IFS

Institute for Fiscal Studies

Gregory F. Veramendi

Working paper

26/45

Smartphones and children's mental health

Economic and Social Research Council

SMARTPHONES AND CHILDREN'S MENTAL HEALTH

Gregory F. Veramendi*

July 2026

Abstract

Children's mental health has deteriorated across high-income countries since the early 2010s, coinciding with the mass adoption of smartphones among young people. I provide the first quasi-experimental evidence on delaying children's smartphone ownership through a community-led policy. Since 2024 the Smartphone Free Childhood movement has coordinated families across the United Kingdom through a "Parent Pact" to delay giving children a smartphone until age 14. Using a matched difference-in-differences design, I compare changes in children's mental-health needs in English schools where a large share of parents signed with changes in similar schools where very few did. Where at least one fifth of families signed, school-recorded social, emotional and mental health need falls by around 8 percent of the counterfactual. At the same time, child psychiatric disability claims, determined independently by the Department for Work and Pensions, also fall by about 6 percent in the same communities. The effects are larger among younger children who are less likely to already have a phone and in neighborhoods where more parents sign the pledge. The findings support coordinated ownership delay as a near-zero-cost complement to school phone bans and statutory age limits, and they give parents causal evidence for delaying a child's first smartphone through the primary-school years.

JEL codes: I12, I18, I21, I38, J13.

Keywords: smartphones, child mental health, ownership delay, parental coordination, special educational needs, disability benefits.

*Royal Holloway, University of London. gregory.veramendi@rhul.ac.uk. The draft has benefited from comments and discussions with Arnaud Chevalier, Melanie Lührmann, Darcey Snape, Justine Elbourne-Cload, and Daisy Greenwell. This work received no external funding. Smartphone Free Childhood provided signup data used in this paper, but had no role in its design, analysis, interpretation, or reporting, imposed no conditions on the use of the data or the reporting of results, and the author received no financial or other benefit. The Longitudinal Education Outcomes (LEO) analysis drawn on in Section 2.4 was carried out in the Secure Research Service, part of the ONS. Total word count: approximately 20,000.

1 Introduction

Young people's mental health has deteriorated markedly across high-income countries since the early 2010s (Blanchflower et al., 2025), a trend that coincides with the rapid adoption of smartphones and social media (Haidt, 2024). In England, the prevalence of probable mental disorder among 8–16 year-olds rose from 12.5 percent in 2017 to 20.3 percent in 2023.[1] The policy debate has centered on two levers. Schools have adopted phone bans in growing numbers, but a school ban restricts use only during school hours, leaving pupils' total daily screen time essentially unchanged (SMART Schools, 2025). Several governments have gone further, legislating or preparing minimum ages for social media.[2] A prominent concern with these bans is that they may push young people toward less regulated platforms and less safe websites. There is, in short, little causal evidence on whether smartphones harm children, and on which policies are effective (Orben et al., 2025).

This paper provides evidence on both, evaluating an intervention that delays smartphone *ownership* for children until age 14. The Smartphone Free Childhood (SFC) movement, founded in February 2024, expanded rapidly across the United Kingdom. By early 2026, over 165,000 parents across more than 15,600 schools in the United Kingdom had signed its "Parent Pact," a joint pledge among the families at a school to delay smartphones. Parents fear that withholding a smartphone unilaterally will socially isolate their child. The pact solves this coordination problem by converting a private intention into a visible, school-year-specific commitment. Compliance can be striking. At Cunningham Hill Junior School, Year 6 smartphone ownership fell from 75 to 12 percent within a year, and to 3 percent the following year, against 73 percent nationally at the same ages.[3]

[1] NHS Digital, "Mental Health of Children and Young People in England, 2023: Wave 4 Follow Up to the 2017 Survey."

[2] Australia legislated a minimum age of 16 for social media accounts in late 2024, in force since December 2025. The United Kingdom announced its own under-16 ban in June 2026, with enforcement planned from 2027. France's National Assembly approved an under-15 ban in January 2026, and Denmark's government agreed in November 2025 to legislate one.

[3] Amelia Gentleman, "What happened to the first city that tried to ban smartphones for under-14s?" *The Guardian*, May 2025, reports the fall to 12 percent; the headteacher reports 3 percent for the following Year 6 cohort (interview with the author, July 2026). The national comparator is my calculation from the Ofcom

I use quasi-experimental variation in the fraction of parents signing the pact across state primary schools in England. Drawing on the per-school registration counts from the SFC, I classify a school as *treated* if at least 20 percent of its families had signed by June 2026 (329 schools across 118 local authority districts). Similarly, I classify a school as *low-uptake* if fewer than 5 percent signed (roughly 12,000 schools). An important challenge is that parents at treated schools are more engaged and, on average, more affluent than those at low-uptake schools. I therefore match each treated school to a set of low-uptake schools with similar mental-health needs and affluence before 2024. I then use an event study to trace how outcomes evolve at treated schools relative to their matched controls after the movement's emergence in February 2024. Two administrative outcomes measure children's mental-health needs from different sources. The first is the number of pupils with social, emotional and mental health (SEMH) needs at each school recorded in the annual school census. The second is the number of families near a school claiming Disability Living Allowance (DLA) for children with a psychiatric condition, a benefit adjudicated by the Department for Work and Pensions. In both, treated units track their matched controls closely in the pre-treatment years that are not used for matching.

My main finding is that children's recorded mental-health needs fall substantially where the pact is widely adopted. In schools where at least a fifth of families signed, SEMH need falls by around 8 percent of the counterfactual relative to matched low-uptake schools. The estimate is similar when I compare high- and low-uptake schools within the same district and year. The reduction grows from roughly 7 percent in the first post-treatment year to 9 percent in the second. I find similar changes in mental-health need when it is determined outside of schools. Child psychiatric DLA claims fall by a similar 6 percent in the same communities, corroborating the school result through an independent measurement. As with SEMH needs, DLA claims fall by 4 percent and 8 percent in the first and second post-treatment years. As I measure total pact signup in June 2026, these estimates are likely

Children's Media Literacy Tracker 2024 parents' survey respondent-level data (fieldwork October–November 2024); the published data tables report 79 percent for the 10–12 age band (Ofcom, 2025).

conservative because many parents might not have signed the pact until after I measure the children's needs.

I then examine who is most affected. First, the effects are strongest for children young enough to be affected by the intervention. DLA mental-health claims drop by about 9 percent for 5 to 10 year-olds, whose smartphone ownership is still being delayed. For adolescents (aged 11 to 15), who mostly already owned a phone when the movement started, the estimate is half the size or less. Once children are given support, they usually keep it for years. If I assume that delaying a phone affects only the flow of new benefits, the 10 percent fall in the stock of behavioral claims implies a roughly 20 percent drop in new benefit claims (Section 7). Second, the heterogeneity in the effects is intuitive. The reduction in need strengthens with the fraction of parents signing, from about 3 percent where a tenth of families signed to 8 percent at the one-fifth threshold. It is larger for girls, in line with the social-media literature.

The results are robust to a large set of alternative specifications and falsification tests. First, a synthetic difference-in-differences estimator (Arkhangelsky et al., 2021) and a triple-differences estimator give similar estimates. Second, there is no effect on outcomes that are not plausibly affected by phone ownership. Adult depression in the same communities is flat, as is the autistic-spectrum school category. Within the DLA caseload, the developmental disorders are not affected, while behavioral conditions fall by about 10 percent. Third, re-dating treatment to a placebo pre-treatment year passes the falsification test for all estimators. Fourth, reassigning treatment to same-size, same-clustering sets of never-treated communities places the DLA estimate below every placebo draw and the SEMH estimate is in the far left tail. Finally, the estimates are robust to the low-uptake threshold, the match ratio, matching with replacement, the pre-period length, and dropping the three media-prominent pledge towns.

The estimates should be interpreted with three caveats. First, engaged parents might change their children's mental health trajectory beyond the effect of the smartphone. To explain my results, engagement would have to affect mostly children aged 5 to 10, reduce the receipt of support only for the mental-health outcomes studied here, and begin with the SFC movement in 2024 (Section 7). Second, as in the school-ban literature, the estimates are intention-to-treat effects, since I observe the intended policy but not how it was carried out. Third, I only observe effects in the short run, since the data end two years after the start of the movement. Aside from these caveats, the findings are in line with the hypothesis that early smartphone ownership harms children's mental health, and they indicate that delaying ownership can substantially improve their well-being.

This paper provides, to my knowledge, the first econometric evidence on coordinated smartphone-ownership delay, and it speaks to three literatures. The first is the causal evidence on social media. Braghieri et al. (2022) show that the staggered rollout of Facebook across US colleges worsened the mental health of its young-adult student users, and Allcott et al. (2020) find that experimentally deactivating Facebook improved well-being among adults. Broadband and mobile-internet rollout studies document negative effects on young people across countries (Donati et al., 2025; Golin, 2022; Churchill & Johnson, 2026; Jain & Stemper, 2026). I add causal evidence of the impact of smartphones on primary-school-age children, the group at the center of current policy debates.

The second is the school phone-policy literature. Bans during the school day raise test scores for low-achieving students in England (Beland & Murphy, 2016), while a Swedish replication finds no effects (Kessel et al., 2020). In the most complete evaluation, staggered bans across Norwegian middle schools leave average educational and mental-health outcomes unchanged, but they reduce psychological health-care use among girls, lower bullying for both sexes, and raise girls' grades, with the largest gains for disadvantaged girls (Abrahamsson, 2026). Spanish regional bans reduce bullying (Beneito & Vicente-Chirivella, 2022), and a Florida county ban raised suspensions at first and test scores later (Figlio & Özek, 2025). The most directly relevant evaluation for the UK setting, the SMART Schools Study (SMART Schools, 2025), finds no mental-health benefit from UK in-school restrictions. These policies restrict phone use during school hours but leave ownership and use outside of schools untouched. I add the first evaluation of an intervention on ownership, which operates around the clock and covers the exposure that school bans cannot affect.

The third is the psychology literature documenting associations between device use and children's mental health. A systematic review and meta-analysis of 41 studies covering about 42,000 children and young people finds problematic smartphone use in roughly one in four, with about threefold odds of depression and anxiety and higher odds of poor sleep (Sohn et al., 2019). A systematic review of mobile and wireless device use more broadly finds suggestive associations with poorer mental health, most consistently for bedtime use and partly mediated by sleep, while emphasizing the heterogeneity and high risk of bias of the underlying observational studies (Girela-Serrano et al., 2024). These associations cannot separate the effect of the smartphone from selection into use. This paper provides causal estimates that back the patterns psychologists have described and teachers observe directly: improved attention and behavioral regulation when smartphone ownership is delayed (Section 7).

The remainder of the paper proceeds as follows. Section 2 describes the institutional context, Section 3 the data, and Section 4 the empirical strategy. Section 5 reports the results and Section 6 the falsification and robustness tests. Section 7 discusses interpretation and scope, and Section 8 concludes.

2 Institutional Background

2.1 The Rise of Smartphones and Children's Mental Health

Children's mental health has deteriorated sharply since the early 2010s. In England, the share of 8–16 year-olds with a probable mental disorder rose from 12.5 percent in 2017 to 20.3 percent in 2023, so that roughly one in five children is now affected.[4] Recorded rates of anxiety and depression, together with referrals to children and young people's mental health

[4]NHS Digital, "Mental Health of Children and Young People in England, 2023: Wave 4 Follow Up to the 2017 Survey."

services, have risen in parallel, with the steepest increases among girls and adolescents. The deterioration also appears in objective measures that are not readily explained by changes in diagnosis or help-seeking. Hospital admissions for self-harm among children aged 10–14 in England more than doubled between 2011/12 and 2021/22, from 124 to 307 per 100,000. The suicide rate among 10–24 year-old females rose by more than 80 percent between 2012 and 2018, to its highest recorded level.[5] The trend is not specific to the United Kingdom. Across high-income countries the mental health of the young has worsened over the same period. Blanchflower et al. (2025) find that the long-standing U-shape of well-being over the life cycle has vanished, with the young now reporting the lowest well-being of any age group in dozens of countries. The onset of this decline, around 2011–2012, coincides closely with the period in which smartphones and social media became near-universal among young people, the coincidence that motivates both the research literature on its causes and the policy responses discussed below.

Smartphone ownership among children has risen in step. By autumn 2024, four in ten UK children aged 8–9 owned their own smartphone, rising to 58 percent at age 10 and 89 percent at age 11, the final year of primary school.[6] Many parents buy the first smartphone around the transition to secondary school, citing safety and coordination, and by then ownership is nearly universal. This age profile is central to what follows. An ownership-delay pledge can bind only for children who have not yet acquired a device. A genuine effect of delayed ownership should therefore concentrate among younger children, whose acquisition was actually postponed, rather than among older children who would already own a smartphone at the start of the initiative.

[5] Self-harm: NHS Digital, Hospital Episode Statistics, as compiled in the Nuffield Trust indicator "Hospital admissions as a result of self-harm in children and young people" (ages 10–14: 124 per 100,000 in 2011/12, 307 in 2021/22). Suicide: Office for National Statistics, "Suicides in the UK: 2018 registrations" (females aged 10–24: 1.8 per 100,000 in 2012 to 3.3 in 2018, an 83 percent rise, the highest recorded); see also Bould et al. (2019), "Rising suicide rates among adolescents in England and Wales," *The Lancet*, which documents the corresponding rise at ages 10–19.

[6] Ofcom Children's Media Literacy Tracker 2024, parents of 3–17s, fieldwork October–November 2024 (Ofcom, 2025); the age-band figures are from the published data tables, and the single-year-of-age figures are my calculations from the respondent-level data.

2.2 The Smartphone Free Childhood Movement

The Smartphone Free Childhood (SFC) campaign began with an Instagram post in February 2024. Co-founder Daisy Greenwell wrote that she did not want her eight-year-old daughter to have a smartphone as classmates had begun to get theirs, and that she felt trapped by the prevailing norm. She was looking for other parents who would break it with her by delaying together. The WhatsApp group she and Clare Fernyhough created in response reached its 1,023-member capacity within 24 hours.[7] The speed of the response demonstrates how widely parents shared the concern before SFC existed. Some who later signed the pact had likely already resolved to delay a smartphone, so the signup measure captures pre-existing intent as well as the effect of the pact's coordination. Local SFC groups formed at schools across the United Kingdom within weeks. Greenwell and her husband, Joe Ryrie, went on to found Smartphone Free Childhood as an organization, which in September 2024 launched a "Parent Pact" tool through which parents pledge not to give their child a smartphone before age 14 and to delay social media until age 16.[8] By early 2026, over 165,000 parents from more than 15,600 schools across the United Kingdom had signed. Google search interest confirms the movement's timing (Figure 1). Searches for "smartphone free childhood" are essentially absent in the United Kingdom before its February 2024 launch, corroborating the treatment date. Searches for a "safe phone" for children emerge a year earlier, indicating that awareness had been building in the year before SFC launched.[9]

[Figure 1 about here.]

A series of prominent town-wide pledges accompanied the movement, beginning in St Albans

[7]Greenwell's first-person account is "The Revolution Has Begun in the UK," *After Babel*, `https://www.afterbabel.com/p/the-revolution-has-begun-in-the-uk`; see also "Smartphone Free Childhood: the unstoppable rise," *Positive News*, `https://www.positive.news/lifestyle/smartphone-free-childhood-suddenly-we-were-in-the-eye-of-the-storm/`. The campaign drew on an earlier Spanish parent movement, *Adolescencia Libre de Móviles*.

[8]The second element of the pledge is unlikely to drive the results in this paper. The children studied are mostly too young for the social-media age to bind (less than 14), and the effects concentrate among 5–10 year olds, with a much smaller effect for the 11–15 year olds closest to that margin (Section 5).

[9]Monthly Google Trends search interest for the United Kingdom, the two terms queried together so they share one normalization (one-time export, July 2026).

in May 2024, where the city's primary headteachers jointly urged parents to delay. Appendix Table 9 lists the major initiatives.

Why parents sign is central to what the pact measures. A parent who wants to delay faces a coordination problem. Withholding a smartphone unilaterally risks isolating the child from peers whose social life increasingly runs through the device. The pact converts a private intention into a visible, school-year-specific commitment. According to the organizers, a parent signs less to bind themselves than to provide evidence to other families in the same year at the same school. The specificity of the school-year assurance is something that a general community group cannot provide.[10] Whether a school then reaches high coverage tends to depend less on formal school policy than on whether there is at least one motivated parent champion and supportive local group dynamics. Schools can support this process, but cannot substitute for it.[11] Campaigns can stall when the organizing parent is seen as judging other parents, and teachers are often reluctant to tell uninterested parents what to do. Many attempts fail to take off. Adoption is therefore largely dependent on parental engagement and highly uneven across schools, even within a single town. This school-level variation is what the empirical design exploits (Section 3.1).

2.3 Special Educational Needs and the SEMH Category

Special educational needs (SEN) provision in England is governed by the Children and Families Act 2014 and the statutory SEND Code of Practice (Children and Families Act, 2014; DfE & DoH, 2015). My school outcome is the social, emotional and mental health (SEMH) category, one of the Code's four broad areas of need, which covers anxiety, depression, conduct and attachment difficulties, self-harm, and ADHD. Support operates at two tiers. Most identified pupils receive *SEN support*, arranged by the school's Special Educational Needs Co-ordinator (SENCo) with no external sign-off. Children with greater needs may receive an

[10] Personal communication, Smartphone Free Childhood, July 2026.

[11] For example, see Appendix Table 14 for variation in signup across St. Albans schools, the first prominent town-wide pledge.

Education, Health and Care (EHC) plan. Only the local authority can issue an EHC plan, after a statutory and clinically evidenced assessment, and the plan binds the council and school to specified provision (DfE & DoH, 2015; Latimer et al., 2025). Identified need has risen steeply. By 2025 roughly one in five pupils under 16 held SEN support or an EHC plan, and EHC plans recorded under SEMH more than doubled between 2016 and 2025 (Latimer et al., 2025). My census measure combines the two tiers within the SEMH primary-need category, per 1,000 enrolled pupils.

Because SEMH is identified by a school-level decision rather than a clinical diagnosis, the threshold at which a pupil is recorded varies across schools. Hutchinson (2021) and Hutchinson (2025) show that differences between schools account for roughly two-thirds of whether an otherwise-similar child is identified with SEN, far more than local authorities. SEMH is among the most discretionary categories, with girls recorded markedly less often than boys despite comparable underlying rates of disorder. This makes the *level* of recorded SEMH difficult to compare across schools. It does not prevent making within-school comparisons assuming that school-level SEMH policy does not change in treated schools relative to control schools in 2024. Section 4 sets out how the panel estimators handle school-level variation in determining SEMH.

2.4 Child Disability Benefits and the DLA System

Child Disability Living Allowance (DLA) is the main cash benefit for children under 16 with disabilities in England and Wales. It is tax-free, non-contributory, not means-tested, and worth about £5,700 a year on average (DWP, 2025a; Latimer et al., 2025). Awards are decided by a Department for Work and Pensions (DWP) decision-maker on a paper-based assessment of the child's care and supervision needs, guided by departmental medical guidance and usually without an examination (DWP, 2025b). The decision rests on the parent-completed claim form and supporting evidence. For the psychiatric grouping the relevant evidence is clinical and developmental. The only educational document the DWP lists is a local-authority document, either a statement of SEN or an EHC plan (DWP, 2025c).[12] In practice the benefit usually comes first. Among children who receive both child DLA and an EHCP or statement of SEN, more than 60 percent start DLA before the EHCP and 80 percent start it no later, so the EHCP is not typically the gateway document for receiving DLA. Both forms of support are quite persistent. For example, of children starting an EHCP (child DLA) at age 5, 84 percent (about 75 percent) still hold it at age 15.[13] The rules are identical across England and Wales, so eligibility does not vary across local areas in the way SEN and EHC practice does.

Caseloads have grown steeply in the last 15 years. The share of under-16s receiving the benefit rose from 2.8 to 7.2 percent between 2010 and 2025, with recent growth concentrated in autism, ADHD, and learning difficulties (Latimer et al., 2025). My DWP outcome focuses on claims with a psychiatric disabling condition. Section 3 sets out the construction and composition of the outcome.

3 Data

3.1 Treatment: Parent Pact Uptake

I measure treatment from the Smartphone Free Childhood Parent Pact registration data. For each registered school, SFC publishes the cumulative number of families who have signed the pact and the school's signup share (signed families relative to enrollment), through a public

[12] A statement of SEN was the legal predecessor to the EHC plan, issued by the local authority after a statutory assessment, before being phased out in favor of EHC plans under the 2014 SEND reforms. Both are local-authority statutory documents, distinct from the school-decided, non-statutory SEN-support tier described in Section 2.3.

[13] Longitudinal Education Outcomes (LEO) timeline summaries calculated by Darcey Snape (Institute for Fiscal Studies) in the ONS Secure Research Service and shared with the author (see also Latimer et al. 2025). Start dates are measured at January of each year. "EHCP" here covers EHC plans and the earlier statements of SEN. Counts below ten are suppressed. The start-order statistics pool the 1998–2008 birth cohorts, the persistence statistics the 1998–2002 cohorts. Two qualifications apply. An EHCP marks the formal statutory tier of SEN provision, reached through a lengthier process for more complex needs, and most children who eventually receive one are registered for school SEN support well beforehand, so these timelines compare DLA with the statutory document rather than with SEN identification generally. The cohorts also largely predate the steep recent rise in SEN and child-DLA rates.

interface keyed to the DfE Unique Reference Number.[14]

The number of SFC registrations likely understates the number of parents actually delaying ownership. Signing requires a parent to find the online tool and register personal details, so delaying families plausibly outnumber families who registered. Cunningham Hill Junior School illustrates this difference. Its surveyed Year 6 smartphone ownership fell from 75 to 12 percent in the 2024 cohort, and to about 3 percent in the next. The school's registered parent-pact proportion reached only about 38 percent of enrollment by June 2026.[15] Signup counts therefore understate how far behavior actually shifted.

I classify a primary school as *treated* if its June 2026 signup share is at least 20 percent and as *low-uptake* if its share is below 5 percent. The 20 percent cut is an ex ante choice anchored on the flagship school. Registered uptake at Cunningham Hill stood near 40 percent and reported ownership almost entirely collapsed. Hence, the tipping point for community-wide compliance lies somewhere below 40 percent, and I set the threshold at half of that. The results are robust to using other thresholds.[16] Among the roughly 16,700 state primary schools, 372 reach the 20 percent threshold. The 329 with the complete census series required by the balanced panel (next subsection) form the treated group, spread across 118 local authority districts with at most 14 in any one. Approximately 12,000 schools have negligible uptake and constitute the low-uptake pool from which matched controls are drawn.[17] Most registered schools have low intensity, with a median share of 2 percent among schools with any pacts, and the high-uptake group is moderately concentrated, with the ten most active districts holding three in ten treated schools. Appendix Table 10 gives the full distribution,

[14] SFC Parent Pact, https://central.smartphonefreechildhood.org. I retrieved a national snapshot in June 2026 covering 15,653 registered schools across all 61 SFC regions.

[15] The fall to 12 percent is the 2024 Year 6 survey reported in Gentleman (2025); see Section 2. For the following Year 6, the school's headteacher reports that ownership dropped to 3 percent (interview with the author, July 2026). The 38 percent figure is the parent-pact proportion published by SFC, pacts per enrolled child in June 2026.

[16] The effect is already present at the 10 percent threshold and rises with uptake (Table 3). Appendix Table 18 varies the control-pool cutoff.

[17] To the best of my knowledge, there is no comparable competing delay initiative. The only other national pledge, Delay Smartphones, gathered roughly 5,000 signatures nationally against the Parent Pact's 165,000, so the low-uptake control pool is unlikely to conceal unmeasured coordinated delay activities.

and Appendix A tabulates the sample construction from the 15,653 registered schools to the analysis samples. Uptake varies widely even within the pledge towns, from around 40 percent of families at the schools with most engagement to near zero at others (Appendix Tables 14 and 15). The design exploits this school-level variation and dropping the three media-prominent towns leaves the estimates essentially unchanged (Appendix Table 18). Schools between 5 and 20 percent are held out of the binary comparison and enter only the dose-response specifications, which estimate the effect at uptake thresholds of 10, 20, 30, and 40 percent.

Table 1 compares the two groups on pre-treatment observables. High-uptake schools are markedly more affluent (free school meals 10.5 versus 26.0 percent, IMD-2019 deprivation scores 10.8 versus 23.2), they begin with somewhat lower SEMH (25.0 versus 30.1 per 1,000) and markedly lower child DLA claiming (32.1 versus 53.2 per 1,000), while mean enrollment is balanced. Our different estimators account for the differences between these schools (Section 4).

[Table 1 about here.]

3.2 School Outcome: SEMH Needs

The DfE School Census, conducted each January, records for every state-funded school in England the number of pupils identified with each special-educational-need category. My school outcome is the SEMH rate: pupils recorded with social, emotional and mental health needs, combining the SEN-support and EHC-plan tiers (Section 2.3), per 1,000 enrolled.

I assemble a panel of approximately 16,700 primary schools and estimate on the balanced subpanel of 14,260 schools with a complete SEMH series across the January censuses from 2019 to 2026, so that group composition is fixed across the pre- and post-treatment windows. Of the 372 high-uptake schools that define treatment (Section 3.1), 329 have a complete series and form the matched estimation sample (Appendix A tabulates the sample selection). I begin in 2019 because the DfE suppressed small SEMH counts (below five) in the earlier censuses but stopped from January 2019. About a quarter of schools have a non-randomly suppressed SEMH value in the 2015–2018 censuses, whereas 2019 onward is complete.

SEMH accounts for about a fifth of identified primary-school need (Appendix Table 11 sets out the categories). The autistic-spectrum category, recorded by the same school staff in the same census yet not plausibly affected by a smartphone pledge, later serves as the primary placebo (Section 6). Figure 2, Panel A, plots the two national series.

3.3 DWP Outcome and Placebos

One concern is that identification of SEMH is a school decision and schools with engaged parents may record need differently. I therefore add child DLA claims as a second outcome, including independent placebo outcomes that a phone delay should not affect. The DLA decision-maker is a DWP caseworker outside the school, applying eligibility rules that are common across England and Wales (Section 2.4). A school cannot reclassify a DWP benefit, so a similar fall in DLA claims corroborates the school result through an independent decision-maker.

The DWP publishes counts of children receiving DLA with a recorded psychiatric main condition, split by disability component, age band, and sex, for each Lower-layer Super Output Area (LSOA), a neighborhood unit of roughly 1,500 residents, at a quarterly frequency. Each count is the stock of current claimants at the snapshot date, not the flow of new benefits. At ages 5–15 the psychiatric caseload is about 60 percent "learning difficulties" (a developmental classification including autism), just under 30 percent behavioral disorder, about 10 percent hyperkinetic/ADHD, and effectively zero psychoneurosis (the anxiety-and-depression category) and psychosis.[18] The ideal outcome would be behavioral claims among 5–10 year-olds. This outcome would target the age range and the emotional and behavioral conditions affected by a phone delay.

[18] Composition at the May 2024 snapshot, from the DWP Stat-Xplore disability-category breakdown; the published categories are independently rounded and small-cell suppressed, so their shares do not sum exactly to the total caseload.

The DWP releases these counts through its Stat-Xplore portal with a short publication lag, under two privacy protections: any cell with fewer than five claims is replaced with a zero, and every published figure is perturbed. Small cells will therefore be mostly filled with zeros. I make three choices to limit this suppression. First, I aggregate neighborhoods into school catchments, assigning each LSOA to its geographically nearest state primary school (about two LSOAs per catchment). Pooling that data to catchments lifts most child psychiatric counts above the suppression floor. I retrieve each catchment count from Stat-Xplore as a single custom-geography total, so it is perturbed once rather than assembled from many separately perturbed cells (Appendix A). The catchments are anchored on all schools whether or not they appear in the census panel, so the DWP and school outcomes rest on slightly different treated sets.[19] Second, I average the four quarterly snapshots within each calendar year. The stock of claimants changes little from one quarter to the next, so averaging removes noise and aligns the series with the annual school census. Third, I use the 5–15 psychiatric aggregate as the main outcome, and split it one dimension at a time: the behavioral component (psychoneurosis, behavioral disorder, and hyperkinetic/ADHD) against the developmental component (learning difficulties and psychosis),[20] ages 5–10 against 11–15, and boys against girls. When calculating percent changes, the estimated effects on the main outcome will be attenuated because the outcome includes older children and conditions unaffected by a phone delay.

After this aggregation, zero cells are close to balanced between treated catchments and their matched controls, 8.9 versus 7.7 percent of catchment-quarters for the 5–15 psychiatric total.[21] A cross-check against the near-unsuppressed MSOA marginals indicates most recorded zeros are close to an actual zero.[22] Imputing every zero with four claimants does

[19]The nearest-school catchment is an approximation: families often attend a non-nearest school, which dilutes the treatment contrast. In addition, any pact-engaged families at non-nearest schools place treated families in control catchments. Both measurement issues should attenuate my estimates.

[20]Appendix A.6 maps the correspondence with the school SEN categories.

[21]See Appendix Table 12 for the suppression rate of each cell. The ideal 5–10 behavioral cell is suppressed for over 70 percent of treated catchment-quarter observations, which is why it is not used for estimation.

[22]The MSOA marginals are a clean benchmark because their cells are large enough that only 0.1 percent are recorded as zero. Summed across LSOAs in their native (2021) geography, the national caseload with

not meaningfully change the results (Appendix Table 18).

Alongside the autistic-spectrum category recorded in the school census (Section 2.3), I use two placebo outcomes that are decided outside the school. The first is the developmental component of the DLA caseload (learning difficulties, which subsumes autism, together with psychosis). The DWP adjudicates it under the same rules as the behavioral component, but it describes structural conditions unaffected by a phone delay. The second is adult depression recorded in GP registers for adults aged 18 and over, from the NHS Quality and Outcomes Framework (QOF, Appendix A). A primary-school pledge cannot affect adult depression. A fall in the developmental component would suggest the effect is not specific to behavioral conditions, and a fall in adult depression would point to an area-level confound rather than an effect on children.

Panel B of Figure 2 plots the national evolution of the seven DLA outcomes described above, alongside the school series in Panel A. All of these series rise steeply over this period.

[Figure 2 about here.]

4 Empirical Strategy

4.1 Design and Estimator

I date treatment to February 2024, the month the Smartphone Free Childhood movement began (Section 2). Observations are labeled by their measurement date (see Appendix A), so the January 2024 census is the last pre-treatment school observation and the January 2025 and 2026 censuses are the two post-treatment ones. The quarterly DLA series, averaged to calendar years for estimation (Section 3), contributes eight post-treatment quarters, February

zero cells left at zero is 264,700 in May 2018 and 665,700 in November 2025, against an MSOA benchmark of 262,600 and 663,900; the zero-fill matches the benchmark to within 1 percent because the few genuinely suppressed cells carry little mass and are further offset by DWP's cell-key perturbation of the published marginals. Filling every zero cell at four instead raises the May 2018 total to 293,300, about 12 percent above the benchmark. Figure 2 plots the England-wide catchment series, whose national sum matches this benchmark to within 1 percent; the matched estimation uses the catchment panel built on native 2021 LSOA geography, with no crosswalk.

2024 through November 2025, forming the 2024 and 2025 post-treatment years.

I compare high-uptake schools (Parent Pact signup share at least 20 percent) to low-uptake schools (below 5 percent), as defined in Section 3.1. High- and low-uptake schools differ in level and were converging before the pact, so my primary estimator is a matched difference-in-differences (Dustmann et al., 2017). I match each treated unit to three low-uptake controls by nearest-neighbour matching on the Mahalanobis distance of the covariates, without replacement. The covariates are the same for every outcome: the outcome's own pre-treatment levels (in each of 2021–2023), neighborhood affluence (the 2019 Index of Multiple Deprivation, IMD-2019), the free-school-meal share, and log size.[23] On the matched sample I estimate the event study

$$Y_{it} = \alpha_i + \beta_t + \sum_{k \neq -1} \gamma_k \, \mathbf{1}[\text{treated}_i] \cdot \mathbf{1}[t = T_0 + k] + \varepsilon_{it}, \tag{1}$$

reading the post-treatment coefficients γ_k, relative to the last pre-treatment period, as the effect. Effects are reported by period and pooled across the post-treatment window. I express each effect as a percentage of the counterfactual level implied by the estimated fixed effects. This base is the level the treated units would have reached had they followed the control time path (Section 5). Adding local-authority-by-year fixed effects to (1) identifies the effect from high- versus low-uptake schools within the same district and year.

I report estimates from two alternative estimators that have different identification assumptions. I estimate synthetic difference-in-differences (Arkhangelsky et al., 2021), which reweights the low-uptake donor pool so that the weighted controls track the treated schools' pre-treatment trajectory. Its donor pool is the covariate-only match of the triple difference

[23] For a school, these are the IMD-2019 where the school is located and size measured by enrollment. For a catchment, IMD-2019 and the FSM share are the anchor school's values, and size is the catchment's resident child population. IMD-2019 is England's official small-area deprivation index, a weighted combination of seven domains: income (22.5 percent of the index), employment (22.5), education, skills and training (13.5), health and disability (13.5), crime (9.3), barriers to housing and services (9.3), and living environment (9.3) (McLennan et al., 2019). A lower score means a more affluent neighborhood. IMD-2019 is measured well before treatment: its income data are from 2015/16 and its health-domain benefit counts from March 2016. Each school's neighborhood is the 2011 LSOA containing its coordinates, the same coordinate rule that assigns neighborhoods to catchments (Appendix A).

below, so no outcome levels enter its match and the weights alone fit the pre-treatment trajectory. The estimator, its weights, and the numeric estimates are set out in Appendix B. I also estimate a triple difference that does not require matching on mental-health outcomes. It compares each mental health outcome to a within-unit placebo outcome, SEMH against the autistic-spectrum category in schools and the behavioral against the developmental component in catchments. Unit-by-year fixed effects absorb shocks common to both categories (Section 6.1).

Treatment is spatially clustered, so I cluster standard errors by local authority district, of which 118 contain a treated school. As a distribution-free robustness check I also report a placebo-in-space permutation test (Section 6).

4.2 Identification and Falsification

The threat to a causal reading is selection. High-uptake communities could differ from low-uptake communities on unobserved trajectories, not merely on levels. I provide four kinds of evidence to support the parallel-trends assumption.

Pre-trends. Matching on levels does not by itself rule out selection on trends. Two problems could remain. Treated communities could already be on a different trajectory, which would continue after treatment and look like an effect. And matching on the years just before treatment can select units at a temporarily high or low level, which then reverts. This is the regression-to-the-mean bias of Daw & Hatfield (2018). Four checks address these problems. First, the matched series track each other closely through the whole pre-period, including 2019 and 2020, which the match never uses (Section 5). Controls selected on a transiently high or low 2021–2023 level would sit apart from the treated group in those years and converge into the matching window, and no such gap appears. Second, re-dating treatment to a year inside the pre-period returns a null, the direct test of a continuing trend (Table 7). Third, the placebo-in-space distribution bounds the bias that matching could produce under the null: the placebo mean is less than a twentieth of the real estimate, with the opposite sign (Section 6.3). Fourth, the triple difference drops the matching on pre-treatment outcomes altogether. It compares the outcome with a placebo category inside the same school or catchment, and unit-by-year fixed effects absorb any trend common to the two categories (Section 6.1). The pre-trends for the triple-difference estimate are flat without any matching on mental-health outcomes and estimates are similar to the main difference-in-differences estimates.

Within-district variation. Adding local-authority-by-year fixed effects to (1) identifies the effect from high- versus low-uptake schools in the same district and year. This absorbs every district-level time-varying confound, the most likely channel for area-level selection. The effect survives this restriction. The same comparison also tests for spillovers. If the delay norm spread from high-uptake schools to nearby low-uptake schools, the within-district comparison group would be partly treated and the estimated effect would shrink. I do not find any such attenuation (Section 5).

School-level selection and recording. District-by-year effects do not account for differences in school policy between schools. This matters because recorded SEMH is a judgment made by the school itself (Section 2.3). The remaining worry is therefore a school-specific change, in parental engagement or in recording practice, that coincides with the pledge. School fixed effects absorb each school's own recording threshold. Measurement error that is unrelated to treatment only widens the confidence intervals. The remaining threat is recording error that is correlated with treatment. First, take-up is decided by parents, not by the school. Within a single pledge town the signup share ranges from near zero to about forty percent across schools, so uptake does not measure school policy (Section 3.1). Second, if a school changed how it records need, the change would affect the developmental categories as well as SEMH, and the developmental categories do not change (Section 6). Third, the DWP outcome is decided outside the school, and it falls together with the school result (Section 5).

Placebo outcomes and the age signature. A genuine smartphone effect makes sharp predictions about where it should appear and where it should not. It should reduce SEMH and child psychiatric claims. It should not affect conditions unrelated to a phone, and I test this by re-estimating the model on the placebo outcomes of Section 3: the autistic-spectrum school category, the developmental DLA component, and recorded adult depression (QOF, 18+) in the same areas. It should also be larger among younger children, whose smartphone acquisition is still being postponed, than among older children who mostly already own a device (Section 5.2).

5 Results

My primary estimator is the matched two-way fixed-effects event study of Section 4: high-uptake schools (at least 20 percent Parent Pact uptake) against low-uptake controls (below 5 percent), matched on the outcome's own 2021–2023 levels, the IMD-2019, the free-school-meal share, and log size. Years before 2021 are not used by the match and serve as a pre-trend check.

5.1 Effects on Mental Health

The school and DWP outcomes are complementary. The school census covers every school without suppression and spans several categories of need, but it cannot distinguish children's ages and is recorded by the school itself. The DLA benefit is determined outside the school and observed by age and sex, at the cost of small-count suppression at fine geography.

School outcomes. The social, emotional and mental health (SEMH) need of children responds to pact uptake. Figure 3 (top row) plots SEMH, showing the weighted group means on the left and the differenced event-study coefficients on the right. For SEMH, treated and matched-control schools track each other closely through the pre-period, both roughly doubling between the January 2019 and January 2024 censuses as national SEMH identification rose, and the treated series falls below the control after the movement began in February 2024. The differenced coefficients are flat and near zero in the pre-period, and a significant difference opens after the start of the movement. The average post-treatment SEMH reduction is 2.9 per 1,000 pupils, about 8 percent of the counterfactual, defined as the level the treated schools would have reached on the control time path (Table 2).The drop in SEMH grows over time, from 2.4 per 1,000 in the first post census (2025) to 3.4 in the second (2026), consistent with uptake growing over time.[24]

Adding local-authority-by-year fixed effects, which absorb district-level trends, leaves the pre-trend flat and the reduction at least as large (Table 2, column 2). The comparison therefore holds within districts, between high- and low-uptake schools in the same local authority and year, and is not the result of comparing affluent areas with poorer ones.

[Figure 3 about here.]

[Table 2 about here.]

DWP outcome. The second outcome in Figure 3, child psychiatric Disability Living Allowance (DLA) claims, is decided outside the school. The SEMH measure is a school classification, recorded by the same staff whose engagement may facilitate the pact's uptake. To test whether the reduction reflects children's wellbeing rather than school recording practice, I estimate effects on the DLA benefit. A DWP caseworker decides each claim under rules that are common across England and Wales. I assign every neighborhood (LSOA) the Parent Pact uptake of its nearest primary school and run the same matched difference-in-differences, comparing high-uptake ($\geq$20 percent) with matched low-uptake (<5 percent) neighborhoods (Figure 3, bottom row; Table 2, columns 5–6).

In high-uptake areas, child psychiatric claims fall by about 3.0 per 1,000 children, around 6 percent of the counterfactual, on a flat pre-trend. The reduction grows from about 4 percent

[24] Note that my measure of uptake is a cumulative June 2026 snapshot that dates partially-treated schools as treated from the outset, so the pooled estimate can be considered a lower bound.

in the first post-treatment year to about 8 percent in the second. A school cannot reclassify a DWP benefit, so this corroborates the school result through an independent measurement system at a similar magnitude.

One might worry that the school and DWP results are linked through the EHC plan, the single school-related document that the DWP includes in its decision (Section 2.4). EHC plans are issued by the local authority and account for about 12 percent of the primary-school SEMH count.[25] Re-estimating the headline SEMH event study on the school-decided SEN-support tier alone (dropping EHC) gives a same-signed reduction,[26] so the similarity of the school and DWP results does not depend on the shared EHC document.

5.2 Dose-Response and Heterogeneity

The splits in this section probe the mechanisms driving my results. The dose-response asks whether the effect scales with local adoption. The condition split asks whether the reduction sits in the diagnoses that a phone channel could plausibly move. The age split asks whether the effect sits with the children whose ownership is actually being delayed. The sex split asks whether the more affected children are the ones the social-media literature points to. The affluence splits describe who the reductions reach within the adopting communities.

Pact uptake. The effect strengthens with pact uptake (Table 3). The matched SEMH reduction is about 3 percent of the counterfactual where a tenth of families signed and 8 percent at the one-fifth threshold. The DWP outcome shows the same pattern, about 4 percent at one in ten and 6 percent at one in five, the headline threshold (Table 3, Panel B). The estimates at the two highest thresholds (30 and 40 percent) are consistent with the growth but are noisy, as few schools reach these uptake levels. An effect that scales with local adoption is consistent with a community-level treatment.

[25] January 2024 census, state-funded primaries: 127,260 pupils on SEN support and 18,060 with an EHC plan under the SEMH primary need, so the EHC tier is 12.4 percent of the SEMH count.

[26] A reduction of 4.9 percent of the counterfactual against 8.2 for the full measure, and 4.7 percent within districts, on flat pre-trends.

[Table 3 about here.]

Behavioral conditions. The child DLA caseload divides into a behavioral component, the conditions that a phone channel could plausibly affect, and a developmental remainder that serves as a placebo (Section 3; Table 6). A genuine smartphone effect should therefore concentrate in the behavioral component, and it does. The behavioral reduction is about 10 percent of its counterfactual and is robust to district-by-year fixed effects (Table 4). The component also carries the heaviest small-cell suppression, with about 45 percent of treated catchment-quarters recorded as zero, so it is measured less precisely than the psychiatric total (Appendix Table 12).

Age gradient of the DLA effect. An effect of delaying ownership works through children whose ownership is actually being delayed. Smartphone ownership rises from about four in ten children at ages 8–9 to nine in ten at age 11 (Section 2). At ages 11–15 the pact therefore binds only at the band's lower edge, the children who reach age 11 without a phone. In these first post-treatment years the effect should be much smaller there than at ages 5–10. It should grow as pledged cohorts age into the band. Table 4 estimates the matched difference-in-differences separately by age band, each band re-matched on its own pre-treatment levels. Child psychiatric claims fall by about 9 percent of the counterfactual at ages 5–10 (Figure 4). At ages 11–15, where ownership is already near-universal, the estimate is half the size or less (-3.9 percent). Adults provide a further check. Recorded depression among registered adults is essentially unchanged (the adult row of Table 4). An area-wide confound, such as an engaged community expanding local services, would also affect adults. The unchanged adult series is inconsistent with this channel.

[Figure 4 about here.]

Sex. The social-media literature finds larger harms among girls, and the Norwegian school-ban evidence finds mental-health gains concentrated among girls (Abrahamsson, 2026). A genuine effect should therefore be larger for girls. The reduction is larger for girls, about 8 percent of their counterfactual. For boys it is about half the size and noisier, with an interval that spans zero (Table 4).

Affluence. Affluence splits are reported in Appendix Table 17. The SEMH reduction is larger in the less affluent treated schools (−14 against −2 percent), while the DLA reduction is larger in the more affluent catchments (−8 against −4 percent), and the smaller cells carry wide intervals. The opposite gradients are puzzling, and I do not have an explanation for them. Section 7 discusses the heterogeneity of the effects, including how they depend on local affluence.

[Table 4 about here.]

6 Robustness and Falsification

The tables in this section report every estimate with its 95 percent confidence interval, so each check can be read directly against zero.

6.1 Alternative Estimators

The headline estimates are robust to estimators that take different approaches to matching treated and control units (Table 5). All three designs match treated to control units on the same covariates: the IMD-2019, the free-school-meal share, and log size. The baseline matched TWFE event study additionally matches on the three most recent years of pre-treatment mental-health outcomes.

The two alternative estimators avoid matching on outcome levels, and each handles differential pre-treatment trends in its own way. Synthetic difference-in-differences (Arkhangelsky et al., 2021) still uses the pre-treatment mental-health outcomes, through reweighting rather than matching: it reweights the covariate-matched controls so that the weighted control series tracks the treated series through the pre-period. Its identifying assumption is that this reweighting removes the unobserved differential trends, so that absent the pact the weighted controls would have tracked the treated series. The triple difference is the only estimator that uses no pre-treatment mental-health outcomes. It compares each mental-health outcome with a developmental outcome in the same unit: SEMH against the autistic-spectrum category in schools, and the behavioral against the developmental component in catchments. Unit-by-year fixed effects absorb any shock or trend common to a unit's two categories, so identification requires no matching on outcomes.

I compare the estimators in claims or pupils per 1,000 rather than percent, because their percent bases differ. The level designs measure the effect against the full psychiatric counterfactual, while the triple difference measures the behavioral–developmental gap against the smaller behavioral base, so the percent figures are not comparable across estimators. In claims or pupils per 1,000 the SEMH reduction is -2.9 under the matched TWFE, -1.7 under SDID, and -2.0 under the triple difference; the DLA reduction is -3.0, -2.4, and -1.7. The DLA triple difference is not directly comparable to the level estimate on the psychiatric total, which aggregates behavioral and developmental claims. If the entire effect runs through behavioral conditions, the two should imply a similar reduction in claims. All three estimators agree in sign, and the triple difference identifies the effect without using pre-treatment outcomes at all. Figure 5 plots the event studies under all three.

[Table 5 about here.]

[Figure 5 about here.]

6.2 Placebo Tests

I subject the result to placebo tests in outcomes and in time. A third, spatial placebo underlies my inference and is set out separately in Section 6.3.

Placebo outcomes. Table 6 compares the effects for the two mental-health outcomes, school SEMH and the behavioral DLA component, against the two within-system placebo outcomes: the autistic-spectrum school category and the developmental DLA component. The DLA rows partition the caseload into disjoint components, so their sampling errors are independent, unlike a total-versus-component comparison. Both placebos are small, with intervals that include zero, under the matched and the district-by-year specifications, while the moving outcomes fall by 8 to 10 percent. Autism is the natural school placebo, identified by the same gatekeepers in the same census as SEMH, and its DLA analogue is in the developmental component (Section 3). A school could in principle relabel pupils between SEMH and autism, which would depress SEMH mechanically. As the DWP-adjudicated behavioral component also falls together with SEMH, the triple difference results cannot be simply explained by school relabeling.

[Table 6 about here.]

Placebo in time. Because both the baseline design and SDID matches on pre-treatment mental health outcomes, a placebo in time is a key falsification test. It asks whether treated communities were already diverging from their matched controls before the pact existed. Matching on pre-treatment outcomes aligns the two groups at the match date but does not make their trends parallel, and can even induce a spurious trend through regression to the mean (Daw & Hatfield, 2018).

I re-date the treatment to February 2022 and re-estimate on pre-treatment data alone. The placebo uses only years before the fake date for matching, and the placebo effect is measured on the last two pre-treatment years (2022 and 2023). For school SEMH the placebo is close to zero, +0.1 per 1,000 under the matched TWFE and +0.2 under SDID, against the genuine −2.9 and −1.7 at the true date (Table 7). Neither design manufactures an effect where there was no treatment. For the DWP outcome the fake-date estimate is marginally negative, −1.2 per 1,000 under the matched TWFE and −0.9 under SDID, against reductions of −3.0 and −2.4 at the true date. While we cannot reject the null in either case, these may reflect a small differential trend. As both the matched TWFE and SDID use the same three pre-treatment outcomes when treatment is assigned to 2022, it is perhaps not surprising that their placebo in time estimates are similar as well.

The triple difference does not use pre-treatment outcomes at all. It compares the behavioral and developmental components within the same catchment, so any differential trend common to the two components is differenced out. At the true date it is −1.7 per 1,000, a 9.4 percent drop in behavioral claims, and its fake-date placebo is +0.4, centered on zero. The school triple difference against autism shows the same, −2.0 per 1,000 at the true date against +0.2 at the fake one. The DLA result is therefore not driven by the differential trend that the level design placebos hint at.

[Table 7 about here.]

6.3 Inference: A Calibrated Placebo-in-Space Permutation

Because treated units cluster within districts, analytic inference alone may be optimistic about the effective number of independent clusters. I therefore complement the district-clustered confidence intervals with a design-based placebo-in-space exercise, in the spirit of the few-treated-clusters literature (Conley & Taber, 2011; Ferman & Pinto, 2019; MacKinnon & Webb, 2020). Each replication draws a placebo "treated" set of never-treated (<5 percent) schools that reproduces the real treated schools' count and per-district distribution, rematches it to the remaining low-uptake units on the same covariates, and recomputes the matched difference-in-differences. Because each placebo set is matched and estimated exactly as the real sample, the exercise also bounds regression to the mean: any bias that matching on pre-treatment levels produces under the null appears in the placebo distribution, so a real estimate in its tail exceeds that bias. For the DWP outcome the drawn schools are the placebo catchments themselves, since the DLA panel is built on school catchments, so the school and DWP outcomes share one draw mechanism. The permutation p-value is the share of placebo estimates at least as extreme as the actual one,

$$p = \frac{1}{B+1} \sum_{b=0}^{B} \mathbf{1}[|\hat{\tau}_b| \geq |\hat{\tau}|] \,, \tag{2}$$

where $\hat{\tau}_0 \equiv \hat{\tau}$ is the actual estimate. Equation (2) defines the uncentered two-sided comparison. I also report the one-sided comparison and a two-sided comparison re-centered on the placebo mean, which replaces $|\hat{\tau}_b|$ and $|\hat{\tau}|$ with absolute deviations from the mean placebo estimate.

The draw is calibrated to reproduce the spatial clustering of the real treatment. The 329 high-uptake schools sit in 118 districts with a mean nearest-neighbour distance of 5.3 km, and each replication assigns the real per-district counts to a fresh sample of never-treated districts, matching both the number of treated districts and the within-district concentration (placebo nearest-neighbour distance 5.5 km). The DLA permutation applies the same school-level draw on the catchment panel, assigning the real per-district counts of the 330 treated catchments to fresh never-treated districts and drawing placebo catchments from the full national pool of <5 percent catchments, re-matched each draw.

Both headline estimates lie in the left tail of their placebo distributions (Figure 6). The DLA estimate is more negative than every one of the 200 placebo draws, and the three comparisons give $p \leq 0.005$ (panel B). For SEMH the one-sided and mean-centered comparisons give $p = 0.010$, and the uncentered two-sided comparison of equation (2) gives $p = 0.015$ (panel A). The placebo distributions sit just positive of zero (means +0.11 per 1,000 for SEMH and +0.60 for DLA), so the regression to the mean that the matching procedure could produce under the null points toward positive estimates, opposite to the reductions I find, and cannot account for them.

[Figure 6 about here.]

6.4 Specification Robustness

Both headline effects are stable across the low-uptake threshold (<5 versus <10 percent), the match ratio (1:3 versus 1:5), and matching with or without replacement (Appendix Table 18). Both headline reductions stay near 3 per 1,000 across the rows. Extending the SEMH pre-period back to 2015 restricts the panel to the schools with complete early data and more than doubles the standard error, leaving the point estimate in the same range but the interval spanning zero. Imputing every catchment-quarter recorded as zero (a true count of 0 to 4 under the disclosure rules) at its upper bound of 4 changes the headline estimate from -3.39 to -3.15 per 1,000. The headline catchment cell is rarely at the suppression floor and its zeros are near-balanced between treated and control catchments, so the bound barely binds. Estimating the DLA event study on the raw quarterly claim snapshots, rather than the calendar-year means used throughout, gives the same pooled reduction of -3.4 per 1,000, with weak evidence of anticipation in the quarters just before February 2024 (Figure 7).[27] Dropping the three media-prominent pledge towns from the treated group leaves every estimate essentially unchanged. The effects are a feature of high-uptake communities nationally, not of a handful of famous towns.

7 Discussion

Taken together, the evidence is consistent with a causal effect of coordinated ownership delay on children's mental health. Two independent measurement systems record falls of similar size (8.2 and 6.0 percent of the contemporaneous counterfactual), each growing across the post-treatment window. The effect strengthens with uptake. It survives the absorption of all district-level trends. It is absent in the placebo outcomes, the other special-educational-need categories in the same children and recorded depression among adults in the same communities.

[27] Any anticipation would be consistent with the search evidence of Section 2.2, where searches for a "safe phone" for children emerge over 2023, before the movement's February 2024 launch (Figure 1).

The size of the estimates raises a fair question: why should delaying one device affect recorded mental-health need by several percent within two years? Part of the answer is mechanical. Both outcomes are stocks that turn over slowly. Support, once begun, persists for years (Section 2.4), so the caseload changes almost entirely through new identifications and newly granted benefits. A back-of-the-envelope calculation combining the LEO persistence rates with the observed national caseload puts new behavioral benefits at roughly 30 percent of the behavioral caseload per year. The estimated 10 percent fall in the behavioral stock within two years then implies a roughly 20 percent drop in *new* benefits below the counterfactual.[28] The delay likely only affects the flow margin, which is also consistent with the estimates growing across the post-treatment window.

A second part of the answer is dilution. The main outcome includes older children and developmental conditions unaffected by a phone delay (Section 3). Converting the estimates into claim counts over the 330 treated catchments, the total reduction is about 480 claims, of which the 5–10 age band accounts for about 330 and the behavioral component for about 320.[29] The groups a phone delay should affect therefore carry most of the count reduction. Measured against their own counterfactuals, claims fall by about 9 percent among 5–10 year-olds and 10 percent in the behavioral component, compared with the 6 percent headline. The two cuts can also be combined. If the entire 5–10 reduction sits in behavioral conditions, as the flat developmental placebo indicates, my estimates imply about a 20 percent drop in DLA behavioral benefits for 5–10 year-olds.[30] The stock-to-flow conversion above applies to

[28]The LEO timelines of Section 2.4 (birth cohorts 1998–2002) imply that 98 percent of holders continue each year. Holders aged 15 age out of the 5–15 band annually, about a tenth of the caseload, since ages 11–15 hold half of it. Few children turn 5 already receiving a behavioral benefit, because attention-deficit and conduct conditions are rarely recognized before school age. Applying these rates to the observed national behavioral caseload, which rose from about 181,000 in 2023 to 289,000 in 2025, implies roughly 71,000 new benefits in 2024 and 87,000 in 2025, about 30 percent of the caseload per year. Two years of new benefits are then about half the 2025 stock, so a given change in the stock implies a change in new benefits about twice as large.

[29]Implied counts multiply each per-1,000 estimate by the 2023 child population of the 330 treated catchments (157,851 children aged 5–15, of whom 81,767 are aged 5–10). Each split is matched on its own outcome, so the split counts need not sum exactly to the total.

[30]The published counts do not cross the condition split with age, so the 5–10 behavioral counterfactual must be approximated. Because the recent claims growth concentrates in behavioral conditions, I take the composition from the matched control catchments after treatment. Behavioral conditions are about 42

this cut as well. A 20 percent fall in the stock implies a roughly 40 percent drop in newly granted behavioral benefits for 5–10 year-olds. The actual effect is therefore likely much larger than the 6 percent headline suggests.

The psychological literature gives the effect concrete channels (Haidt, 2024): smartphones displace sleep, fragment attention through continual interruption, train compulsive checking through algorithm design, and crowd out the independent play in which children learn to regulate themselves. The causal economics evidence points the same way for the effect of social media on adults (Braghieri et al., 2022; Allcott et al., 2020). These channels map onto exactly the margins that fall in this paper. The SEMH category explicitly includes attention-deficit disorders alongside anxiety and attachment difficulties, and the DLA behavioral component is dominated by hyperkinetic and conduct conditions, the diagnoses closest to attention and self-regulation. These are exactly the behavioral changes reported by teachers both when smartphones were initially given to children historically and when ownership is delayed today (Haidt, 2024). For example, Justine Elbourne-Cload, the headteacher of Cunningham Hill Junior School (Section 2) reports that once children no longer received smartphones, “we have seen less evidence of dysregulation for those children and better attention spans”, and now counts “three, four children out of about four hundred that get very dysregulated, compared to a much higher number” before the change.[31] This paper provides the causal evidence for the changes that teachers observe on the ground and that psychologists have described.

The estimates cover the communities the pact has reached. Uptake concentrates in affluent areas. The treated catchments are markedly more affluent than England as a whole, and even the less affluent treated half only reaches about the national median. The estimates therefore do not speak to the most disadvantaged children. Within the range the pact covers, the exploratory affluence splits place the larger school effect in the less affluent treated

percent of their psychiatric claims over 2024–2025, up from 34 percent before treatment. Applying that share to the 5–10 counterfactual of about 3,800 claims gives roughly 1,600 claims, and the 330-claim reduction removes about a fifth of that.

[31] Interview with the author, July 2026.

schools and the larger DLA effect in the more affluent catchments (Appendix Table 17). The two cuts are measured on different margins and neither is precise enough to bound the other. Unfortunately, the analysis does not give a clear indication of how the effects vary by affluence.

The estimates bear on a live fiscal question. Latimer et al. (2025) document steep growth in spending on children's disability benefits and high-needs education. Real-terms child DLA spending doubled to over £4 billion by 2025, and the high-needs schools budget reached £12 billion. The reductions estimated here, about 8 percent of recorded SEMH need and 6 percent of child psychiatric claims in adopting communities, bear directly on both budgets if they extend beyond those communities. The design measures a fall in recorded need and paid claims. Reading that fall as resource savings requires that it reflect reduced underlying need. The two-gatekeeper evidence is consistent with that reading, but a full costing must confirm it.

8 Conclusion

This paper provides the first quasi-experimental evidence on a community-level effort to delay children's smartphone ownership. The Smartphone Free Childhood Parent Pact coordinates families within a school to withhold smartphones until age 14, and its take-up varies enormously across England's primary schools. I identify its effect from that variation. In schools where at least a fifth of families signed, recorded social, emotional and mental health need falls by about 8 percent of the counterfactual relative to matched low-uptake schools. Child psychiatric disability claims, adjudicated by the DWP outside the school, fall by about 6 percent in the same communities. The reductions strengthen with uptake and over time. They concentrate among the younger children whose acquisition is still being delayed and in the behavioral conditions that a phone channel would affect.

In-school phone bans, the policy evaluated most often, leave total exposure unchanged and show no average mental-health benefit. Coordinated ownership delay is a different policy. It changes the norm for when children first receive a smartphone, and it costs almost nothing to operate. The evidence here indicates that where communities adopt it intensively, children's mental-health need falls.

Ownership delay may be more effective when paired with complementary policies. In the UK, the rise in mental-health need and smartphone ownership also coincided with the closure of many outdoor clubs and public spaces under local-authority budget cuts. These were the places where children once developed self-regulation and social skills, which are often lacking in the children with the greatest need. Teachers and Parent Pact organizers alike stress that removing the phone is not sufficient on its own. As one school governor put it to a headteacher in this sample, 'That's great that you're taking away screens, but what are you putting in place?' Future research should examine whether complementary programs, such as restored outdoor play, make ownership delay more effective.

References

Abrahamsson, S. (2026). Smartphone Bans, Student Outcomes and Mental Health. *Journal of Human Resources*, forthcoming.

Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2020). The Welfare Effects of Social Media. *American Economic Review*, 110(3), 629–676.

Arkhangelsky, D., Athey, S., Hirshberg, D.A., Imbens, G.W., & Wager, S. (2021). Synthetic Difference-in-Differences. *American Economic Review*, 111(12), 4088–4118.

Beland, L.P. & Murphy, R. (2016). Ill Communication: Technology, Distraction & Student Performance. *Labour Economics*, 41, 61–76.

Beneito, P. & Vicente-Chirivella, O. (2022). Banning Mobile Phones in Schools: Evidence from Regional-Level Policies in Spain. *Applied Economic Analysis*, 30(90), 153–175.

Blanchflower, D.G., Bryson, A., & Xu, X. (2025). The Declining Mental Health of the Young and the Global Disappearance of the Unhappiness Hump Shape in Age. *PLOS ONE*, 20(8), e0327858.

Braghieri, L., Levy, R., & Makarin, A. (2022). Social Media and Mental Health. *American Economic Review*, 112(11), 3660–3693.

Children and Families Act 2014, c. 6, Part 3. London: The Stationery Office.

Churchill, B.F. & Johnson, K.R. (2026). Broadband Internet Access and Adolescent Mental Health in the U.S. NBER Working Paper 34614.

Conley, T. & Taber, C. (2011). Inference with "Difference in Differences" with a Small Number of Policy Changes. *Review of Economics and Statistics*, 93(1), 113–125.

Daw, J.R. & Hatfield, L.A. (2018). Matching and Regression to the Mean in Difference-in-Differences Analysis. *Health Services Research*, 53(6), 4138–4156.

Department for Education & Department of Health (2015). *Special Educational Needs and Disability Code of Practice: 0 to 25 Years*. London: DfE/DoH.

Department for Work and Pensions (2025). *Disability Living Allowance (DLA) for Children*. GOV.UK. https://www.gov.uk/disability-living-allowance-children.

Department for Work and Pensions (2025). *Medical Guidance for DLA Decision Makers (Child Cases): Staff Guide*. GOV.UK. https://www.gov.uk/government/publications/medical-guidance-for-dla-decision-makers-child-cases-staff-guide.

Department for Work and Pensions (2025). *Disability Living Allowance for a Child Under 16: Information Booklet* (DLA1 Child Notes, edition 10/25). GOV.UK.

Donati, D., Durante, R., Sobbrio, F., & Zejcirovic, D. (2025). Lost in the Net? Broadband Internet and Youth Mental Health. *Journal of Health Economics*, 103, 103017.

Dustmann, C., Schönberg, U., & Stuhler, J. (2017). Labor Supply Shocks, Native Wages, and the Adjustment of Local Employment. *Quarterly Journal of Economics*, 132(1), 435–483.

Ferman, B. & Pinto, C. (2019). Inference in Differences-in-Differences with Few Treated Groups and Heteroskedasticity. *Review of Economics and Statistics*, 101(3), 452–467.

Figlio, D. & Özek, U. (2025). The Impact of Cellphone Bans in Schools on Student Outcomes: Evidence from Florida. NBER Working Paper 34388.

Girela-Serrano, B.M., Spiers, A.D.V., Ruotong, L., Gangadia, S., Toledano, M.B., & Di Simplicio, M. (2024). Impact of Mobile Phones and Wireless Devices Use on Children and Adolescents' Mental Health: A Systematic Review. *European Child & Adolescent Psychiatry*, 33(6), 1621–1651.

Golin, M. (2022). The Effect of Broadband Internet on the Gender Gap in Mental Health: Evidence from Germany. *Health Economics*, 31(S2), 6–21.

Goodyear, V.A., et al. (2025). School Phone Policies and Their Association with Mental Wellbeing, Phone Use, and Social Media Use (SMART Schools): A Cross-sectional Observational Study. *The Lancet Regional Health – Europe*, 51, 101211.

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

Hutchinson, J. (2021). Trends in SEN Identification: Contexts, Causes and Consequences. *Journal of Research in Special Educational Needs*, 21(1), 19–38.

Hutchinson, J. (2025). *Identifying SEND*. London: Education Policy Institute (funded by the Nuffield Foundation).

Jain, R. & Stemper, S. (2026). Growing Up Online: Global Evidence on (3G) Mobile Internet and Adolescent Learning. Working paper, March 2026 (first version 2023; previously circulated as "3G Internet and Human Capital Development").

Kessel, D., Hardardottir, H.L., & Tyrefors, B. (2020). The Impact of Banning Mobile Phones in Swedish Secondary Schools. *Economics of Education Review*, 77, 102009.

Latimer, E., Sibieta, L., & Snape, D. (2025). Support for Children with Disabilities and Special Educational Needs. Chapter 5 in *The IFS Green Budget 2025*. London: Institute for Fiscal Studies.

MacKinnon, J.G. & Webb, M.D. (2020). Randomization Inference for Difference-in-Differences with Few Treated Clusters. *Journal of Econometrics*, 218(2), 435–450.

McLennan, D., Noble, S., Noble, M., Plunkett, E., Wright, G., & Gutacker, N. (2019). *The English Indices of Deprivation 2019: Technical Report*. London: Ministry of Housing, Communities and Local Government.

Ofcom (2025). Children and Parents: Media Use and Attitudes Report 2025, with the Children's Media Literacy Tracker 2024 data tables and respondent-level data (parents of 3–17s, fieldwork 1 October to 25 November 2024). London: Office of Communications.

Orben, A., et al. (2025). *Feasibility Study of Methods and Data to Understand the Impact of Smartphones and Social Media on Children and Young People*. London: UK Department for Science, Innovation and Technology.

Sohn, S.Y., Rees, P., Wildridge, B., Kalk, N.J., & Carter, B. (2019). Prevalence of Problematic Smartphone Usage and Associated Mental Health Outcomes amongst Children and Young People: A Systematic Review, Meta-analysis and GRADE of the Evidence. *BMC Psychiatry*, 19, 356.

Table 1. Pre-treatment characteristics of high- and low-uptake primary schools.

	High uptake ($\geq$20%)	Low uptake (<5%)	Std. diff.
Number of schools	329	12,012	
SEMH per 1,000 (2021–23 mean)	25.0	30.1	-0.25
Free school meals (%)	10.5	26.0	-1.29
IMD-2019 deprivation score	10.8	23.2	-1.02
Mean enrollment	280.3	283.5	-0.02
Child psychiatric DLA per 1,000 (2021–23 mean)	32.1	53.2	-0.96
behavioral component	9.2	20.0	-0.88
developmental component	21.8	32.8	-0.75
Urban (%)	64.7	73.3	-0.19

Notes: Pre-treatment (academic year 2023/24, January 2024 census) characteristics of treated schools (Parent Pact signup share at least 20%) and low-uptake schools (share below 5%). SEMH is averaged over 2021–2023. The child DLA rows are the school's catchment values, claims per 1,000 children aged 5–15 (calendar-year means of the quarterly stock, averaged over 2021–2023), for the psychiatric total and its behavioral and developmental components. Standardized difference is the difference in means over the pooled standard deviation.

Table 2. Effect of high pact uptake on the two headline outcomes: school SEMH and child psychiatric DLA.

	SEMH		Child DLA psych.	
	(1)	(2)	(3)	(4)
Average post-treatment effect	$-8.2\%^{***}$	$-11.3\%^{***}$	$-6.0\%^{***}$	$-7.1\%^{***}$
	(2.9)	(4.1)	(1.7)	(2.0)
LAD×year FE	No	Yes	No	Yes
Treated units	329	329	330	330
Observations	10,479	10,479	10,560	10,560

Notes: Each entry is the average of the two post-treatment matched TWFE event-study coefficients as a percent of the FE counterfactual (the level the treated units would have reached on the control time path). Each outcome uses its own matched sample (treated $\geq$20% pact uptake matched to three $<$5% controls). School SEMH is matched on its 2021–2023 levels, IMD-2019, FSM share and log enrollment. Child psychiatric DLA is matched on the catchment's 2021–2023 claim levels, the anchor school's IMD-2019 and FSM share, and log child population. All columns include unit and period fixed effects. Standard errors (LAD-clustered, in the same percent units) in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 3. Dose-response: effect by pact-uptake threshold.

	Treated if uptake at least			
Outcome	≥10%	≥20%	≥30%	≥40%
Panel A: School SEMH (per 1,000 pupils)				
SEMH	−2.9%*	−8.2%***	−8.7%	−24.9%***
	(1.7)	(2.9)	(5.4)	(9.4)
Treated schools	1,099	329	116	40
Panel B: Child psychiatric DLA (per 1,000 children 5–15)				
Child DLA psych.	−3.6%***	−6.0%***	−7.7%***	−7.2%**
	(1.0)	(1.7)	(2.3)	(3.5)
Treated units	1,073	330	118	40

Notes: Each cell is the average post-treatment effect as a percent of the FE counterfactual, from a separate matched TWFE event study comparing units above the stated pact-uptake threshold to matched <5% controls (1:3). School SEMH is matched on its 2021–2023 levels, IMD-2019, FSM share and log enrollment. Child DLA is estimated on the school-catchment panel at every threshold. Treated catchments are those whose anchor school's uptake is at or above the threshold, matched on their 2021–2023 claim levels, the anchor school's IMD-2019 and FSM share, and log child population. Controls are drawn from all <5% catchments in the national catchment cube, so the dose progression holds the geographic design fixed across columns. The ≥20% DLA rung is the headline design of Table 2. Standard errors (LAD-clustered, in percent-of-counterfactual units) in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4. DLA effect of high pact uptake by age, sex, and the behavioral component.

Outcome	Matched	+LAD×yr
Child psychiatric DLA (5–15)	−6.0%***	−7.1%***
	[−9.4, −2.7]	[−11.1, −3.2]
Behavioral component	−10.3%***	−8.9%***
	[−15.4, −5.1]	[−15.1, −2.7]
By age		
Ages 5–10	−8.7%***	−5.6%*
	[−13.4, −4.1]	[−11.7, 0.5]
Ages 11–15	−3.9%**	−3.0%
	[−7.8, −0.1]	[−8.2, 2.1]
By sex		
Boys	−3.6%*	−3.7%
	[−7.8, 0.7]	[−9.8, 2.4]
Girls	−7.8%***	−8.7%**
	[−13.3, −2.2]	[−15.5, −2.0]
Adults (top of the age gradient; external to the child caseload)		
Adult depression (QOF, 18+)†	0.6%	0.2%
	[−0.4, 1.5]	[−0.7, 1.1]

Notes: Each cell is the average post-treatment effect as a percent of the FE counterfactual, from a matched TWFE event study (treated ≥20% pact uptake to three <5% controls, with candidates drawn from all catchments in the national cube). Each outcome is matched on its own 2021–2023 catchment claim levels, the anchor school's IMD-2019 and FSM share, and log child population, the same own-outcome design as Table 2. The behavioral component (psychoneurosis, behavioral disorder, hyperkinetic/ADHD) is the part of the caseload a smartphone-delay channel would plausibly move. The developmental remainder appears as a placebo in Table 6. Boys and girls split the all-psychiatric caseload. "+LAD×yr" adds local-authority-by-year fixed effects. 95% confidence intervals (LAD-clustered) in brackets. Treated units: 330 catchments. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. †Adult depression (QOF, 18+) is GP-recorded prevalence at LSOA level (723 treated LSOAs), matched on its own 2021–2022 levels and IMD-2019. Adults are not a component of the child DLA caseload.

Table 5. The two headline outcomes under three estimators: matched TWFE, synthetic difference-in-differences, and triple differences.

	Matched TWFE	SDID	Triple difference
School SEMH	-2.92^{***}	-1.69^{*}	-1.97^{*}
	(1.02)	(1.02)	(1.09)
	[−4.92, −0.91]	[−3.70, 0.31]	[−4.11, 0.18]
	[−8.2% of base]	[−4.8% of base]	[−6.1% of base]
Child psychiatric DLA	-3.02^{***}	-2.35^{***}	-1.68^{**}
	(0.86)	(0.74)	(0.73)
	[−4.70, −1.35]	[−3.80, −0.90]	[−3.11, −0.25]
	[−6.0% of base]	[−4.7% of base]	[−9.4% of base]

Notes: Effects in claims or pupils per 1,000 (standard errors in parentheses: LAD-clustered for TWFE and the triple difference, jackknife for SDID), averaged over the two post-treatment years. Section 6.1 describes the three estimators and their samples. Brackets give the 95 percent confidence interval (per 1,000) and the estimate as a percent of its own base: the FE counterfactual for TWFE and SDID, the treated moving category's post-treatment mean for the triple difference. Treated units: 329 schools and 330 catchments for the matched TWFE, and 325 schools and 328 catchments for the SDID and triple-difference panels. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 6. Main outcomes against placebo outcomes.

	Matched	+LAD×yr
Main outcomes		
School SEMH	$-8.2\%^{***}$	$-11.3\%^{***}$
	[−13.9, −2.6]	[−19.4, −3.2]
Behavioral DLA component	$-10.3\%^{***}$	$-8.9\%^{***}$
	[−15.4, −5.1]	[−15.1, −2.7]
Placebo outcomes		
Autistic spectrum (school)	0.5%	−3.4%
	[−6.0, 7.0]	[−11.0, 4.2]
Developmental DLA component	−2.2%	−1.4%
	[−6.0, 1.7]	[−6.5, 3.7]

Notes: Average post-treatment effect as a percent of the FE counterfactual, with 95% confidence intervals (LAD-clustered) in brackets, from matched TWFE event studies on own-outcome matched samples (treated ≥20% pact uptake to three <5% controls). School outcomes are matched on their own 2021–2023 levels, IMD-2019, FSM share and log enrollment. Catchment outcomes are matched on their own 2021–2023 claim levels, the anchor school's IMD-2019 and FSM share, and log child population. "+LAD×yr" adds local-authority-by-year fixed effects. The DLA row is the behavioral component rather than the 5–15 total so that the main and placebo rows partition the caseload into disjoint components with independent sampling error (the total appears in Tables 2 and 4). $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 7. Placebo-in-time test for the two headline outcomes.

	Real effect (per 1,000 [95% CI])	Placebo (fake February 2022) (per 1,000 [95% CI])
School SEMH		
Matched TWFE	−2.92 [−4.92, −0.91]	0.11 [−1.68, 1.90]
SDID	−1.69 [−3.70, 0.31]	0.20 [−1.60, 2.01]
DDD: SEMH vs autistic spectrum	−1.97 [−4.11, 0.18]	0.22 [−1.68, 2.11]
Child psychiatric DLA		
Matched TWFE	−3.02 [−4.70, −1.35]	−1.16 [−2.33, 0.02]
SDID	−2.35 [−3.80, −0.90]	−0.89 [−2.02, 0.24]
DDD: behavioral vs developmental	−1.68 [−3.11, −0.25]	0.43 [−0.86, 1.72]

Notes: Each cell is the estimate in pupils or claims per 1,000 with its 95 percent confidence interval. "Real effect" uses the genuine February 2024 treatment date and averages the two post-treatment years. "Placebo" re-dates the treatment to February 2022 and is estimated on pre-treatment data only, with the matched TWFE rows re-matched on the 2019–2021 panel years (Section 6.2 describes the designs). Confidence intervals are LAD-clustered for the TWFE and DDD rows and unit-jackknife for the SDID rows.

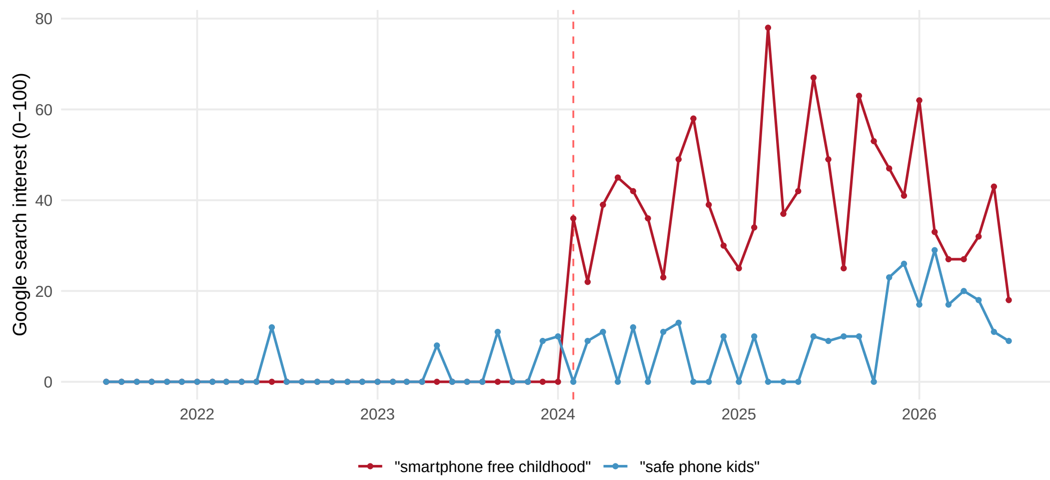

Figure 1. Google search trends on SFC and phone safety.

Notes: Monthly Google Trends search interest in the United Kingdom for “smartphone free childhood” and “safe phone kids,” the two queried together on one 0–100 scale. The dashed red line marks February 2024, the movement’s launch. “Smartphone free childhood” registers no searches before that date.

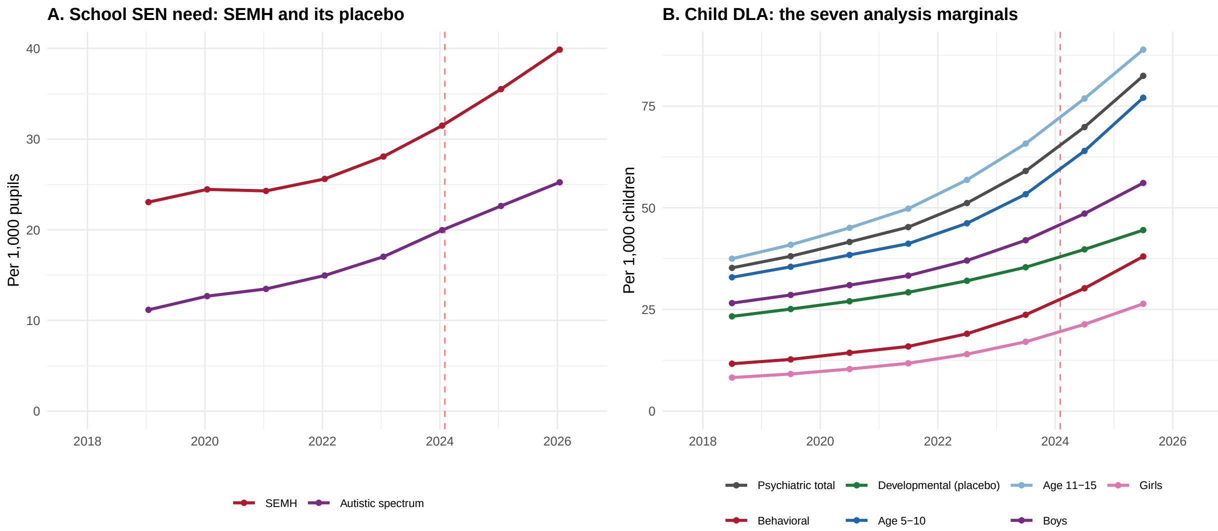

Figure 2. Evolution of the analysis outcomes, England.

Notes: Panel A: SEMH and the autistic-spectrum category, per 1,000 state-primary pupils by January census (the series begin at the January 2019 census, when coverage stabilizes at the full ∼16,700 primaries). Panel B: the seven child DLA marginals, per 1,000 children (annual calendar-year means of the quarterly claim stock), England-wide over all school catchments: the ages 5–15 psychiatric total, its behavioral and developmental split, its age split (5–10 and 11–15, each per its own population), and its sex split. The dashed vertical line (February 2024) marks the start of the movement.

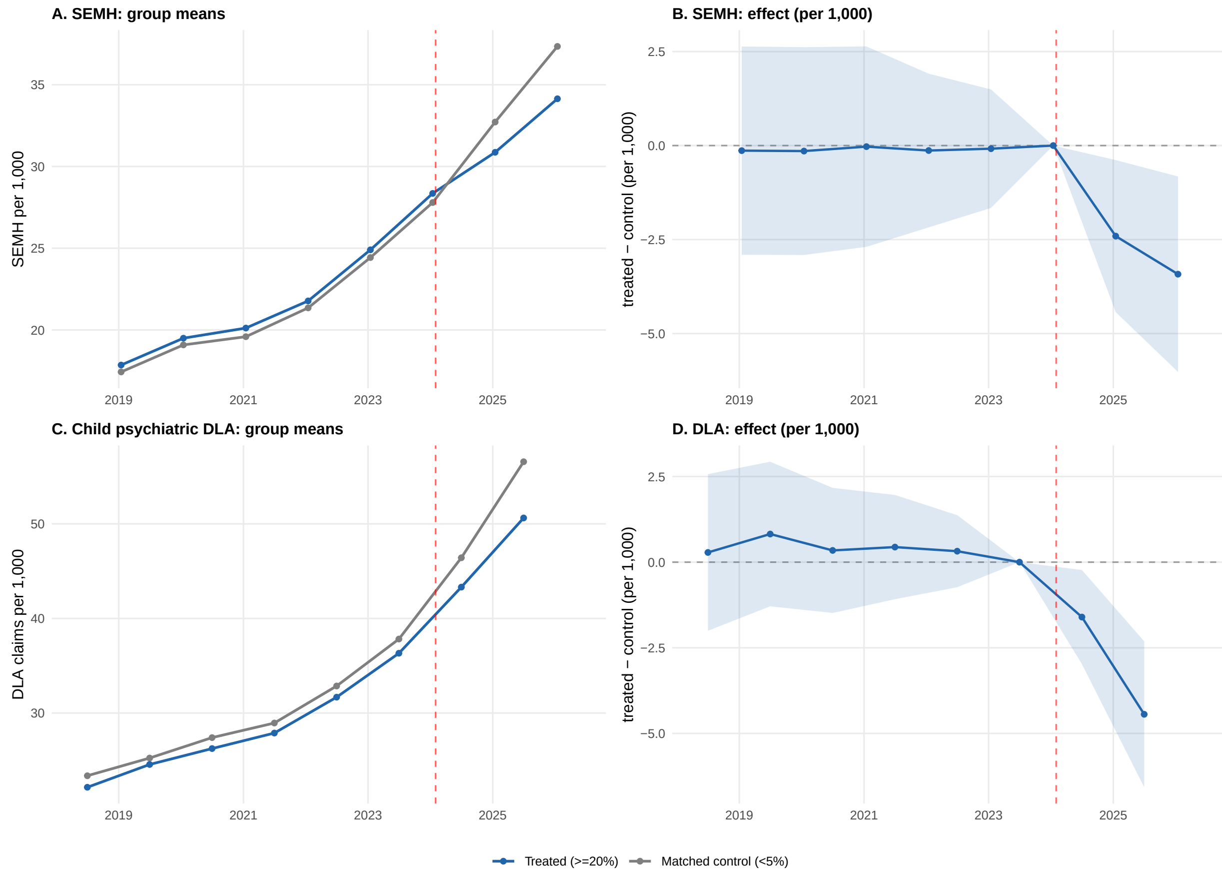

Figure 3. Matched event studies for the two headline outcomes.

Notes: School SEMH (top) and the DWP outcome, child psychiatric DLA claims (bottom). High-uptake ≥20% versus matched low-uptake <5% units. Left: group means for treated and matched-control units (each unit weighted equally, as in the matching). Right: the differenced two-way fixed-effects event-study coefficients (treated minus control, per 1,000), with 95% confidence intervals (LAD-clustered). Both series are annual: SEMH by January census year, and child psychiatric DLA as calendar-year means of the quarterly DWP claim stock, with each neighborhood assigned to its nearest state-primary catchment. The dashed vertical line (February 2024) marks the start of the movement, the boundary between the last pre-treatment and first post-treatment annual observation.

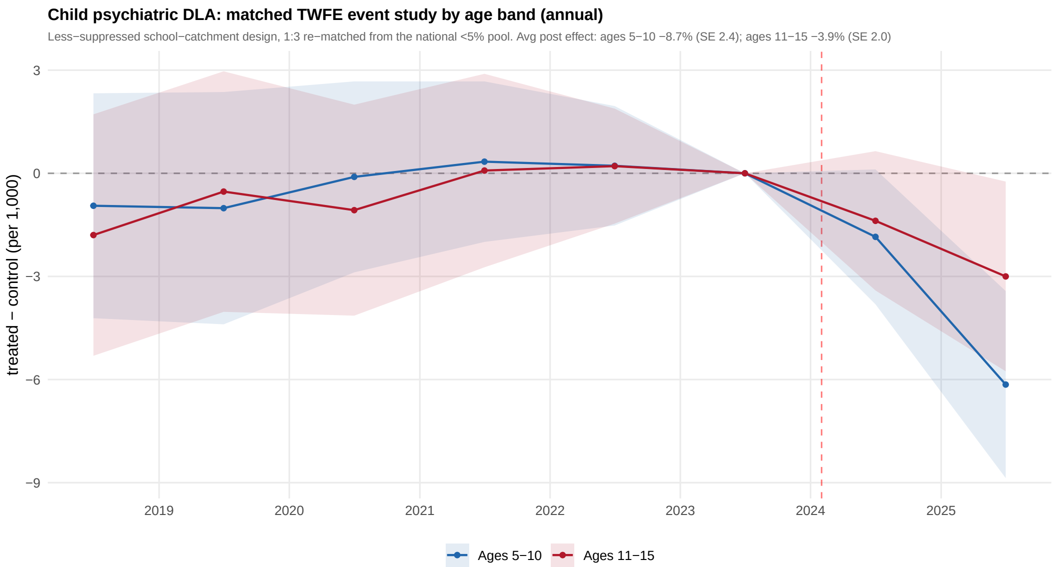

Figure 4. Matched event study for child psychiatric DLA claims, ages 5–10 versus 11–15.

Notes: High-uptake (≥20%) versus matched low-uptake (<5%) school catchments (1:3, annual calendar-year means, each band re-matched on its own pre-treatment levels), as the differenced two-way fixed-effects coefficient (treated minus control, per 1,000 of the age group). Each neighborhood (LSOA) is assigned to its nearest state primary school and aggregated to the catchment. The dashed vertical line (February 2024) marks the start of the movement. Table 4 reports the corresponding estimates.

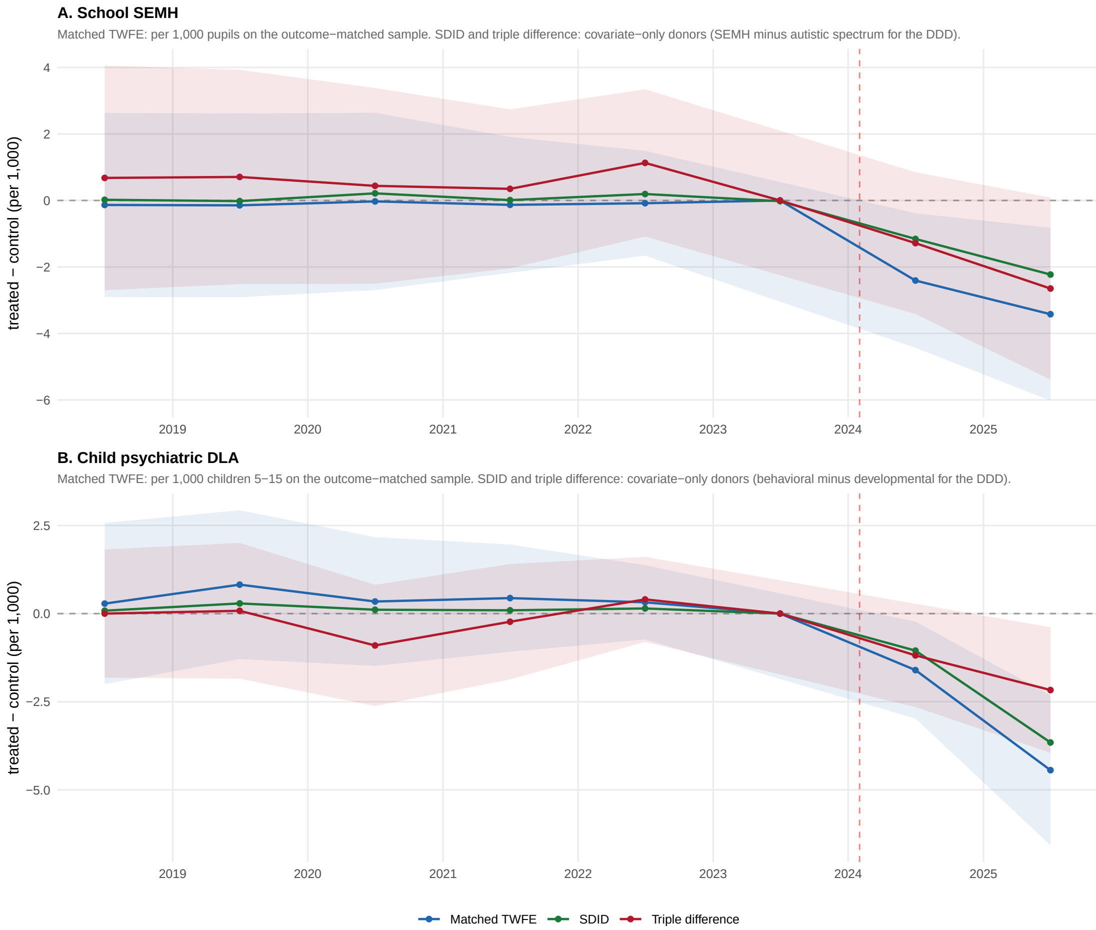

Figure 5. Event studies for the two headline outcomes under the three estimators.

Notes: Panel A: school SEMH per 1,000 pupils. Panel B: child psychiatric DLA claims per 1,000 children aged 5–15. Matched TWFE: differenced event-study coefficients on the matched sample, with 95 percent LAD-clustered intervals. SDID: treated minus synthetic control on the covariate-only donor set, recentered so that its weighted pre-period mean is zero (no analytic per-period interval). Triple difference: per-period coefficients on the covariate-only match, per 1,000 of the moving category (SEMH minus autistic spectrum in Panel A, behavioral minus developmental in Panel B), with 95 percent intervals. The dashed vertical line (February 2024) marks the start of the movement.

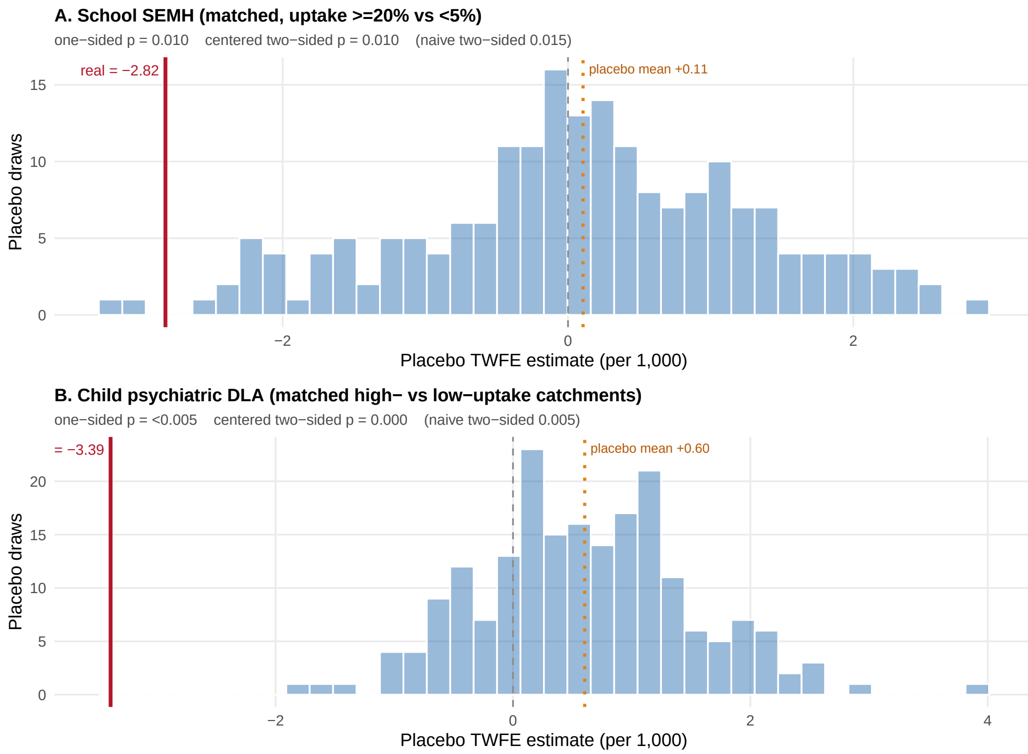

Figure 6. Placebo-in-space permutation distributions for the two headline outcomes.

Notes: (A) school SEMH and (B) child psychiatric DLA. Each histogram is the distribution of placebo matched-TWFE difference-in-differences estimates across calibrated placebo draws, in which treatment is reassigned to a same-size, same-clustering set of never-treated communities and re-matched each draw. The solid red line is the actual estimate, the dashed gray line is zero, and the dotted line is the placebo mean. The child DLA permutation draws its placebo sets from the full national pool of never-treated school catchments. Each panel prints the one-sided, mean-centered two-sided, and uncentered two-sided permutation p-values.

A Data Appendix

A.1 Sample Construction

Table 8 describes the sample selection.

[Table 8 about here.]

A.2 Timeline of UK Smartphone Initiatives

Table 9 lists the major UK smartphone ownership-delay and in-school ban initiatives over the study period, distinguishing headteacher-led ownership-delay pledges, grassroots parent pacts, and in-school bans (which restrict use on premises but not ownership; Section 2).

[Table 9 about here.]

A.3 Additional Descriptive Tables

Table 10 gives the cross-school distribution of Parent Pact sign-up shares and the geographic clustering of the high-uptake group (Section 3.1). Table 11 sets out the census primary-need categories and their frequencies among primary pupils.

[Table 10 about here.]

[Table 11 about here.]

A.4 School Census: Additional Variables

Free School Meals (FSM). Eligibility for Free School Meals is determined by whether a pupil's parent or guardian is in receipt of qualifying benefits, principally Universal Credit with a household income below £7,400 per year (after tax, excluding benefits). FSM eligibility is the standard proxy for socioeconomic disadvantage in English school data. I use the current FSM rate (the share of pupils eligible for free school meals), averaged over the 2021–2023 matching window, as the school-level socioeconomic covariate in the matched design.

Other school characteristics. Additional covariates include total enrollment, school type (community, academy, voluntary-aided, voluntary-controlled, free school), Ofsted rating, and the rates of other SEN categories (communication and interaction, cognition and learning, sensory and physical needs).

Census timing. The January census is conducted in the third week of January each year, and SEN results are typically published the following June. Academic years run September–August. The "year" variable in the replication panel's data files refers to the year in which the academic year begins (year = 2024 denotes AY 2024/25, with a January 2025 census). The paper's text and exhibits label observations by their measurement date instead, so that census appears as year 2025 (Section 4).

A.5 DLA and QOF Data

DLA CYP psychiatric claims. Section 3 sets out the construction of the DLA outcome. Two details supplement it here. First, DLA is an under-16 benefit, and claimants transition to Personal Independence Payment at age 16, so the under-16 caseload is the complete count of child claims. The published data also report a small residual 16–17 band, which I do not use. Second, Table 12 tabulates the share of recorded-zero cells by treatment group, for the raw LSOA panel and for every catchment-cube marginal. In the raw LSOA panel, recorded zeros are markedly more common in the treated neighborhoods (25.6 against 10.0 percent of cells), which are more affluent and carry lower caseloads. Any resulting bias runs against the estimated reduction. As caseloads grow, cells clear the suppression floor (the all-cell zero share falls from 20.1 percent in 2018 to 3.9 percent in 2025), which raises recorded treated rates mechanically relative to control. The suppressed-cells-at-4 row of Table 18 bounds the estimate under the opposite extreme. In the catchment cube, pooling neighborhoods keeps the headline total far below the suppression floor and the treated and matched-control zero shares close (8.9 against 7.7 percent), while the sparser splits (the behavioral component, the 11–15 band, the sex split) carry more zeros on both sides.

[Table 12 about here.]

Adult depression (QOF, placebo). As an adult placebo I use recorded depression prevalence among GP-registered adults (18+) from the Quality and Outcomes Framework (QOF), mapped from practice to LSOA level using patient-registration shares. No values are imputed. The COVID-disrupted FY 2020/21 register was not collected and is left as a gap (interpolating it would add no information and only re-weight its neighbouring years). FY 2023/24 is left out of the panel altogether, because the QOF depression register changed from a prevalence to an incidence basis that year and is not comparable. The event study therefore references the last observed pre-treatment year (FY 2022/23), so the pre-versus-post anchor is always a genuine observation. Matching uses the 2021 and 2022 levels. QOF runs annually from 2013, giving a long pre-treatment window. I use it rather than working-age psychiatric disability (PIP) claims, which begin only in 2022 and would leave just two pre-treatment periods. A primary-school smartphone-delay pact cannot affect adult depression, so a movement here would indicate a confounding area-level shock rather than an effect on children.

A.6 Composition of the SEN categories and DLA conditions

The school SEN measure and the DLA measure each record a single category per child, but the two taxonomies differ, so I set out here what each contains and how they correspond.

School SEN categories. The school census assigns every pupil on SEN support or with an Education, Health and Care plan a single primary type of need, following the Department for Education's SEND Code of Practice (2015). I group the census types into the categories used in the paper. Social, emotional and mental health (SEMH), the headline moving outcome, is described in the Code (paragraph 6.32) as covering anxiety, depression, self-harm, substance misuse, eating disorders, attachment difficulties and, explicitly, attention deficit disorder and attention-deficit hyperactivity disorder (ADHD). Speech, language and communication needs (SLCN) covers difficulty in producing or understanding speech and language and in applying the social rules of communication. The placebo categories are cognition and learning (specific, moderate, severe and profound learning difficulty, the first of which includes dyslexia, dyscalculia and dyspraxia), autistic spectrum (ASD), and physical and sensory need (physical disability together with hearing, visual and multi-sensory impairment). I report these specific need types rather than the Code's four broad areas, because the broad "communication and interaction" area combines SLCN with the autistic spectrum, the paper's placebo category.

DLA conditions. The DWP records one main disabling condition per claim. Five are psychiatric: learning difficulties, psychosis, psychoneurosis (anxiety and mood disorders), behavioral disorder, and hyperkinetic syndrome (ADHD).[32] I split these into a behavioral component (psychoneurosis, behavioral disorder and hyperkinetic syndrome, the emotional and behavioral conditions a smartphone-delay effect would plausibly move) and a developmental component (learning difficulties and psychosis, structural conditions it should not affect, psychosis being negligible among children). Among child claims these are about 38 percent behavioral and 60 percent developmental. The developmental component serves as the within-caseload placebo; the small residual of physical, non-psychiatric conditions (for example asthma, epilepsy and cerebral palsy) is too small-cell suppressed at the catchment level to report separately: an attempted matched estimate on it carried both a failed pre-trend (joint $p = 0.03$) and a standard error near 7 percent of its counterfactual, too imprecise to inform the falsification, so it is not reported. Unlike its working-age successor PIP, DLA has no separate autism category, so autistic-spectrum claims are recorded chiefly under learning difficulties; autism therefore sits in the developmental component, consistent with its placebo status in the school data.

Correspondence. Table 13 maps the more common conditions across the two systems. The emotional and behavioral conditions (ADHD, anxiety and mood, conduct) are moving outcomes in both: SEMH in the school data and the behavioral component in DLA. The structural conditions (learning difficulty and autism) and the physical and sensory conditions are placebos in both. Speech and communication need has no DLA analogue and appears only in the school data.

[Table 13 about here.]

A.7 Parent Pact Registrations in the Headline Pledge Towns

The following tables report cumulative SFC Parent Pact registrations per school (as of 12 June 2026) in the three headline pledge towns, alongside a Hertfordshire comparison benchmark. They document the wide within-town variation in sign-up rates that motivates measuring treatment by school-level uptake rather than by town.

[Table 14 about here.]

[Table 15 about here.]

[32] The Department publishes these category labels without definitional text (Stat-Xplore metadata, "Main Disabling Condition"), so the clinical composition of the administrative categories is not further documented; the parenthetical glosses (hyperkinetic syndrome as ADHD, psychoneurosis as anxiety and mood disorders) follow the standard usage of research with these data. None of the paper's claims rests on the finer contents of the behavioral-disorder category: the developmental–behavioral boundary is drawn at learning difficulties and psychosis versus the remainder, and the absence of anxiety and depression as recorded main conditions is a property of the published counts themselves.

A.8 Software

The analysis was carried out in R, using the `data.table`, `fixest`, `MatchIt`, `synthdid`, `arrow`, and `ggplot2` packages. Claude Fable 5 and Claude Opus 4.8 (Anthropic) assisted with code development and manuscript editing. Responsibility for all errors remains my own.

B Additional Results

B.1 Synthetic difference-in-differences

My robustness estimator is the synthetic difference-in-differences (SDID) estimator of Arkhangelsky et al. (2021), which solves

$$(\hat{\tau}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \arg \min_{\tau,\mu,\alpha,\beta} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\,\tau\right)^2 \hat{\omega}_i\,\hat{\lambda}_t, \tag{3}$$

where $W_{it} = \mathbf{1}[i \in \text{treated}] \cdot \mathbf{1}[t > T_0]$, the unit weights $\hat{\omega}_i$ are chosen so that the weighted donor units track the treated schools' pre-treatment trajectory, and the time weights $\hat{\lambda}_t$ emphasize the pre-treatment periods most comparable to the post period. SDID builds a synthetic control that matches the treated pre-trend, so a residual pre-trend in an unweighted event study is handled by reweighting. I apply SDID on the covariate-only donor sets of the triple difference (FSM, IMD-2019, and log size, with no outcome levels entering the match), so the weights, not the match, fit the treated pre-trend, and re-dating treatment falsifies the design directly (Table 7).

Table 16 reports the SDID estimates with unit-jackknife standard errors, computed on the covariate-only donor sets and expressed against the FE-counterfactual base of the matched TWFE tables. The estimates are same-signed and smaller than the matched TWFE ones. School SEMH is −1.69 against −2.92 per 1,000, and the DLA total is −2.35 against −3.02. The splits reproduce the TWFE pattern. The reduction concentrates at ages 5–10 and in the behavioral component, the developmental placebo is near zero, and the girls' estimate is the more precise of the sex pair.

[Table 16 about here.]

B.2 Additional Results Tables and Figures

This section collects the exhibits summarized in the main text: the affluence heterogeneity of Section 5.2 (Table 17) and the specification robustness of Section 6 (Table 18).

[Table 17 about here.]

[Table 18 about here.]

[Figure 7 about here.]

Table 8. Sample selection from the SFC registry to the analysis samples.

Stage	Schools
Registered schools, SFC Parent Pact snapshot (June 2026, all nations)	15,653
Signup share $\geq$20 percent (all nations)	482
of which England, with a valid DfE URN	431
of which state-funded primary schools (the treatment definition)	372
of which in the balanced 2019–2026 census panel (the treated group)	329
Catchments anchored on high-uptake schools (DLA analysis)	330

Notes: The balanced panel requires a complete SEMH series across the January 2019–2026 censuses (14,260 schools, of which 12,012 are $<$5 percent low-uptake schools). Each school's IMD-2019 value is that of the 2011 LSOA containing its GIAS coordinates (point-in-polygon against the 2011 boundary polygons), the same coordinate rule that assigns each LSOA to its nearest-school catchment. The handful of schools without usable coordinates fall back to their published census code, and no school drops for missing matching data. The DLA catchments are anchored on the high-uptake schools directly, whether or not the school appears in the census panel (Section 3), so the school and catchment counts differ for data-availability reasons alone.

Table 9. Timeline of UK smartphone ownership-delay and school ban initiatives

Date	**Area**	**Type**
Feb 2024	SFC movement launches nationally	Parent movement
May 2024	**St Albans** (20→33 primary schools)	Ownership delay (headteacher-led)
Jun 2024	Southwark (18/20 secondary schools)	School ban (headteacher-led)
Sep 2024	**SFC Parent Pact tool launched**	**Parent pact (national)**
Sep 2024	Ormiston Trust (44 schools)	School ban
Sep 2024	NI Dept. of Education guidance	Government + parent pacts
Sep 2024	Hitchin (town-wide, all 25 primaries invited)	Ownership delay (school-led)
Sep 2024	Farnham (individual schools)	Parent pacts (grassroots)
Dec 2024	**Cambridge CB4** (13 headteachers, 10 schools)	Ownership delay (headteacher-led)
Dec 2024	Brighton & Hove (council debate)	School ban
Jan 2025	Ealing (72 primary schools)	School ban
Feb 2025	Barnet (126 schools, borough-wide)	School ban
May 2025	Harpenden (19 primary schools)	School ban + parent pact
Jul 2025	Hampshire (109 primary schools)	School ban
Mar 2026	Surrey (58 secondary schools, county-wide letter)	Ownership delay (headteacher-led)
Jun 2026	England (Children's Wellbeing and Schools Act, s.36)	Statutory school guidance

Table 10. Distribution and geographic clustering of Parent Pact sign-up rates across primary schools.

Panel A: cross-school distribution of sign-up share	
Primary schools in panel	14,260
with at least one registered pact	8,630 (60.5%)
Median share among schools with pacts (%)	2.0
90th / 99th percentile of share (%)	8.0 / 28.6
High uptake (≥20%, treated)	329 (2.3%)
Intermediate (5–20%)	1,919 (13.5%)
Control (<5%)	12,012 (84.2%)
Panel B: geographic clustering of high-uptake schools	
High-uptake schools	329
Districts with ≥1 high-uptake school	118
exactly one	51
two to five	52
six or more	15
Most high-uptake schools in a single district	14
Share in the ten most concentrated districts (%)	30.7

Notes: Parent Pact sign-up shares (registrations over enrollment, public SFC API as of 12 June 2026) for the balanced panel of state-funded primary schools. Schools without an SFC listing count as zero. Panel A gives the cross-school distribution. Panel B describes the geographic concentration of the high-uptake (treated) schools across local authority districts, the level at which standard errors are clustered throughout the paper.

Table 11. Special educational need by primary type of need: state primary schools, January 2024.

Area of need	DfE primary-need codes	Per 1,000 pupils	% of SEN
Social, emotional and mental health (SEMH) †	SEMH	31.5	18.4
Cognition and learning	SpLD, MLD, SLD, PMLD	39.4	23.1
Speech, language and communication	SLCN	59.4	34.8
Autistic spectrum disorder	ASD	20.0	11.7
Physical and sensory ‡	PD, HI, VI, MSI	7.9	4.6
Other or not assessed	OTH, NSA	12.6	7.4
All SEN		170.8	100.0

Notes: Pupils with special educational needs in state-funded primary schools, by their recorded primary type of need (each pupil has one), combining the SEN-support and EHC-plan tiers, January 2024 census. Codes: SpLD/MLD/SLD/PMLD, specific/moderate/severe/profound-and-multiple learning difficulty; SLCN, speech, language and communication needs; PD, physical disability; HI/VI/MSI, hearing/visual/multi-sensory impairment; ASD, autistic spectrum disorder; OTH, other; NSA, SEN support with type of need not assessed. †School outcome. ‡Used as a placebo (medically determined, so a smartphone pledge cannot affect it).

Table 12. Recorded zeros in the child-DLA data: share of cells recorded as zero, treated versus control.

	All quarters		Post years (2024–25)	
	Treated	Control	Treated	Control
Panel A: LSOA quarterly cells (psychiatric total, ages 5–15)				
All LSOA-quarters	25.6	10.0	14.6	3.9
first year (2018), all cells	20.1			
last year (2025), all cells	3.9			
Panel B: catchment-cube quarterly cells, by marginal				
Psychiatric total (5–15)	8.9	7.7	3.1	2.2
Behavioral component	45.3	39.4	24.7	18.2
Developmental component	16.2	15.9	9.5	9.1
Ages 5–10	28.7	25.4	18.2	13.0
Ages 11–15	23.2	22.3	10.9	11.1
Boys	15.6	12.9	7.4	5.7
Girls	46.5	44.9	28.3	23.5

Notes: Under the DWP's Stat-Xplore small-cell rule a true count of one to four is withheld and returned as zero, so a recorded zero is either a genuine zero or a suppressed small count. Panel A covers the raw LSOA quarterly panel (treated: nearest school's pact uptake $\geq$20%; control: <5%), with the all-cell zero share also shown for the first and last panel years. Panel B covers the catchment cube on the headline estimation sample (treated catchments against their matched 1:3 controls), whose counts are retrieved as single custom-geography totals, so suppression applies to the catchment cell itself. Cells are unit-quarters over May 2018 to November 2025. The post-treatment columns restrict to calendar years 2024–2025, the post years of the annual estimation.

Table 13. Correspondence of common conditions across the school SEN and DLA classifications.

Condition	School SEN category	DLA main condition (component)	Role
ADHD / hyperkinetic	SEMH	Hyperkinetic syndrome (behavioral)	moving
Anxiety, depression	SEMH	Psychoneurosis (behavioral)	moving
Conduct, behavioral	SEMH	Behavioral disorder (behavioral)	moving
Autism	Autistic spectrum (ASD)	Learning difficulties† (developmental)	placebo
Learning difficulty	Cognition and learning	Learning difficulties (developmental)	placebo
Physical, sensory	Physical and sensory	Non-psychiatric (physical)	placebo

Notes: Each system records a single category per child. "Moving" denotes the emotional and behavioral categories a smartphone-delay effect is expected to reduce. "Placebo" denotes the structural, developmental and physical categories it should not affect. The "Anxiety, depression" row is a taxonomy correspondence only: psychoneurosis is essentially absent among child DLA claims (Section 3), so in-sample the DLA side of that row carries no cases and the behavioral component consists of behavioral disorder and hyperkinetic/ADHD. †DLA, unlike its working-age successor PIP, has no separate autism category, so autistic-spectrum claims are recorded chiefly under learning difficulties.

Table 14. SFC Parent Pact registrations: St Albans treated schools

School	Pacts	Share (%)
Alban City School	152	39.6
Cunningham Hill Junior School	92	38.3
Prae Wood Primary School	83	18.3
Wheatfields Junior Mixed School	83	23.1
St Peter's School	80	22.9
Bernards Heath Infant and Nursery School	77	25.5
St Michael's Church of England Voluntary Aided Primary Sch	73	38.0
Bernards Heath Junior School	66	18.5
Oakwood Primary School	63	19.8
Cunningham Hill Infant School	58	32.2
Skyswood Primary & Nursery School	57	24.1
Redbourn Primary School	48	11.9
Garden Fields Junior Mixed and Infant School	47	9.0
Aboyne Lodge Junior Mixed and Infant School	37	19.8
The Abbey Church of England Voluntary Aided Primary School	35	21.2
Wheatfields Infants' and Nursery School	35	10.9
Fleetville Junior School	32	9.0
How Wood Primary and Nursery School	32	14.3
St Adrian Roman Catholic Primary School	28	13.2
St Alban & St Stephen Catholic Primary School & Nursery	27	6.0
Camp Primary and Nursery School	25	10.7
Mandeville Primary School	25	6.3
Maple Primary School	23	10.6
Mount Pleasant Lane Primary School	21	9.2
Fleetville Infant and Nursery School	15	4.9
St Helen's Church of England Primary School	14	6.6
Killigrew Primary and Nursery School	13	3.2
Sandridge School	10	6.2
Colney Heath Junior Mixed Infant and Nursery School	7	3.3
Beech Hyde Primary School and Nursery	5	2.8
Windermere Primary School	5	2.8
Margaret Wix Primary School	4	2.7
St John Fisher Catholic Primary School, a Voluntary Academ	4	2.6
Bowmansgreen Primary School	2	0.6
Saint Bernadette Voluntary Catholic Primary Academy	2	—
London Colney Primary & Nursery School	0	—
Park Street Church of England Voluntary Aided Primary Scho	0	—
Total	1380	

Notes: Cumulative Parent Pact registrations per school as of 12 June 2026, from the public SFC Parent Pact API (counts are per child; share = pacts over enrollment as recorded by SFC). The Pact tool launched in September 2024, after the St Albans pledge, so these registrations formalize rather than define treatment. Schools without an SFC listing are shown with zero pacts.

Table 15. SFC Parent Pact registrations: replication towns and control benchmark

School	Pacts	Share (%)
Panel A: Hitchin (treated primaries)		
William Ransom Primary School	242	57.3
Samuel Lucas Junior Mixed and Infant School	186	44.3
Highover Junior Mixed and Infant School	121	28.3
St Andrew's Church of England Voluntary Aided Primary Scho	62	24.6
Wilshere-Dacre Junior Academy	41	15.4
Strathmore Infant and Nursery School	35	17.6
Whitehill Junior School	28	11.7
Highbury Infant School and Nursery	20	9.8
Mary Exton Primary School	13	7.2
Oughton Primary and Nursery School	11	5.0
Purwell Primary School	10	6.0
Our Lady Catholic Primary School	5	3.2
Panel B: Cambridge CB4 (treated primaries)		
Milton Road Primary School	132	31.9
Mayfield Primary School	115	28.2
Arbury Primary School	34	8.9
Kings Hedges Primary School	18	4.5
Shirley Community Primary School	16	4.6
Chesterton Primary School	10	5.5
The Grove Primary School	9	3.7
St Laurence Catholic Primary School	6	2.5
St Luke's CofE Primary School	1	1.1
Orchard Park Community Primary School	0	—
Panel C: group comparison		
Group	Total pacts	Median share (%)
St Albans (treated primaries) (37 schools, 35 with pacts)	1380	10.8
Hitchin (treated primaries) (12 schools, 12 with pacts)	774	13.5
Cambridge CB4 (treated primaries) (10 schools, 9 with pacts)	341	4.6
Hertfordshire excl. pledge towns (all SFC-listed schools) (358 schools, 358 with pacts)	4707	2.4

Notes: Same source and definitions as Table 14. The control benchmark covers all Hertfordshire schools listed by SFC outside the pledge towns (St Albans, Harpenden, Hitchin), including independent and secondary schools. Hertfordshire is the highest-pact county nationally, so the benchmark overstates typical control-school uptake elsewhere in England. Mean and 90th-percentile shares for the control benchmark are 4.8% and 12.8%, versus means of 14.4%, 19.2%, and 10.1% for St Albans, Hitchin, and Cambridge CB4 respectively.

Table 16. Synthetic difference-in-differences estimates on the covariate-only donor sets, with the matched TWFE estimate for comparison.

Outcome	SDID	(s.e.)	[95% CI]	% of cf	TWFE
School SEMH	−1.69	(1.02)	[−3.70, 0.31]	−4.8%	−2.92
Child psychiatric DLA (5–15)	−2.35	(0.74)	[−3.80, −0.90]	−4.7%	−3.02
Ages 5–10	−3.96	(0.97)	[−5.86, −2.06]	−8.7%	−4.00
Ages 11–15	−1.73	(1.25)	[−4.18, 0.71]	−3.1%	−2.19
Behavioral	−1.95	(0.56)	[−3.04, −0.86]	−9.8%	−2.04
Developmental (placebo)	−0.28	(0.55)	[−1.37, 0.80]	−0.9%	−0.64
Boys	−0.98	(0.62)	[−2.19, 0.22]	−2.9%	−1.20
Girls	−1.11	(0.45)	[−1.99, −0.23]	−7.1%	−1.22

Notes: Each row applies the synthetic difference-in-differences estimator of Arkhangelsky et al. (2021) to a covariate-only 1:3 donor set (treated ≥20% pact uptake versus <5% controls; school SEMH matched on FSM, IMD-2019 and log enrollment, the child DLA marginals on one shared catchment set matched on the anchor school's FSM and IMD-2019 and log child population; no outcome levels enter any match). The unit and time weights of the estimator, not the match, fit the treated pre-treatment trajectory. This is the same donor restriction as the triple difference, so re-dating treatment falsifies the design directly (Table 7). School SEMH is per 1,000 pupils on the 2018–2025 panel years. Child DLA marginals are per 1,000 children as catchment annual means. Standard errors are unit jackknife. "% of cf" expresses the SDID estimate against the same FE-counterfactual base as the matched TWFE tables, for comparability. "TWFE" is the corresponding matched TWFE estimate (per 1,000).

Table 17. Effect by affluence.

	Treated units, by affluence		
Outcome	All treated	More affluent	Less affluent
Panel A: School SEMH (split at median FSM)			
SEMH	$-8.2\%^{***}$	-2.2%	$-13.7\%^{***}$
	(2.9)	(4.9)	(4.0)
Mean FSM (%)	10.3	5.4	15.1
Panel B: Child psychiatric DLA (split at median IMD-2019)			
Child DLA psych.	$-6.0\%^{***}$	$-8.3\%^{***}$	$-3.9\%^{*}$
	(1.7)	(2.5)	(2.3)

Notes: Each cell is the average post-treatment effect as a percent of the FE counterfactual, from a matched TWFE event study (treated ≥20% pact uptake to three <5% controls), for all treated units and for treated units split at their median pre-treatment affluence. Panel A splits schools at the median FSM share (9.1%). Panel B splits treated catchments at their median anchor-school IMD-2019. Standard errors (LAD-clustered, percent-of-counterfactual units) in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 18. Specification robustness for the two headline outcomes.

Specification	SEMH est. (s.e.)	Child DLA psych. est. (s.e.)
Main: matched, ≥20% vs <5%	−2.82 (1.11) [−5.01, −0.64]	−3.39 (1.15) [−5.64, −1.14]
Low-uptake pool <10%	−2.48 (1.12) [−4.67, −0.29]	−2.86 (1.14) [−5.09, −0.63]
Match ratio 1:5	−2.98 (1.02) [−4.98, −0.99]	−3.52 (1.08) [−5.63, −1.40]
Matching with replacement	−3.18 (1.17) [−5.47, −0.88]	−3.75 (1.16) [−6.03, −1.47]
Excl. St Albans, Hitchin, Cambridge	−3.23 (1.14) [−5.47, −0.99]	−3.24 (1.06) [−5.31, −1.16]
Extended pre-period (2015+)	−2.96 (2.36) [−7.59, 1.67]	–
Suppressed cells = 4 (DLA)	–	−3.15 (1.18) [−5.47, −0.83]
Quarterly panel (DLA)	–	−3.39 (1.14) [−5.62, −1.17]

Notes: Each cell is the matched TWFE difference-in-differences coefficient (treated × post, per 1,000) with its LAD-clustered standard error in parentheses and 95% confidence interval in brackets, under the stated specification (main match 1:3). School SEMH is per 1,000 pupils. Child psychiatric DLA is per 1,000 children aged 5–15, on the school-catchment panel in every row. Every DLA row draws its 1:3 controls from all catchments in the national catchment cube, matched on the outcome's 2021–2023 levels, the anchor school's IMD-2019 and FSM share, and log child population (the headline design). The suppressed-cell row imputes every catchment-quarter recorded as zero (a true count of 0–4 under the DWP disclosure rules) at its upper bound of 4. The quarterly-panel row estimates the same pooled treated *imes* post difference-in-differences on the raw quarterly claim snapshots (November 2023 the last pre-treatment quarter, February 2024 the first post) rather than the calendar-year means, on the same matched catchments. "Excl. St Albans, Hitchin, Cambridge" drops the three media-prominent pledge towns from the treated group. The extended pre-period balances the school panel back to the 2015 census and is unavailable for child DLA, whose series begins in 2018.

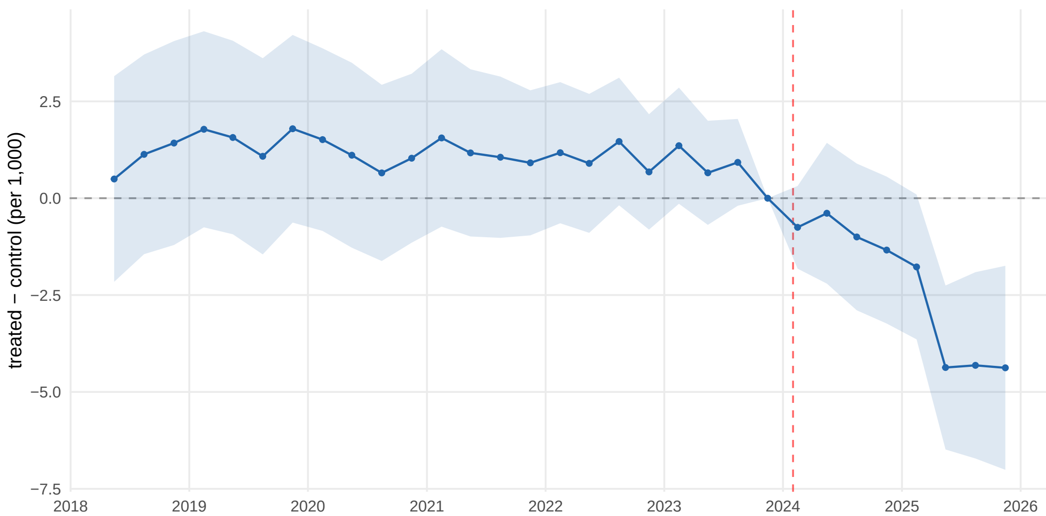

Figure 7. Quarterly child psychiatric DLA event study.

Notes: Differenced two-way fixed-effects coefficients (treated minus control, per 1,000 children aged 5–15) on the raw quarterly DLA snapshots (February, May, August and November), rather than the calendar-year means used in the main text, for the headline matched catchments (≥20% versus matched <5% pact uptake). The reference is the last pre-treatment quarter, November 2023. The dashed red line marks February 2024, the first post-treatment quarter. Shaded bands are 95 percent LAD-clustered confidence intervals. The pooled treated × post difference-in-differences is −3.4 per 1,000 (Table 18).