December 2025, Australia became the first country to set a minimum age of 16 for social media accounts. In this survey I argue that teen social media use is a collective action problem. Causal evidence shows that social media is habit forming and harms mental health. Drawing on work on social traps and a recent model of bad networks (Akerlof, Holden, and Thornton, 2026), I explain why marginal policies may fail. The fight against the tobacco industry offers a useful parallel, although the health and productivity costs of social media are not yet quantified. Worrying statistics emerged from PISA and from the global rollout of 3G mobile internet on the learning costs of digital distraction. Two evaluations of the Australian ban find that most banned teens still use the platforms, with compliance falling over time. A study of school phone bans offers a silver lining. A national ban is necessary but not sufficient. Enforcement at the level of individual accounts, coordinated action in schools and families, and digital literacy taught with a mentor mindset would give the ban the best chance of bringing teenagers along.

JEL classification

D62, D91, I12, I18, I21, L86

Keywords

social media, adolescents, collective action, network externalities, age restrictions, school phone bans

Corresponding author

Stefanie Schurer
stefanie.schurer@sydney.edu.au

* *Acknowledgement:* This survey article is based on current research conducted by the author at the University of Sydney on the costs and benefits of hard government intervention to deal with problem behaviours, including youth social media use. A first version of this paper was presented at lectures organised by the Parents and Citizens Association of St Ives High School on 14 August 2026 and 23 September 2026 (slides available on request). The author thanks Marit Skivenes and Richard Holden for valuable feedback, and Richard Holden for approving the reproduction of the model dynamics of Akerlof, Holden and Thornton (2026) in Figure 2. The author acknowledges funding and support from an internal University of Sydney grant scheme to conduct this research. Artificial intelligence (Claude, Anthropic) was used to redraw Figures 1 and 2 and to refine the manuscript, including grammar checks and proofreading.

I Introduction

Social media is part of almost every Australian teenager's day. In the 2025 Mission Australia Youth Survey (La Sala et al., 2025), 97 per cent of respondents reported daily use and 38 per cent reported three or more hours a day. Parents, teachers and governments worry about what this does to young people's mental health, sleep and learning, and exposure to harmful content and exploitation. Social psychologist Jonathan Haidt hypothesises in his 2024 book The Anxious Generation that social media use is to blame for the mental health crisis among adolescents (Haidt 2024). In December 2025, Australia acted on this worry. The *Online Safety Amendment (Social Media Minimum Age) Act 2024* requires designated platforms to take reasonable steps to prevent people under 16 from holding accounts. More than 20 countries have since moved to restrict children's access to social media. In her recent State of the Union speech, the President of the European Commission von der Leyen proposed an EU Kids Act that would ban social media for children under 13, limit those aged 13 to 15 to accounts supervised by parents and introduced safe design features for older teenagers (European Commission, 2026).

Why is the world jumping suddenly so quickly to hard government intervention regarding youth social media access? This review will make the argument that teen social media use is a burning collective action problem and requires an urgent, concerted effort to find solutions. The key problem is that each teenager decides whether to use social media given what her peers do. A teenager who leaves a platform alone loses contact with friends and with the information that circulates among them. Yet many teenagers would prefer a world in which nobody in their peer group used the platforms. Everyone would be better off under a rule that binds the whole group, but no individual has the incentive or the power to create one. Akerlof et al. (2026) show that a 'bad network', one that makes its members worse off, is easy to establish but hard to dislodge. Bursztyn et al. (2025) show in incentivised experiments that many active users of TikTok and Instagram would prefer these platforms not to exist.

If the problem is collective, individual solutions will fail. Information campaigns, parental control software and small taxes change the incentives of one user at a time. They leave the equilibrium in place. A hard rule that moves the whole peer group at once can, in principle, break the trap. A rule that is not enforced cannot. The first Australian evidence illustrates this point. Two studies, one published in the *British Medical Journal* (Barnes et al., 2026) and one

circulated as a National Bureau of Economic Research working paper (Bursztyn et al., 2026), find that most banned teens continued to use social media in the months after the ban took effect. The second study finds that compliance fell from about 27 per cent in April 2026 to about 6 per cent in June 2026.

The fight over teen social media access resembles the fight over tobacco access in the twentieth century. Both products are habit forming, both spread through adolescent peer groups, and in both cases the firms knew more about the harm than their users did. Like the tobacco companies in 1998, Meta has now settled with almost every US state (Mawajdeh and Rameswaram, 2026). The two cases may still differ in one important way. The health care costs and lives lost to smoking were well documented by 1998, but it is early days in understanding the medical and human capital costs, or as some victims described in their own court case against Meta of a "lost childhood", of excessive social media use.

The article proceeds as follows. Section II presents a framework for what is at stake. Section III defines the collective action problem and sets out the model of Akerlof et al. (2026). Section IV reviews the evidence for the claim that social media harms wellbeing, that it is habit forming, and that users feel trapped. Section V reviews the early evidence on the Australian ban. Section VI turns to schools. Section VII draws lessons from tobacco control. Section VIII sets out implications for policy.

II What is at stake? A conceptual framework

Figure 1 summarises how I see the problem as a human development economist concerned with the productivity and wellbeing costs of unregulated social media use among teenagers. At its centre are the long-run outcomes that matter for young people and, through them, for the whole economy: health, skills, identity, a sense of belonging and wellbeing. Skills formed in adolescence build on earlier skills and shape adult productivity, including both cognitive and non-cognitive skills (Cunha and Heckman, 2007, Heckman 2006). High schools and universities are critical learning places where labour-market relevant skills are being shaped (Schurer 2017, Kassenboehmer et al. 2018). The non-cognitive skills of young people shape their future labour market outcomes (Heckman et al. 2006), and both cognitive and non-cognitive skills of a population predict long-run economic growth (Hanushek and Woessmann, 2008, 2015; Balart et al. 2018). Mental health problems, on the other hand, experienced already in childhood reduce earnings in adulthood (Goodman et al., 2011) and stay within families across generations (Johnston et al., 2013), with potential long-term damage.

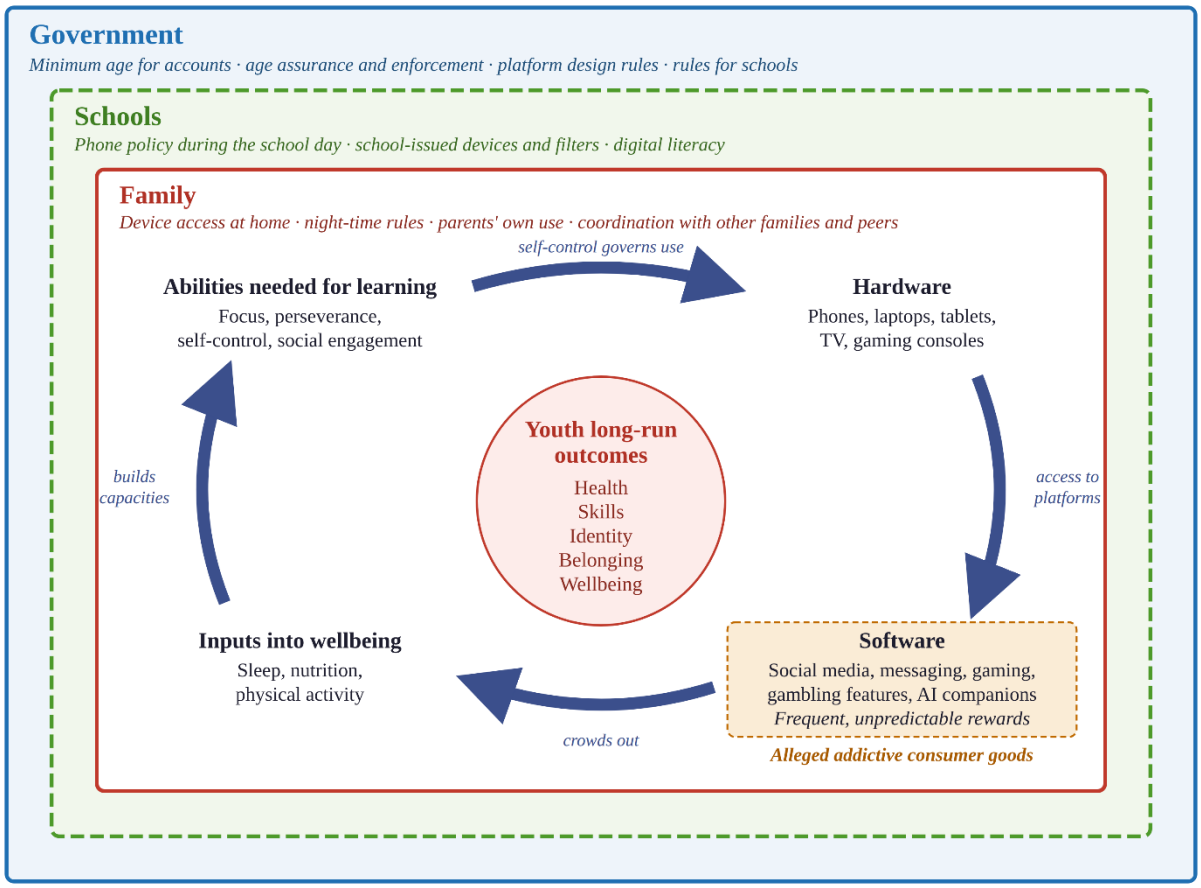
Around the centre sits a cycle of how digital technologies, both hard- and software, will ultimately affect the economy's full potential. Hardware such as phones, laptops, tablets and gaming consoles gives access to software and the internet. The software includes social media, messaging, games with gambling features and, increasingly, artificial intelligence (AI) companions, and unlimited access to everything that the internet has to offer. Heavy use of this software competes for time with the basic inputs into wellbeing: sleep, nutrition and physical activity. These inputs shape the abilities needed for learning, such as focus, perseverance, self-control and social engagement. Weaker self-control in turn makes it harder to regulate device use, which closes the cycle.

The cycle is embedded in three institutions. Families can control access to devices, software and internet at home. Schools can control device use during the school day and may require students to use a laptop in class. Governments set the rules for access to platforms and how schools use digital technologies. Each actor can intervene in the cycle, but as I and others have argued, none can do so alone, because teenagers' desire for digital technology use will depend on what their peer networks do. If families block access to certain online information, or soft- and hardware, then teenagers can bypass those restrictions with the help of their peers or school devices. Restrictions at schools are meaningless if digital hygiene in the home is lacking. National or state guidelines are powerless if schools, families and peers do not play along.

How common is smartphone use and online activity among Australian teenagers? Australian data on teenagers' smartphone use shows that unhealthy digital habits, such as chronic nighttime interruptions are widely spread. Already in 2014, so more than 12 years ago, more than 70 per cent of Australian adolescents keep two or more devices in their bedroom overnight (Gamble et al., 2014). Most recent figures suggest that almost all adolescents aged 15-17 years old have at least one device in their bedroom (94%), while among 12-14 years olds 80.8% of girls and 67.7% of boys do so (Australian Bureau of Statistics, 2026).

Recent data also shows that 34% of primary-school-aged children reported nighttime phone use at least once a week, while in high school this share is doubled (66% in girls, 58% in boys) (Centofanti et al., 2024). While in primary school not more than 1 in 20 children (6% in girls and 8% in boys) would never achieve eight hours of uninterrupted sleep, in high school these shares tripled to 17% in girls and 13% in boys. Psychological distress was reported by one in five primary-school-aged children and almost two in five girls in secondary school (38% in girls, 23% in boys).

Figure 1 Digital technology, youth development and the institutions that shape them



Notes: The cycle links access to devices, use of software, basic inputs into teenager wellbeing and the abilities needed for learning. Each institutional layer can intervene in the cycle. Peer networks span all three layers. Source: Author's own illustration.

The key concern is both the quantity and quality of screen time and social media exposure. The most recent figures suggest that Australia's young people spent 14 hours online per week on average, almost 100% report daily social media use, and almost two in five spend more than 3 hours per day on social media. The key concern is how and why social media is used. An increasing number of young people are exposed to extremely violent (self-harm, abuse) or painful content, often through algorithm-driven feeds that prioritise engagement over safety (La Sala et al., 2025). Smart phones and social media are also the gateway to encounter (violent forms of) pornography, with a median age of first exposure of 13 for boys and 16 for girls (Lim et al., 2017; Crabbe et al., 2024), and online gambling opportunities that reportedly are sought by one in three 12- to 17-year-old Australians (Saunders and Harrington, 2025).

Mental health issues, loneliness and deprivation of sleep are all potential consequences of heavy screen time use and are the sort of consequences highlighted in recent book The Anxious Generation (Haidt, 2024). Although hard evidence is scarce, heavy screen use might impact the biology of brain development. Data emerging from the Adolescent Brain Cognitive Development Study, which follows almost 10,000 United States children from age 9 or 10 into adolescence, shows that Children with more screen time at baseline had more symptoms of attention deficit hyperactivity disorder two years later (Shou et al., 2025). They also had a thinner cortex in three regions involved in attention, working memory and cognitive control. Cortical thickness partly accounts for the association between screen time and symptoms. Thus, heavy screen time may have long-term and irreversible consequences for the skill development of children.

However, correlation is not causation. It might as well be that teenagers who overuse social media most are the ones who have already mental health issues or have thinner cortical thickness and are thus more vulnerable to self-control issues that social media triggers. Odgers (2024) in a recent book review in Nature made exactly this point. She argued that the existing evidence does not support the claim that social media is causing an epidemic of mental illness among adolescents. She notes that hundreds of studies have produced no, small or mixed associations, and that most of the data are correlational. Where studies follow young people over time, the direction runs the other way: adolescents who already have mental health problems use social media more often, or in different ways, than their healthy peers.

These considerations make clear why policymakers, educators and parents alike are confused while seeking solutions to address their teenagers' (potential) over-use of digital technologies. In the next section, I will outline why it is so hard to address this problem, and why it can be considered a classic case of a collective action problem, while the follow up section will discuss the limited causal evidence, that does exist, on the impact of social media use on mental health and its addictive properties.

III Teen Social Media as a Collective Action Problem

(i) What is a collective action problem?

A collective action problem arises when choices that are rational for each individual produce an outcome that the members of a group judge worse than an alternative they could reach together, and no individual can bring about that alternative by changing her own behaviour. Two conditions must hold for the collective action problem to occur. First, each individual does

better by following her private incentive, given what the others do. Second, all or most members of the group would be better off if everyone acted differently.

Olson (1965) showed that large groups rarely organise themselves to pursue a shared interest. Each member gains little from contributing and can benefit from the efforts of others, so large groups need coercion or selective incentives to act. Hardin (1968) described this tragedy of the commons in the context of grazing and maintaining collective resources like meadows. Each herder gains the full benefit of adding an animal to a shared pasture but bears only a fraction of the cost of overgrazing. The pasture is ruined even though no herder wants that outcome. Hardin argued that appeals to conscience do not solve such problems. His remedy was mutual coercion, mutually agreed upon by those affected.

Platt (1973) generalised Hardin's argument, explaining that “social traps” can explain why certain behaviours that are rewarding in the short run might lead to long-run loss for the group (see also Cross and Guyer, 1980). Platt distinguished social traps from one-person traps, in which an individual's own short-term rewards lead to her own long-term loss, as with addiction. Dawes (1980) formalised group traps as social dilemmas, in which each individual does better by defecting, but all do better if all cooperate.

Schelling (1973) modelled these social traps as binary decisions with externalities, where individual choices affect someone else’s choice, and vice versa. One of his examples was taken from professional ice hockey players in the United States. Until the late 1960s, few wore helmets, yet most would have favoured a rule that would require every player to wear one.

Teen social media use is in some way very similar, but unlike helmets, it fits both one-person and social traps. It is both a one-person trap, because a substantial share of use reflects self-control problems (Section IV). It is also a social trap, because for each teenager, being on a platform is better than being the only one off it. For the group, a world in which no one is on the platform may be better than a world in which everyone is. Haidt (2024) uses the term collective action problem in the same sense.

(ii) A Formal Model of Bad Networks

Akerlof, Holden and Thornton (2026) provide a formal model of this argument applied to teen social media use. They are motivated by the evidence linking platforms such as Instagram and TikTok to poor mental health. They model how social networks like Instagram or TikTok generate 'rat races' of online presence and discuss what the externalities created by them means for public intervention.

The generalised version of the model starts with the observation that teenagers (referred to a agents) decide whether to join a social network. In their model q is defined as the share of teenagers who join and $\varphi(q)$ an increasing function with $\varphi(0) = 0$ and $\varphi(1) = 1$. An Agent i's utility dependent in the binary choice of either joining the platform ($x_i = 1$) or not ($x_i = 0$):

$$U(x_i = 1) = a\,\varphi(q) + \varepsilon_i \quad \text{if agent } i \text{ joins } (x_i = 1) \qquad (1)$$

$$U(x_i = 0) = b\,\varphi(q) \quad \text{if agent } i \text{ stays off } (x_i = 0). \qquad (2)$$

The parameter a captures the network externality for members and b the externality for non-members. The term ε_i is an idiosyncratic taste for the network, distributed symmetrically around zero on the interval $[-c, c]$. The authors assume $a > b$, so that participants benefit more from the externality than non-participants.

A network generates positive externalities if $a > 0$ (good network) and negative externalities if $a < 0$ (bad network). Without idiosyncratic tastes ($c = 0$), individuals prefer joining when anyone else joins ($q > 0$) and are indifferent when no one joins ($q = 0$). In this case, either everyone joins or no one joins, and the authors show that there are two (Nash) equilibria. With idiosyncratic tastes ($c > 0$), equilibrium participation is determined by equation (3) below.

Aggregate welfare is the sum of the aggregate welfare of those joining and those not joining. If the network has positive externalities, then aggregate welfare rises with the number of people joining, while social harms are produced in a negative spillover network as more participants join.

The logic of the model is illustrated in Panel (a) of Figure 2 (based on Schelling (1973) for the linear case $\varphi(q) = q$). The two lines show the payoff to joining (blue, solid line) and to staying off (red, dashed line) as the share of peers on the platform rises. The 'joining' line lies above the 'staying-off' line as long as at least someone else joins ($q>0$), so each teenager prefers to join whatever others do. Both lines slope down, so each additional member makes everyone worse off which leads to a collective loss, but the individual loss is greater for the teenagers who stay off the platform. The group moves from point A, where no one is on the platform, to point B, where everyone is, and everyone ends up worse off than at A.

The collective loss is the difference between the payoff when no one joins and the payoff of the members. This figure illustrates the collection action problem: The more teenagers join, the smaller the incentive for teenagers, who have not joined yet, to stay off the platforms, but the more join the greater the collective loss.

So why is social media an example of a bad network? Akerlof et al. (2026) explain that social media generates a 'rat race' of social comparison that is welfare-reducing for participants ($a<0$), while the experience for non-participants is amplified further by social exclusion ($b<a$), which teenagers care strongly about to avoid.

Joining gives agent i a payoff of $a\varphi(q) + \varepsilon_i$. Staying off gives $b\varphi(q)$. She joins if the first exceeds the second, which rearranges to $\varepsilon_i > (b - a)\varphi(q)$. Because the model assumes $a > b$, the right-hand side is negative and grows more negative as q rises. Two things follow:

- Some people who dislike the network ($\varepsilon_i < 0$) still join. Their dislike is outweighed by the cost of being left out.
- The bar for joining falls as more peers join. Each new member lowers the threshold for the next. This drives the cascade.

Equilibrium participation is the share of agents whose taste exceeds the threshold:

$$q^{NE} = 1 - F((b - a)\,\varphi(q^{NE})), \qquad (3)$$

where F is the distribution function of tastes for the network ε_i and NE stands for *Nash Equilibrium*.

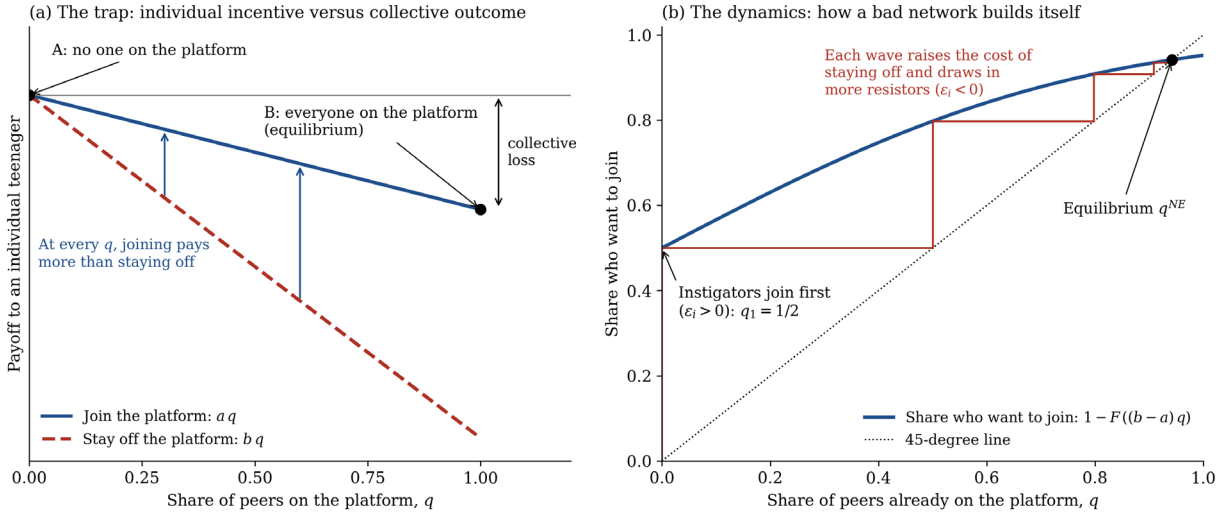
The paper distinguishes three types of potential participants: The network instigators ($\varepsilon_i > 0$), the network resistors ($\varepsilon_i < 0$), and neutral agents ($\varepsilon_i = 0$), terms borrowed from Granovetter (1978). Relating this concept to heterogeneous preferences over social comparison, instigators would correspond to individuals who benefit from the rat race directly either through visibility and/or status (e.g. influencers). Resistors could be individuals who distaste social comparisons or care strongly about privacy risks. Resistors only join the network if a large enough number of peers join the network and when the benefit of social connection outweighs their intrinsic distaste for the platform. When $c=0$, all agents are neutral; when $c>0$, there is a combination of instigators and resistors.

The process of how the network evolves is explained in Fig. 2 (b). Three key results emerge from the model. First, bad networks are easy to establish. Instigators join even when no one else is on the network. They have a zero threshold. In the model, half of the teenager population would be in that group ($q=0.5$) at the starting point of the network would be willing to join the network. Their presence raises the cost of staying off (the blue solid line), which draws in some of the resistors, which raises the cost again, and a cascade of joining follows.

Thus, the second result is that a bad network is always too large in equilibrium. Akerlof et al. (2026) propose that in equilibrium participation exceeds the level that maximises aggregate welfare, and that a positive mass of members would be better off if the network did not exist.

Third, the trap can be complete with everyone joining. If idiosyncratic tastes are weak relative to the gap between the two externalities that the network generates ($c \leq a - b$), everyone joins. If they are weaker still ($c < -a$), everyone would be better off with no network at all.

Figure 2 Teen social media as a collective action problem: the trap and its dynamics



Notes: Illustration of the model in Akerlof et al. (2026) with φ(q) = q. Panel (a) uses a = −1 and b = −3 and omits the idiosyncratic term ε_i. Panel (b) uses a = −1, b = −3 and ε_i normally distributed with mean zero and standard deviation 1.2; the staircase traces the cascade from an empty network. Parameter values are illustrative. Source: Author's illustration based on Akerlof et al. (2026) and Schelling (1973). Figure drawn with Anthropic Claude with permission for reprinting the model dynamics from Akerlof, Holden and Thornton.

Why would a social network satisfy $0 > a > b$? The authors argue that this happens in rat races. Each agent chooses whether to join and whether to exert costly effort, at cost C, in a zero-sum competition for esteem. Joining the network raises the weight an agent places on esteem by α, which captures how salient likes, followers and views make social comparison. The network also delivers a connection benefit β per member.

The teenager's utility function of being in the network is expressed in Eq. (4), where q_e is the share of teenagers who exert effort and q_x the share of teenagers on the network.

$$U(x_i, e_i, q_e, q_x) = \underbrace{(1 + \alpha \cdot x_i)R(e_i, q_e)}_{\text{Esteem component}} + \underbrace{\beta \cdot q_x \cdot x_i}_{\text{Connection component}} - \underbrace{C \cdot e_i}_{\text{Cost of effort}} + \epsilon_i \cdot x_i \qquad (4)$$

When α is large enough that members compete ($\alpha \geq C - 1$), i.e. the returns to effort exceed its cost, the model reduces to equation (1) with $a = \beta - \alpha - 1$ and $b = -1$. The network is bad when $\beta - 1 < \alpha < \beta$. The rat race outweighs the value of connection, yet staying off means falling behind in a competition that others continue to play. Hence, the negative aspect of the network (the rat race) outweighs the positive aspect of the network (connecting peers).

Raising α weakly increases the size of the network. A platform can therefore grow by making social comparison more salient, even though this harms its users. Akerlof, Holden and Thornton (2026) argue that media-rich social platforms like Instagram or TikTok strategically use visibility metrics (likes, views, follower counts), driving the competitive "rat race". This has direct implications for regulation of a bad network.

(iii) What the Model Implies for Policy

Economists typically approach an externality with a Pigouvian tax to reduce consumption. Akerlof et al. (2026) show that a tax equal to the marginal externality at the social optimum restores efficiency when network effects are linear. However, in the case of nonlinear network effects there can be several equilibria, and a tax calibrated to the marginal externality may fail to fix a network that is already established. When almost everyone is on the network, a single resistor who leaves barely changes q, the share of people on the network. She compares being on a bad network with everyone else to being off it while everyone else stays, and the second option looks worse. Many resistors would have to leave at once, and none of them internalises the benefit of a collective exit. Dislodging an established bad network may therefore require a more extreme policy, perhaps a tax high enough to destroy the bad equilibrium, or an outright ban. The authors note that policy may need to be dynamic, with a large initial intervention to tip the network followed by a smaller correction.

Taxing teenagers is politically difficult, or in the case of a free consumer good like Instagram, Facebook or TikTok impossible, so a quantity restriction such as Australia's ban can take the place of a tax. Status-seeking early adopters act as instigators who draw resistors onto the platform. By keeping the instigators off, the ban removes the impulse that tips a peer network into the high-participation equilibrium. The model also carries a warning, visible in Figure 2(b). A ban works only while it holds participation down. In the model, if enforcement lapses, the same cascade rebuilds the network from wherever enforcement leaves off.

IV Why Social Media is Not an Ordinary Consumer Good

Standard consumer theory treats time spent on a platform as revealed preference. If teenagers choose social media, they must value it. The model in Section III rests on three premises that challenge this presumption: participation harms users, use is hard to control, and users are trapped. This section reviews the evidence for each.

(i) Social Media Harms Mental Health

The most convincing causal evidence comes from the early expansion of Facebook. Between 2004 and 2006, Facebook became available at United States colleges in a staggered order. Braghieri et al. (2022) combine the roll-out dates with repeated waves of the National College Health Assessment using a difference-in-differences estimation approach. The arrival of Facebook increased depression by 9 per cent and generalised anxiety disorder by 12 per cent. The use of mental health services rose among the students most at risk. The authors trace the effects to unfavourable social comparisons that the platform made possible.

Experimental evidence points in the same direction. Allcott et al. (2020) recruited 2,743 adult Facebook users in the United States and randomly paid some of them to deactivate their accounts for the four weeks before the 2018 midterm elections. Deactivation improved an index of subjective wellbeing by 0.09 standard deviations, with small but significant improvements in happiness, life satisfaction, depression and anxiety. The authors note that this is about 25 to 40 per cent of the effect of psychological interventions such as self-help therapy and group training. Deactivated users spent more time offline, including with friends and family. Several weeks after the experiment ended, they still used the Facebook app 22 per cent less than the control group. Both studies concern young adults or adults.

(ii) Social Media is Habit Forming

Allcott et al. (2022) ask whether people use social media because they value it or because they struggle to stop. They recruited about 2,000 adults in the United States with Android smartphones and measured their use of six apps: Facebook, Instagram, Twitter, Snapchat,

web browsers and YouTube. The experiment ran from March to July 2020. Participants were randomised independently into two treatments.

The first treatment paid participants US$2.50 per hour for reducing their use over three weeks. It reduced use by 56 minutes per day while the payments were in effect, a fall of 39 per cent. In the three weeks after the payments ended, the group still used the apps 19 minutes less per day. The authors interpret this persistence as evidence of habit formation.

The second treatment gave participants a tool to set limits on their own future use. It reduced use by 22 minutes per day, or 16 per cent, over 12 weeks. Although there was no incentive to do so, 78 per cent of participants set binding limits and kept using them. On average, participants were willing to pay US$4.20 for three weeks of access to the tool. The authors conclude that people have self-control problems and are at least partly aware of them. Their model implies that self-control problems account for 31 per cent of social media use.

This finding matters for welfare analysis. If almost a third of use reflects self-control problems, observed demand overstates the value users place on the product. In Platt's (1973) terms, social media is a one-person trap as well as a group trap. The evidence comes from adults. Direct causal evidence for adolescents is scarce.

(iii) Social Media Users are Trapped

Bursztyn et al. (2025) provide the most direct evidence that social media users are trapped. In incentivised experiments with United States college students, they measure the value of TikTok and Instagram in two ways. The first measure asks how much a user must be paid to deactivate her own account, holding everyone else's use fixed. By this measure, the platforms generate positive welfare. The second measure asks what a user would need to be paid to deactivate if all her peers deactivated as well. This accounts for the loss that non-users suffer when their peers remain on the platform

The authors calculated that the average student needed to be paid about US$59 to leave TikTok for four weeks but would pay about US$28 to have TikTok deactivated for everyone.

Bursztyn et al. (2025) call this a product market trap, in line with Akerlof et al. (2026)'s model predictions (see their Lemma 2).

(iv) Are bans the only solution to solve the product market trap?

Haidt (2024), who argues that the move of childhood onto smartphones in the early 2010s contributed to rising anxiety and depression among adolescents, requires swift and hard intervention through collective action. He proposes four avenues for intervention: First, parents should give their children no smartphone before high school, which is age 14 in the US. Second, governments should ban social media access before age 16. Third, schools should ban phones to make them distraction free zones for learning and interaction. Fourth parents and schools alike should encourage more independence and free play. Haidt stresses that a family that adopts these norms alone imposes a social cost on its own child. The norms work when many families adopt them together, and when schools come along on the journey.

In the next section, I will discuss whether Australia's social media account ban was able to move teenagers' compliance above the point at which quitting becomes self-reinforcing. If compliance stays below that tipping point, it will erode. Section V shows that this is what happened in the first months of the Australian ban. In Section VI, I will discuss solutions at the school level.

V Does Australia's Ban Work? Early Evidence

(i) What is Australia's Social Media Account Ban?

Australia's *Online Safety Amendment (Social Media Minimum Age) Act 2024* was enacted in December 2024 and took effect in December 2025. It applies to platforms designated as age restricted, including Facebook, Instagram, Threads, TikTok, Snapchat, X, YouTube, Reddit, Kick and Twitch. Messaging services, Discord, Roblox and YouTube Kids are not covered. Platforms must take reasonable steps to prevent users under 16 from creating or keeping accounts. The Act does not prescribe a technology. Platforms that fail to comply face civil penalties of up to A$49.5 million (Barnes et al., 2026). The Act imposes no penalty on young users or their parents, and parents cannot consent on behalf of a younger child. The ban restricts accounts, not content. Young people can still view most content without logging in.

(ii) Causal Evidence

Barnes et al. (2026) recruited 436 adolescents aged 12 to under 17 between September and December 2025. Recruitment ran through channels aimed at parents, including social media advertising, previous trial participants, newsletters and the networks of research and health organisations. Parents gave consent and adolescents gave assent. Participants completed surveys immediately before the ban and about three months later. The study retained 408 participants, or 94 per cent, at follow up.

Three months after the ban, more than 86 per cent of participants under 16 had used a banned platform in the previous week. The majority of them did so through their own accounts: 54 per cent of those aged 12 to 13 and 68 per cent of those aged 14 to 15. Daily social media use among 14- and 15-year-olds fell from 78 to 69 per cent. Among those aged 16 and over, who were not subject to the ban, it rose from 80 to 89 per cent. Age checks were weak. Among teens still using banned platforms, 24 to 39 per cent reported that a platform had simply asked their age, and 13 to 27 per cent had uploaded a selfie. Requests for official identification were rare. A minority used workarounds such as a fake account (15 to 19 per cent), a private browser (6 to 11 per cent), or someone else's account (29 per cent of those aged 12 to 13 and 9 per cent of those aged 14 to 15). Use of virtual private networks was rare. More participants reported spending more time on messaging apps and outdoors than reported spending less.

The study's main causal analysis uses a sharp regression discontinuity design at the age threshold of 16. It finds no statistically significant discontinuity in daily use or in time spent on social media. The estimates rest on 87 observations within the optimal bandwidth for daily use, and the confidence intervals are wide. The authors report that the study was underpowered and that some tests raise concerns about the continuity assumption. The findings therefore do not rule out modest effects. They do rule out the large reduction in use that the policy intended.

Bursztyn et al. (2026) surveyed 507 Australian teenagers aged 14 to 18 between March and April 2026, recruited through an opt-in online panel run by the youth membership platform Student Edge. A mechanism survey of 464 teenagers followed in April 2026, and a follow-up wave of 792 teenagers in June 2026. The comparison of interest is between banned 14- and 15-year-olds and unbanned 16- to 18-year-olds in the same panel.

In April 2026, 63.8 per cent of banned 14- and 15-year-olds reported using a banned platform in the past week, compared with 87.1 per cent of unbanned 16- to 18-year-olds. Using the older group as the counterfactual implies compliance of about 27 per cent. A randomised response

module, which protects individual answers, gave almost identical prevalence. Knowledge of the ban was near universal (98 per cent), and 22 per cent of them erroneously believed there were personal consequences for breaking it. Three quarters described circumvention as easy or very easy, and 60 per cent had not had their account removed. Social pressures are the main reason why teenagers do not comply. Among banned teens who still used the platforms, 52 per cent gave at least one social reason, most often that their friends still used them or a fear of missing out. Among teens who believed that four or five of their five closest friends still used banned platforms, 80.9 per cent used them too. Among those who believed that three or fewer did, the share was 36.5 per cent. When asked directly, 13–18-year-old teens require 66 per cent of their peers to stop using social media before complying with the ban themselves, a robust finding across all considered reference groups. They believed that only 30 per cent of peers had stopped. Combining the two, the authors estimate that the only stable level of compliance is about 17 per cent, well below the observed 27 per cent.

The June 2026 wave confirmed that the ban was ineffective to reduce network membership. Past-week use among banned 14- and 15-year-olds rose to 92.2 per cent, and implied compliance fell to about 6 per cent. Among 93 teens observed in both waves, implied compliance fell from 25 to 7 per cent. Half of banned teens reported that they had stopped using a banned platform at some point but resumed once they saw that their peers had not stopped. Close to two thirds said that more of their friends used social media than a month earlier.

Who complies matters as much as how many. Of the respondents, 45 per cent said that teens who comply are less popular than those who do not, and only 4 per cent said they are more popular. Among banned 14- and 15-year-olds, non-compliers reported about one and a half times as many Instagram followers as compliers. The authors show formally that when influential teens hold out, the compliance curve bends in a way that keeps equilibrium compliance low. The compliance estimates are consistent with those of Barnes et al. (2026).

Both studies differ in sampling, timing and method, but they agree on the main facts. Most banned teenagers kept using social media. Most kept their own accounts. Age assurance was weak and easy to get around. Teens themselves explain their behaviour with reference to their peers. This is what Akerlof et al. (2026) predicted for a ban that is announced but not enforced at the level of individual accounts. A ban coordinates behaviour only if teens expect that their friends will be off the platforms.

The evidence shows that a weakly enforced ban leaves the bad equilibrium in place. Bursztyn et al. (2026) note that penalties on platforms were increased over the period of their study, yet compliance fell. The size of a fine that a platform might face does not enter a teenager's decision. What enters her decision is whether her own account and her friends' accounts are removed. Their model also implies that age assurance must reach socially central users, because the equilibrium depends on who complies and not only on how many.

It would be wrong, however, to read this evidence as a case against bans. Australia's account-level ban is the first of its kind, and other countries are following. Turkey passed a comparable law in April 2026, Norway announced a bill the same month, and Spain, Denmark and France are drafting their own versions. Skivenes (2026) weighs the arguments for and against the Norwegian proposal. She concludes that a ban is an important first step, because children need protection while they are still young. She also answers the objection that a ban violates children's rights. The UN Convention on the Rights of the Child obliges states to protect children from harm, not only to respect their right to participate. If autonomy means acting from one's own values rather than independence from adults, then removing access to platforms designed to override self-control is a condition for autonomy, not a violation of it. Skivenes is equally clear that a ban alone is not enough. If platform design stays the same, a ban may simply delay the harm to 17- and 18-year-olds and let platforms claim the problem is solved. She argues that the ban should be one part of a wider regime. This regime would include rules on algorithmic design, rapid takedown obligations, child-safety defaults in gaming and messaging apps, and digital literacy in schools, developed with young people at the table.

VI Schools: Distraction, Learning and Phone Bans

If national bans, if poorly enforced, do not work as well, perhaps schools could play the role of supporting institution. Schools have been heavily affected by the mobile (smart) phone epidemic among students and the consequences of distraction for learning are now visible in data. Thus, schools are not only well equipped to reach young people and affect their daily routines, but also have the incentives to address teen social media and smartphone overuse.

(i) Evidence from Programme for International Student Assessment

Data from the OECD Programme for International Student Assessment (PISA) show that performance fell sharply in 2022, the largest decline in the history of the assessment. Between 2018 and 2022, the OECD average fell by 15 points in mathematics and 10 points in reading (OECD, 2023a). The decline in mathematics was about three times larger than any previous

change between consecutive assessments. The pandemic explains part of the decline. The OECD notes, however, that performance in reading and science had been falling in many countries before 2018.

PISA also documents the role of digital distraction in the classroom. On average across OECD countries, 59 per cent of students reported that other students' use of digital devices distracted them in at least some mathematics lessons (OECD, 2023b). Students who were distracted by peers' device use scored 15 points lower in mathematics, which the OECD equates to about three quarters of a year of learning. Students who spent at maximum one hour a day at school on devices for leisure scored 49 points higher in mathematics than those who spent five to seven hours (OECD, 2024). Across OECD countries, 45 per cent of students reported feeling nervous or anxious when their phone was not near them (OECD, 2023b).

Australia performs worse than the OECD average on these measures. In 2022, 40 per cent of Australian students reported getting distracted using digital devices in most or every mathematics lesson (OECD average: 30 per cent), and 37 per cent reported getting distracted by other students using devices (OECD average: 25 per cent) (OECD, 2023c). The PISA 2025 results show little change. In 2025, 39 per cent of Australian students reported that their peers were distracted by digital devices in most or every science lesson (OECD average: 28 per cent). This is despite 85 per cent of Australian students attending schools that ban phones, up from 53 per cent in 2022 (OECD average: 49 per cent) (OECD, 2026).

The OECD figures show that digital distraction in class is common and that it is partly a peer effect, because one student's device use distracts others. But correlation is not causation. The PISA evidence cannot make causal claims that screen time and online engagement caused the decline in test scores observed in OECD data. However, a recent study by Jain and Stemper (2026) can do this. Exploiting the staggered rollout of 3G mobile internet across and within 82 countries between 2000 and 2018 and linking geospatial coverage data to the test scores of 2.5 million 15-year-olds, they offer reliable evidence that mobile internet access harmed children’s test scores. They. Their difference-in-differences design compares students in the same country and type of location (village, town or city) before and after 3G arrived. 3G coverage raised the probability that a student had a smartphone at home by 10 percentage points. It increased time online by about five hours a week, a 17 per cent rise, and most of the increase occurred outside school. Test scores in mathematics, reading and science fell by 0.04 to 0.08 standard deviations. This is equivalent to about a quarter of a year of learning. The losses are driven by exposure

during adolescence rather than childhood, and they are largest among lower-achieving students. There is suggestive evidence of larger losses for girls, for students whose parents did not attend university, and in low- and middle-income countries.

The mechanisms point to social media rather than to learning online. Daily social media use rose by 17 percentage points, while the use of the internet for schoolwork fell by 6 percentage points. Students' sense of belonging at school and their ease of making friends declined by 0.09 to 0.16 standard deviations. These results show that the spread of mobile internet lowered learning at scale, and that most of the extra time online occurred outside the classroom.

(ii) Evidence from School Phone Bans

Hence, schools should be concerned about the use of mobile phones among their students both on and off school grounds. However, schools can only influence the mode and intensity of digital engagement of their students while they are on campus. Many schools have started banning smartphone use during school hours in Australia and worldwide. Do these bans work and in what ways?

Allcott et al. (2026) provide the most comprehensive causal evidence on school phone bans to date. They study about 5,000 United States schools that adopted lockable (Yondr) pouches, in which students' phones stay sealed during the school day. The authors combine location data on phone activity with teacher and student surveys and school records and compare adopting schools with similar schools that had not yet adopted.

The phone ban was indeed effective in reaching its primary objective. The share of students that teachers observed using phones in class for personal reasons fell from 61 to 13 per cent. On-campus phone activity measured from location data fell by about 30 per cent by the third year. However, the first year brought adjustment costs with students clearly struggling with the ban. School suspension rates rose by 15 per cent, and student-reported wellbeing fell by 0.2 standard deviations. Nevertheless, by the second year, wellbeing had recovered and exceeded its initial level by 0.16 standard deviations. This effect is larger than the effect of a four-week Facebook deactivation on adult wellbeing (Allcott et al., 2020).

Perhaps surprisingly, and disappointingly, there were no substantial improvements in tests scores on average. But this could have been due to heterogeneous effects by year levels. The

study found small improvements in mathematics tests scores for high school students (year levels 10-12) and small declines for students in middle schools (year levels 7-9).

Two lessons can be drawn from these findings. First, a rule that is enforced for everyone in the peer group at the same time can change behaviour quickly, even among teenagers. The pouch is a coordination device. It removes the social cost of being the only student without a phone. Second, less phone use does not automatically mean more learning. The study cannot tell us what replaced phone use in class, and a phone ban leaves other devices, such as school laptops, in use.

VII Lessons from Tobacco Control

The parallels between calls for regulation of social media companies and the regulation of the tobacco industry in the twentieth century are striking. First, both products are habit forming, and both are consumed more than users themselves would choose. Gruber and Kőszegi (2001) showed that smokers with self-control issues benefit from taxes that help them follow through on their own plans to quit. Gruber and Mullainathan (2005) found that cigarette taxes made smokers happier. Allcott et al. (2022) found a similar pattern for social media. Users anticipate self-control problems and take up tools that limit their own use.

Second, both products are networking consumer goods for adolescents. Youth smoking responds strongly to the behaviour of peers (Powell et al., 2005; Krauth, 2006). Economists have modelled it as a binary choice with social interactions and possibly multiple equilibria (Brock and Durlauf, 2001). This is the same class of model that underlies Schelling (1973) and Akerlof et al. (2026). Harris and López-Valcárcel (2008) model smoking as a signal of being cool. They find that peer effects make price increases about 60 per cent more effective. Glaeser et al. (2003) call this the social multiplier. Peer effects entrench use when a behaviour is popular. They also speed up its decline once it falls out of favour.

Third, in both cases the firms knew more about the harm than their users did. The tobacco litigation of the 1990s exposed internal industry research on addiction. It ended in 1998, when 52 US states and territories settled with the four largest tobacco companies under the Master Settlement Agreement (MSA). The companies agreed to pay about US$206 billion over 25 years, in part to compensate states for the medical costs of smoking, and to stop marketing to young people. Four states had settled separately beforehand. The settlement also established

and funded the Truth Initiative, an organisation 'dedicated to achieving a culture where all youth and young adults reject tobacco'.

The case against Meta follows a similar pattern. In October 2023, a coalition of state attorneys general led by California sued Meta in *People of* the *State of California v.* Meta *Platforms, Inc.* The states did not target harmful content. They targeted the design of the apps themselves, such as infinite scroll, autoplay, constant notifications, like buttons and beauty filters (Mawajdeh and Rameswaram, 2026). The plaintiffs argued that Meta built these features to hook children and teens. They also alleged that Meta ignored or concealed internal studies, such as Project Myst, which found that vulnerable young users showed compulsive use under stress. Meta denied that it engineered addiction into its products.

In late August 2026, a few days into the trial, Meta settled with 47 states, the District of Columbia and a US territory. New Mexico had reached its own agreement earlier (Mawajdeh and Rameswaram, 2026). Meta agreed to pay about US$17.1 billion over 10 years, although over one third of the settlement payment (US$5 billion) will take effect only if YouTube and TikTok agree to similar terms (Mawajdeh and Rameswaram, 2026).

It also agreed to change its products for teenagers, promising to introduce a daily usage limit of two hours across its apps, which excludes messaging and long-form content, to switch off notifications by default at night and during school hours, and to replace the personalised feed with a reverse chronological one. Meta has also publicly called on Snap, TikTok and YouTube to adopt the same rules for consistency across platforms.

The structure of the settlement fits the model in Section III. A usage limit on one company's platforms invites substitution. A teenager who reaches the limit on Instagram can move to YouTube or TikTok. A rule that binds one firm protects no one if the peer group moves elsewhere. Meta therefore has an interest in extending the rules to its competitors. In the absence of federal legislation, the settlement may become the main set of rules for social media in the United States. It remains an agreement with one company, on terms that the company judged it could live with (Mawajdeh and Rameswaram, 2026).

Tobacco control offers three lessons for social media. The first is that no single instrument worked. Smoking fell through a long series of measures: evidence of harm, taxes, advertising bans, smoke-free laws, restrictions on sales to minors, litigation and, eventually, plain

packaging. Australia was the first country to introduce plain packaging in 2012. It is now the first country with a minimum age for social media.

The second lesson is that smoking declined faster once it lost social status. Christakis and Fowler (2008) show that people quit in connected groups and that the remaining smokers moved to the edges of their social networks. Bursztyn et al. (2026) find the opposite for social media today, because teenagers see those who comply with the ban as less popular. Regulation that ignores status is likely to struggle.

The third lesson is that change takes time. Already in 1950, Doll and Hill (1950) established the link between smoking and lung cancer, yet regulation took decades to catch up with the evidence. Coordinated international action came only with the WHO Framework Convention on Tobacco Control (World Health Organization, 2003).

I admit that the analogy has limits. The evidence on tobacco rests on large and unambiguous physical harm, and the health care costs and lives lost to smoking are well documented. The evidence base on the causal effects of social media on mental health is limited to date (Braghieri et al., 2022), while Odgers (2024) rightfully emphasises that reverse causality might be a defining feature of the social media-mental health nexus. The medical and human capital costs of excessive social media use are not yet known. Social media also offers benefits of connection that cigarettes do not. And social media is free to the user, so price instruments such as taxes, which were central to tobacco control, are not available. Minimum age laws, usage limits and rules on platform design must do more of the work (Akerlof et al., 2026).

Nevertheless, the main lesson from tobacco control still applies. Arguments in favour of individual responsibility are not satisfactory. Change in teenagers social media use requires coordinated, enforced and sustained regulation.

VIII Implications for Policy

The evidence reviewed here supports four conclusions for policy. Each follows from treating teen social media use as a collective action problem.

First, a legal minimum age is the right instrument in principle. Government is the only actor in the system that can move the entire peer group at once. Akerlof et al. (2026) show that marginal

instruments, such as a small tax, may fail to dislodge an established bad network, while a ban can. Hardin (1968) and Ostrom (1990) point to the same requirement from different directions: a rule that binds everyone, backed by monitoring and sanctions.

Second, the ban must be enforced where the decision is made, namely at the level of individual accounts. Both studies on Australia's social media account ban for the under 16s show that self-declared age and selfie checks are easily defeated, and that most banned teens kept their existing accounts. The regulator should require platforms to detect and remove existing underage accounts, to use stronger forms of age assurance, and to report compliance in a way that can be independently audited. Bursztyn et al. (2026) suggest that enforcement should reach socially central users, because their continued presence sustains the pull of the platform. They also propose aligning the age limit with the school year rather than the birth date, so that teenagers are not sitting next to classmates who may legally use the platforms.

Third, the ban must be accompanied by action on close substitutes and on platform design. Messaging apps, Discord, Roblox and games with gambling features sit outside the current list. Barnes et al. (2026) report that teens shifted time towards messaging apps. A narrow ban invites substitution towards platforms that carry similar risks. The government has indicated that the list of covered platforms may be expanded, and the evidence supports doing so where the same network features are present. Rules on design features such as public like and follower counts would further weaken the rat race that makes the network bad in the first place as proposed in Akerlof et al. (2026).

Fourth, government, schools and families must act together. Schools should adopt and enforce phone restrictions for the whole school day and apply the same logic to school-issued and school-required laptops. Private laptops remain currently the Trojan horse through which teenagers access inappropriate content and inappropriate use during class time, often sanctioned by teachers to keep students quiet during class.

The Australian PISA results prompted the Government of Australia's most populous state, New South Wales, to address online engagement and screen time use in the school curriculum heads on. From the start of 2027, the NSW Government will cap the time students in NSW public schools spend learning on laptops, tablets and desktop computers, from 15 minutes a day in Kindergarten to Year 2 to a quarter of the time allocated to each subject in Years 7 to 10 and half in Years 11 and 12. The caps will become mandatory in 2028 (NSW Government, 2026).

As of writing this article, Western Australia followed suit and will introduce its own, unique, screen time policy.

Jain and Stemper (2026) showed that most of the additional time online that followed the arrival of mobile internet occurred outside school. School phone or laptop bans on their own therefore address only part of the problem, and families and peer networks must be involved too.

Families can coordinate with other families, as Haidt (2024) recommends, so that no child is the only one without an account. Yeager (2024) argues that adolescents respond to respect and a sense of purpose more than to control. Rules that are explained and set together with teenagers are more likely to hold. Teenagers themselves appear open to help. Of the teens surveyed by Bursztyn et al. (2026), 72 per cent said they would prefer an app that limits their own use to the ban. Allcott et al. (2022) show that adults take up commitment tools when they are offered. Such tools can complement, but not replace, a coordinated rule.

There will be a central role also for parents to reconsider their own views on and support for government restrictions on their underage children's digital footprint. To date, no data are available on parental attitudes to hard government intervention in children's lives. This, and the overall digital hygiene that parents practice in their homes, is a critical knowledge gap that requires comprehensive data collection.

Thus, teen social media use is not simply a matter of individual choice or parental control. It is a collective action problem in the sense of Hardin (1968), Olson (1965) and Schelling (1973). Causal evidence shows that social media harms mental health and that a substantial share of use reflects self-control problems. Theory and experiments show that users can be trapped in networks they would collectively prefer to leave. Once a peer group is on a platform, no individual teenager and no individual family can leave at low cost.

Australia's minimum age law is the right kind of response to this problem. The first evidence shows that, as implemented, it has not yet shifted the equilibrium. Compliance was low four months after the ban and fell further by six months. Teenagers explain their behaviour with reference to their friends. Evidence from school phone bans shows that coordinated rules can change teenage behaviour quickly when they are enforced for everyone at once.

The lesson for Australia and for the many countries now considering similar laws is clear, and it echoes the history of tobacco control. A ban is necessary but not sufficient. It must be

enforced at the level of individual accounts, extended to close substitutes, and supported by schools and families acting together.

References

Akerlof, R., Holden, R. and Thornton, D.J. (2026), 'Bad Networks', *Journal of Public Economics*, 254, 105574.

Allcott, H., Baron, E.J., Dee, T., Duckworth, A.L., Gentzkow, M. and Jacob, B. (2026), 'The Effects of School Phone Bans: National Evidence from Lockable Pouches', NBER Working Paper No. 35132, National Bureau of Economic Research, Cambridge, MA.

Allcott, H., Braghieri, L., Eichmeyer, S. and Gentzkow, M. (2020), 'The Welfare Effects of Social Media', *American Economic Review*, 110, 629–76.

Allcott, H., Gentzkow, M. and Song, L. (2022), 'Digital Addiction', *American Economic Review*, 112, 2424–63.

Australian Bureau of Statistics (2026), 'Children's Physical Activity, Screen Time and Sleep', Australian Bureau of Statistics, Canberra. Available at: https://www.abs.gov.au/articles/childrens-physical-activity-screen-time-and-sleep (accessed 25 September 2026).

Balart, P., Oosterveen, M., Webbink, D. (2018), Test scores, noncognitive skills and economic growth, Economics of Education Review, Volume 63, 134-153.

Barnes, C., Hall, A., Mantach, S., Oldmeadow, C., Attia, J., Backholer, K., Williams, C., Tefera, Y., Kay, F. and Wolfenden, L. (2026), 'Assessing Early Effects of Australia's Social Media Minimum Age Act on Adolescents' Social Media Use: Observational Study', *BMJ*, 393, e363695.

Braghieri, L., Levy, R. and Makarin, A. (2022), 'Social Media and Mental Health', *American Economic Review*, 112, 3660–93.

Brock, W.A. and Durlauf, S.N. (2001), 'Discrete Choice with Social Interactions', *Review of Economic Studies*, 68, 235–60.

Bursztyn, L., Duckworth, A.L., Jiménez-Durán, R., Leonard, A., Milojević, F., Roth, C. and Sunstein, C.R. (2026), 'Tipping Points and Australia's Social Media Ban', NBER Working Paper No. 35162, National Bureau of Economic Research, Cambridge, MA.

Bursztyn, L., Handel, B.R., Jiménez-Durán, R. and Roth, C. (2025), 'When Product Markets Become Collective Traps: The Case of Social Media', *American Economic Review*, 115, 4105–36.

Centofanti, S., Lushington, K., Wicking, A., Wicking, P. and Dorrian, J. (2024), 'Nighttime Phone Use and Past Exposure to Cyberbullying and Their Impact on Sleep and Psychological Wellbeing in Australian Children Aged 7 to 19 Years', *Adolescents*, 4, 342–61.

Christakis, N.A. and Fowler, J.H. (2008), 'The Collective Dynamics of Smoking in a Large Social Network', *New England Journal of Medicine*, 358, 2249–58.

Crabbe, M., Flood, M. and Adams, K. (2024), 'Pornography Exposure and Access among Young Australians: A Cross-Sectional Study', *Australian and New Zealand Journal of Public Health*, 48, 100135.

Cross, J.G. and Guyer, M.J. (1980), *Social Traps*, University of Michigan Press, Ann Arbor.

Cunha, F. and Heckman, J.J. (2007), 'The Technology of Skill Formation', *American Economic Review*, 97, 31–47.

Dawes, R.M. (1980), 'Social Dilemmas', *Annual Review of Psychology*, 31, 169–93.

Doll, R. and Hill, A.B. (1950), 'Smoking and Carcinoma of the Lung: Preliminary Report', *British Medical Journal*, 2, 739–48.

European Commission (2026), *2026 State of the Union Address by President Ursula von der Leyen*, Strasbourg, 16 September. Available at: https://ec.europa.eu/commission/presscorner/detail/ov/SPEECH_26_1868

Gamble, A.L., D'Rozario, A.L., Bartlett, D.J., Williams, S., Bin, Y.S., Grunstein, R.R. and Marshall, N.S. (2014), 'Adolescent Sleep Patterns and Night-Time Technology Use: Results of the Australian Broadcasting Corporation's Big Sleep Survey', *PLoS ONE*, 9, e111700.

Glaeser, E.L., Sacerdote, B.I. and Scheinkman, J.A. (2003), 'The Social Multiplier', *Journal of the European Economic Association*, 1, 345–53.

Goodman, A., Joyce, R. and Smith, J.P. (2011), 'The Long Shadow Cast by Childhood Physical and Mental Problems on Adult Life', *Proceedings of the National Academy of Sciences*, 108, 6032–7.

Granovetter, M. (1978), 'Threshold Models of Collective Behavior', *American Journal of Sociology*, 83, 1420–43.

Gruber, J. and Kőszegi, B. (2001), 'Is Addiction "Rational"? Theory and Evidence', *Quarterly Journal of Economics*, 116, 1261–303.

Gruber, J. and Mullainathan, S. (2005), 'Do Cigarette Taxes Make Smokers Happier?', *B.E. Journal of Economic Analysis & Policy*, 5, Article 4.

Haidt, J. (2024), *The Anxious Generation: How the Great Rewiring of Childhood Is Causing an Epidemic of Mental Illness*, Penguin Press, New York.

Hanushek, E.A. and Woessmann, L. (2008), 'The Role of Cognitive Skills in Economic Development', *Journal of Economic Literature*, 46, 607–68.

Hanushek, E.A. and Woessmann, L. (2015), *The Knowledge Capital of Nations: Education and the Economics of Growth*, MIT Press, Cambridge, MA.

Hardin, G. (1968), 'The Tragedy of the Commons', *Science*, 162, 1243–8.

Harris, J.E. and López-Valcárcel, B.G. (2008), 'Asymmetric Peer Effects in the Analysis of Cigarette Smoking among Young People in the United States, 1992–1999', *Journal of Health Economics*, 27, 249–64.

Heckman, J.J. (2006), 'Skill Formation and the Economics of Investing in Disadvantaged Children', *Science*, 312, 1900–2.

Heckman, J.J., Stixrud, J. and Urzua, S. (2006), 'The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior', *Journal of Labor Economics*, **24**, 411–82.

Jain, R. and Stemper, S. (2026), 'Growing Up Online: Global Evidence on (3G) Mobile Internet and Adolescent Learning', Working Paper, Department of Economics, University of Zurich, Zurich.

Johnston, D., Schurer, S. and Shields, M. (2013), 'Exploring the Intergenerational Persistence of Mental Health: Evidence from Three Generations', *Journal of Health Economics*, 32, 1077–89.

Kassenboehmer, S., Leung, F. Schurer, S. (2018). 'University education and non-cognitive skill development'. Oxford Economic Papers, 70(2), 538–562

Krauth, B.V. (2006), 'Simulation-Based Estimation of Peer Effects', *Journal of Econometrics*, 133, 243–71.

La Sala, L., Filia, K., Gao, C.X., Baker, D., Browne, V., Brennan, N., Freeburn, T., Boon, B. and Teo, S.M. (2025), *Social Media and Young People in Australia: Findings from the 2024 Mission Australia Youth Survey*, Orygen, Melbourne, and Mission Australia, Sydney.

Lim, M.S.C., Agius, P.A., Carrotte, E.R., Vella, A.M. and Hellard, M.E. (2017), 'Young Australians' Use of Pornography and Associations with Sexual Risk Behaviours', *Australian and New Zealand Journal of Public Health*, 41, 438–43.

Mawajdeh, H. and Rameswaram, S. (2026), 'Meta Suddenly Cares about Kids: Why the Social Media Giant Is Trying to Rope Its Competitors into a Settlement', *Vox*, 3 September. Available at: https://www.vox.com/podcasts/501732/meta-trial-settlement-children-social-media-mental-health-tiktok-youtube-snap

National Association of Attorneys General (1998), *Master Settlement Agreement*, National Association of Attorneys General, Washington, DC.

NSW Government (2026), 'Screentime Limits in All Public Schools to Lift Our Game in the Classroom', Media Release, NSW Government, Sydney, September. Available at: https://www.nsw.gov.au/ministerial-releases/screentime-limited-from-next-year

Odgers, C.L. (2024), 'The Great Rewiring: Is Social Media Really Behind an Epidemic of Teenage Mental Illness?', *Nature*, 628, 29–30.

OECD (2023a), *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*, OECD Publishing, Paris.

OECD (2023b), *PISA 2022 Results (Volume II): Learning During – and From – Disruption*, OECD Publishing, Paris.

OECD (2023c), *PISA 2022 Results (Volumes I and II): Country Notes, Australia*, OECD Publishing, Paris.

OECD (2024), 'Managing Screen Time: How to Protect and Equip Students against Distraction', *PISA in Focus*, No. 124, OECD Publishing, Paris.

OECD (2026), *PISA 2025 Results (Volume I): Country Notes, Australia*, OECD Publishing, Paris. Available at: https://www.oecd.org/en/publications/pisa-2025-results-volume-i-country-notes_2d4ff9ea-en/australia_20d725d6-en.html

Olson, M. (1965), *The Logic of Collective Action: Public Goods and the Theory of Groups*, Harvard University Press, Cambridge, MA.

Online Safety Amendment (Social Media Minimum Age) Act 2024 (Cth).

Ostrom, E. (1990), *Governing the Commons: The Evolution of Institutions for Collective Action*, Cambridge University Press, Cambridge.

People of the State of California et al. v. Meta Platforms, Inc. et al. (2023), United States District Court for the Northern District of California, complaint filed 24 October 2023.

Platt, J. (1973), 'Social Traps', *American Psychologist*, 28, 641–51.

Powell, L.M., Tauras, J.A. and Ross, H. (2005), 'The Importance of Peer Effects, Cigarette Prices and Tobacco Control Policies for Youth Smoking Behavior', *Journal of Health Economics*, 24, 950–68.

Schurer, S. (2017). 'Does education strengthen life skills of adolescents? ' *IZA World of Labor*. June 366.

Saunders, M. and Harrington, M. (2025), *Teenage Gambling in Australia: Rates of Expenditure and Participation among 12–19-Year-Olds*, Discussion Paper, The Australia Institute, Canberra.

Schelling, T.C. (1973), 'Hockey Helmets, Concealed Weapons, and Daylight Saving: A Study of Binary Choices with Externalities', *Journal of Conflict Resolution*, 17, 381–428.

Shou, Q., Yamashita, M. and Mizuno, Y. (2025), 'Association of Screen Time with Attention-Deficit/Hyperactivity Disorder Symptoms and Their Development: The Mediating Role of Brain Structure', *Translational Psychiatry*, 15, 447.

Skivenes, M. (2026), ‘Norway’s Ban on Social Media for Youth: The Arguments For and Against the Proposed Social Media Ban for Under-16s’, Blog post, Centre for Research on Discretion and Paternalism, University of Bergen, 29 May. URL: https://discretion.w.uib.no/2026/05/29/norways-ban-on-social-media-for-youth-the-arguments-for-and-against-the-proposed-social-media-ban-for-under-16s/

World Health Organization (2003), *WHO Framework Convention on Tobacco Control*, World Health Organization, Geneva.

Yeager, D.S. (2024), *10 to 25: The Science of Motivating Young People*, Avid Reader Press, New York.