SMARTPHONES AND CHILDREN'S MENTAL HEALTH

Gregory F. Veramendi*

September 2026

Abstract

I provide quasi-experimental evidence on the effect of delaying children's smartphone ownership through a community-led policy. Since 2024, the Smartphone Free Childhood movement has coordinated UK families through a "Parent Pact" to delay a child's first smartphone until age 14. Using a matched difference-in-differences design, I compare changes in children's mental-health needs across English schools. Treated schools are those where at least one fifth of families signed the pact. Control schools are similar schools where fewer than 5 percent signed. I find that mental-health special needs measured by schools fall by about 8 percent relative to control schools. Similarly, mental-health disability needs independently measured by the Department for Work and Pensions fall by about 6 percent relative to similar communities. The effects are driven by younger children, who are less likely to already own a phone. I also find that by the second year, pupils' illness absence and teacher turnover decline by 7 and 12 percent, respectively. The findings support coordinated ownership delay as a near-zero-cost complement to school phone bans and statutory age limits. They also 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, Daisy Greenwell, and participants of the Helsinki Economics of Education Workshop. This work received no external funding. Smartphone Free Childhood provided the signup data used in this paper. It had no role in the paper's design, analysis, interpretation, or reporting. It imposed no conditions on the use of the data or the reporting of results. I received no financial or other benefit. This work contains statistical data from the Office for National Statistics (ONS), which are Crown Copyright. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates. The Longitudinal Education Outcomes (LEO) analysis drawn on in Section 2.4 was carried out in the Secure Research Service, part of the ONS.

1 Introduction

Young people's mental health has deteriorated across high-income countries since the early 2010s (Blanchflower et al., 2025). The decline 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 focused on two types of restriction. First, schools have adopted phone bans in growing numbers. Their effects on mental health and well-being are mostly neutral or beneficial (Park et al., 2026). Second, several governments are introducing minimum ages for social media.[2] A third approach, in which parents delay a child's first smartphone, complements the government policies. There is, in short, little causal evidence on whether smartphone ownership harms children, and on which policies are effective (Orben et al., 2025).

This paper provides evidence on both questions. I evaluate an intervention that delays children's smartphone *ownership* 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 at more than 15,600 UK schools had signed its "Parent Pact." The pact is 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 high. At Cunningham Hill Junior School, Year 6 smartphone ownership fell from 75 to 12 percent within a year. Ownership fell 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. Denmark's government agreed in November 2025 to legislate a similar ban.

[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

I use variation in pact signup across state primary schools in England. The SFC records the number of pact registrations at each school. I classify a school as *treated* if at least 20 percent of its families had signed by June 2026. The treated group has 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. An event study then compares treated schools with their matched controls after February 2024, when the movement emerged. 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 a child with a psychiatric condition. The Department for Work and Pensions adjudicates these claims. 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.

I then examine who is most affected and how. I start by decomposing the DLA claims

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

into the effects for different populations and types of claims (Table 3). First, 71 percent of the effect is driven by behavioral claims, while ADHD and developmental claims show no clear change. Second, most of the effect occurs for children who are young enough to be affected by the pact. Behavioral disorder at ages 5 to 11 falls by 15.1 percent. Most 12 to 15 year-olds already owned a phone when the movement started, and their behavioral claims show no evidence of a change. Third, claims held for two years or more began before the movement and show no clear impact. In contrast, new behavioral-disorder claims at ages 5 to 11 fall by 21.5 percent. The pact therefore prevents new need rather than reversing existing need. Lastly, 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.

I provide five sets of robustness checks. First, a synthetic difference-in-differences estimator (Arkhangelsky et al., 2021) and a triple-differences estimator give similar estimates. Second, placebo outcomes show no effect. Adult depression in the same communities is unchanged. The autistic-spectrum school category is also unchanged. Within the DLA caseload, developmental disorders are unaffected. Behavioral conditions fall by about 10 percent. Third, a placebo treatment date before 2024 shows no effect. All estimators pass this test. Fourth, I reassign treatment to placebo sets of never-treated communities with the same size and clustering. The DLA estimate lies below every placebo estimate. The SEMH estimate lies in the far left tail of the placebo distribution. Fifth, the estimates are similar under changes to the low-uptake threshold, the match ratio, replacement in matching, and the pre-period length. They are also similar without the three pledge towns that received the most media attention.

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 children aged 5 to 11, 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.[4] 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 consistent with the hypothesis that early smartphone ownership harms children's mental health. They also indicate that delaying ownership substantially improves children's well-being.

To my knowledge, this paper provides the first econometric evidence on coordinated smartphone-ownership delay. It contributes to four literatures.

First, the paper contributes to causal evidence on internet access and screen time, the literature closest to an ownership delay. Broadband and mobile-internet rollout studies document negative effects on young people across countries (Donati et al., 2025; Golin, 2022; Churchill and Johnson, 2026; Jain and Stemper, 2026). Short-run randomized trials of screen-time limits support this evidence. Limiting smartphone screen time to two hours a day for three weeks reduces depressive symptoms among university students (Pieh et al., 2025). The gains fade once the restriction ends. This pattern favors interventions that change exposure for long periods, such as an ownership delay. To my knowledge, only one completed randomized trial includes children under 13 (Schmidt-Persson et al., 2024). It asked 45 of 89 Danish families to limit leisure screen time to three hours a week per person for two weeks. Parent-reported difficulties fell by 0.53 standard deviations relative to controls. These trials restrict single participants or families while their peers keep their phones. I add causal evidence on primary-school-age children, the age group most discussed in current policy debates. I also estimate general equilibrium effects, in which many children at the same school do not get a phone.

Second, the paper contributes to the causal evidence on social media. Braghieri et al. (2022) show that the staggered rollout of Facebook across US colleges worsened students' mental health. Allcott et al. (2020) find that deactivating Facebook in an experiment improved adults' well-being. Short-run randomized trials of social media limits support this

[4]Only two of the seven natural experiments reviewed by Park et al. (2026) have independent measures of phone use at school.

evidence. Limiting social media for a few weeks reduces depressive symptoms among university students (Hunt et al., 2018) and among youth with emotional distress (Davis and Goldfield, 2025). A meta-analysis of ten such trials in adults and students confirms the depression finding (May et al., 2025). Reviews of real-world restrictions lasting days to months find effects of about 0.2 standard deviations in adolescents and adults (Burnell et al., 2025; Liu et al., 2025).[5] These effects are comparable to the advantage of antidepressants over placebo for depression in children and adolescents (Locher et al., 2017). Because participants can guess the hypothesis, self-reported outcomes in these experiments are subject to experimenter demand effects (Ferguson, 2025). In this paper, school staff and Department for Work and Pensions case workers record the outcomes, not the families who signed the pact.

Third, the paper complements the literature on school phone policies. A review of seven natural experiments finds mostly beneficial or neutral effects (Park et al., 2026). Bans during the school day raise test scores for low-achieving students in England (Beland and 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 (Abrahamsson, 2026). The bans reduce girls' use of psychological health care. They also lower bullying for both sexes. They raise girls' grades, with the largest gains for disadvantaged girls. A US study of nearly 5,000 schools finds that in-school bans cut phone use during the school day (Allcott et al., 2026). The bans initially worsen discipline and well-being, which later recover. They do not improve average test scores. Spanish regional bans reduce bullying (Beneito and Vicente-Chirivella, 2022). A Florida county ban raised suspensions at first and test scores later (Figlio and Özek, 2025). A study of statewide bans in three US states uses the National Survey of Children's

[5] A broader review of 27 social media experiments, ranging from brief laboratory exposures to multi-week abstinence, finds a pooled effect indistinguishable from zero (Ferguson, 2025). That pooling assumes the designs estimate one effect. The review does not test that assumption. Its heterogeneity of 75 percent and its range of effects, from -0.37 to 0.80, make the assumption doubtful. Rausch and Haidt (2026) argue that the review blends heterogeneous designs, outcomes and durations. They note that brief abstentions can lower mood temporarily through withdrawal effects, while longer reductions improve mental health.

Health for ages 11 to 17 (Saffer, 2026). With at most two post-ban years, the study finds no clear change in screen time or parent-reported psychological well-being. The most directly relevant evaluation for the UK setting, the SMART Schools Study (Goodyear, 2025), finds no mental-health benefit from UK in-school restrictions. These policies restrict phone use during school hours and leave ownership and use outside school unchanged.

Fourth, the paper contributes to psychology studies of 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 about one in four (Sohn et al., 2019). Problematic use is associated with about threefold odds of depression and anxiety and with higher odds of poor sleep. A systematic review of mobile and wireless device use finds suggestive associations with poorer mental health (Girela-Serrano et al., 2024). The associations are most consistent for bedtime use. Sleep partly mediates them. The review stresses the heterogeneity and high risk of bias of the underlying observational studies. These associations cannot separate the effect of the smartphone from selection into use. Psychologists and teachers report improved attention and behavioral regulation when smartphone ownership is delayed (Section 7). This paper provides causal estimates consistent with these reports.

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. 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.[6] Roughly one in five children is now affected. Recorded rates of anxiety and depression have risen in parallel, as have referrals to children and young people's mental health services. The steepest increases are 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.[7] 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. In dozens of countries, the young now report the lowest well-being of any age group. 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. This coincidence motivates both the research on its causes and the policy responses discussed below.

Smartphone ownership among children has risen over the same period. By autumn 2024, four in ten UK children aged 8–9 owned their own smartphone. Ownership rose to 58 percent at age 10 and 89 percent at age 11, the final year of primary school.[8] Many parents buy the first smartphone around the transition to secondary school, citing safety and coordination. By then, ownership is nearly universal. This age profile is central to the empirical design. An ownership-delay pledge can bind only for children who have not yet acquired a device.

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

[7] 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." The rate at ages 10–14 was 124 per 100,000 in 2011/12 and 307 in 2021/22. Suicide: Office for National Statistics, "Suicides in the UK: 2018 registrations." The rate among females aged 10–24 rose by 83 percent ($3.3/1.8 = 1.83$), from 1.8 per 100,000 in 2012 to 3.3 in 2018. The 2018 rate was 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.

[8] 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. The single-year-of-age figures are my calculations from the respondent-level data.

An effect of delayed ownership should therefore concentrate among younger children. The pledge delayed their first smartphone. Older children already owned one when the initiative started.

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. Her daughter's classmates had begun to get their own. She wrote that she felt trapped by the prevailing norm. She was looking for other parents who would break it with her by delaying together. She and Clare Fernyhough created a WhatsApp group in response. The group reached its 1,023-member capacity within 24 hours.[9] The speed of the response indicates widespread parental concern before SFC existed. Some parents who later signed the pact had likely already resolved to delay a smartphone. The signup measure therefore 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, then founded Smartphone Free Childhood as an organization. In September 2024, the organization launched a "Parent Pact" tool. Through the pact, parents pledge not to give their child a smartphone before age 14. They also pledge to delay social media until age 16.[10] 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. This pattern corroborates the treatment date. Searches for a "safe phone" for children emerge a year earlier. This timing indicates

[9] 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*.

[10] The second element of the pledge is unlikely to explain the results in this paper. Most of the children studied are too young for the social-media age to bind (less than 14). The effects concentrate among 5–11 year-olds (Section 5). The effect is much smaller for 12–15 year-olds, the ages closest to that age limit.

rising awareness in the year before SFC launched.[11]

[Figure 1 about here.]

A series of prominent town-wide pledges accompanied the movement. The first was in St Albans in May 2024, where the city's primary headteachers jointly urged parents to delay. Appendix Table 9 lists the major initiatives.

Parents' motives for signing matter for interpreting the signup measure. A parent who wants to delay faces a coordination problem. Withholding a smartphone unilaterally risks isolating the child from peers whose social life increasingly takes place on their phones. 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 assure other families in the same year group at the school. The specificity of the school-year assurance is something that a general community group cannot provide.[12] High coverage at a school tends to depend less on formal school policy than on local parents. Coverage depends on at least one motivated parent champion and on supportive local group dynamics. School support cannot substitute for this process.[13] Campaigns can stall when the organizing parent is seen as judging other parents. Teachers are often reluctant to tell uninterested parents what to do. Many campaigns never reach high coverage. Adoption is therefore largely dependent on parental engagement and highly uneven across schools, even within a single town. The empirical design uses this school-level variation (Section 3.1).

2.3 Special Educational Needs and the SEMH Category

The Children and Families Act 2014 governs special educational needs (SEN) provision in England (Children and Families Act, 2014). The statutory SEND Code of Practice also governs this provision (Department for Education & Department of Health, 2015). My

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

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

[13] For example, see Appendix Table 17 for variation in signup across St Albans schools, the first prominent town-wide pledge.

school outcome is the social, emotional and mental health (SEMH) category, one of the Code's four broad areas of need. The category covers anxiety, depression, conduct and attachment difficulties, self-harm, and ADHD. Support operates at two tiers. First, most identified pupils receive *SEN support*, arranged by the school's Special Educational Needs Co-ordinator (SENCo) with no external sign-off. Second, children with greater needs may receive an *Education, Health and Care (EHC) plan*. Only the local authority can issue an EHC plan, after a statutory and clinically evidenced assessment. The plan binds the council and school to specified provision (Department for Education & Department of Health, 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 (Latimer et al., 2025). 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.

SEMH is not a clinical diagnosis. The Department for Education describes SEMH as "a descriptor of a range of challenges used within educational contexts" (Antalek et al., 2025). The Department also notes that SEMH needs "are identified largely through educational assessment rather than formal medical diagnosis." The category therefore records a school's judgment that a pupil's difficulties impede learning. This judgment coincides only partly with underlying disorder. Among children whose parents reported special educational needs in 2021, half had needs in the SEMH area (NHS Digital, 2022). Conversely, only a third of children who met diagnostic criteria for a mental disorder had parent-reported special educational needs of any kind (NHS Digital, 2018). The share ranged from a quarter among children with an emotional disorder to three-fifths among those with a hyperactivity disorder. Identified need therefore corresponds more closely to externalizing conditions, which disrupt a classroom, than to internalizing ones.

The threshold for recording SEMH varies across schools. Identification is a school-level decision. Hutchinson (2021) and Hutchinson et al. (2025) compare children with similar

attainment and background. Differences between schools account for roughly two-thirds of the variation in whether such children are identified with SEN. Local authorities account for far less of it. SEMH is among the most discretionary categories. Schools record girls less often than boys, despite comparable underlying rates of disorder. Among children meeting diagnostic criteria for a disorder, parents reported special educational needs for 48 percent of boys (NHS Digital, 2018). The share for girls was 21 percent. School-level discretion makes the *level* of recorded SEMH difficult to compare across schools. Within-school comparisons are still valid under one assumption. School-level SEMH policy must not change in treated schools relative to control schools in 2024. Section 4 discusses the estimators' handling of school-level variation in SEMH identification.

2.4 Child Disability Benefits and the DLA System

Child Disability Living Allowance (DLA) is the main cash benefit for children with disabilities. It covers children under 16 in England and Wales. It is tax-free, non-contributory, not means-tested, and worth about £5,700 a year on average (Department for Work and Pensions, 2025a; Latimer et al., 2025). A Department for Work and Pensions (DWP) decision-maker decides each award on a paper-based assessment of the child's care and supervision needs. The decision-maker follows departmental medical guidance (Department for Work and Pensions, 2025b). The assessment usually involves no examination. The decision is based on the parent-completed claim form and supporting evidence. For the psychiatric grouping the relevant evidence is clinical and developmental. The DWP lists only local-authority educational documents, either a statement of SEN or an EHC plan (Department for Work and Pensions, 2025c).[14] 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. In 80 percent of cases, DLA starts no later than the EHCP. The

[14] A statement of SEN was the legal predecessor to the EHC plan. The local authority issued it after a statutory assessment. The 2014 SEND reforms replaced statements with EHC plans. Both are local-authority statutory documents, distinct from the school-decided, non-statutory SEN-support tier described in Section 2.3.

EHCP therefore usually does not precede DLA. Both forms of support are persistent. For example, of children starting an EHCP at age 5, 84 percent still hold it at age 15. For child DLA, the share is about 75 percent.[15] The rules are identical across England and Wales. Eligibility therefore does not vary across local areas, unlike SEN and EHC practice.

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 (Latimer et al., 2025). The growth concentrated in learning difficulties (which record autism), behavioral disorder, and ADHD. Behavioral disorder has grown by far the fastest. At primary ages its claims rose fivefold between 2018 and 2025, while ADHD claims grew more slowly than the caseload as a whole.[16] My DWP outcome focuses on claims with a psychiatric disabling condition. Section 3 sets out the construction and composition of the outcome.

2.5 What SEMH Needs and DLA Claims Measure

Survey data link both outcomes to severe behavioral symptoms. In Understanding Society, parents in England report whether their child has a special educational need, and why (University of Essex, Institute for Social and Economic Research, 2026). I call a need SEMH-type when the parent names ADHD, behavioral problems or hyperactivity, or mental illness or depression. Psychiatric-type DLA requires a DLA report by an adult in the household and a psychiatric-type reason for the child's need. Parents of children aged 5 to 9 also complete

[15] Darcey Snape (Institute for Fiscal Studies) calculated the Longitudinal Education Outcomes (LEO) timeline summaries in the ONS Secure Research Service. The summaries were 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. First, an EHCP marks the formal statutory tier of SEN provision. Families reach it through a lengthier process for more complex needs. Most children who eventually receive an EHCP are registered for school SEN support well beforehand. These timelines therefore compare DLA with the statutory document only. Second, the cohorts largely predate the steep recent rise in SEN and child-DLA rates.

[16] DWP Stat-Xplore, national child DLA counts by main disabling condition. At ages 5–11, behavioral-disorder claims rose from 32,216 in November 2018 to 160,549 in November 2025. Over the same period, learning-difficulty claims rose from 108,885 to 202,146. Hyperkinetic/ADHD claims rose from 23,505 to 33,132.

the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997).[17] Children with an SEMH-type need average 19.9 total difficulties, against 8.2 for children with no reported need (Appendix Table 14). The adjusted gap is 11.3 points (95 percent confidence interval 9.9 to 12.8). It equals 2.2 standard deviations of the score among children with no reported need. Children with psychiatric-type DLA show a similar gap, 10.7 points (8.2 to 13.2). In both groups, the gap is largest on the hyperactivity scale (Appendix Figure 6). The survey identifies more children with an SEMH-type need than the school census does, but fewer with psychiatric DLA than DWP records show (Appendix Table 16). In the survey, 8.7 percent of children aged 5 to 10 have an SEMH-type need, against 2.9 percent of pupils in the census. Most of the difference comes from children whose parents also name autism or a learning difficulty. Without these children, the survey share falls to 3.9 percent, closer to the census rate. Their symptom gap is similar, at 9.6 points (7.6 to 11.7). For DLA, the survey records psychiatric-type claims for 2.5 percent of children, about half the 5.1 percent in DWP records ($2.5/5.1 = 0.49$). The survey misses claims that parents do not report or that cannot be linked to a particular child.

Both measures rarely capture children whose difficulties are only emotional. Among children with an abnormal total difficulties score, 60.9 percent have any parent-reported need (Appendix Table 15). Only 41.7 percent have an SEMH-type need. About 70 children in the sample have abnormal scores for emotional symptoms, with conduct and hyperactivity scores below the abnormal range. Fewer than 10 of these children have an SEMH-type need, and none has psychiatric-type DLA. Both outcomes therefore track behavioral difficulties, such as conduct problems and hyperactivity, more closely than emotional difficulties such as anxiety and depression.

[17] Understanding Society waves 13 and 15, with interviews from January 2021 to May 2025. The parent-report sample holds 1,290 children aged 5 to 9 in England, rounded to the nearest 10. All estimates are weighted. The gaps adjust for age and sex. Appendix A.7 describes the data, the definitions, and the disclosure rules.

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. It also publishes the school's signup share, the ratio of signed families to enrollment. The data are available through a public interface keyed to the DfE Unique Reference Number.[18]

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. Delaying families therefore 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.[19] Signup counts therefore understate how far behavior actually shifted.

I classify a school as *treated* if its signup share is at least 20 percent. I classify it as *low-uptake* if its share is below 5 percent. Both shares are measured in June 2026. The 20 percent threshold is an ex ante choice based on Cunningham Hill Junior School. Registered uptake at Cunningham Hill stood near 40 percent. Reported Year 6 ownership fell almost to zero. The tipping point for community-wide compliance therefore lies below 40 percent. I set the threshold at half of that level. The results are similar at other thresholds.[20] 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

[18]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.

[19]The fall to 12 percent is the 2024 Year 6 survey result reported in Gentleman (2025), as cited in 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.

[20]The effect is already present at the 10 percent threshold (Table 4). It rises with uptake. Appendix Table 21 reports results for other control-pool cutoffs.

group. They are spread across 118 local authority districts, with at most 14 in any one. Approximately 12,000 schools have negligible uptake. They form the low-uptake pool from which I draw matched controls.[21] Most registered schools have low uptake, with a median share of 2 percent among schools with any pacts. The high-uptake group is moderately concentrated. The ten most active districts hold three in ten treated schools. Appendix Table 10 gives the full distribution. Appendix A tabulates the sample construction from the 15,653 registered schools to the analysis samples. Uptake varies widely even within the pledge towns (Appendix Tables 17 and 18). It ranges from around 40 percent of families at the most engaged schools to near zero at others. The design uses this school-level variation. The estimates are similar without the three pledge towns with the most media attention (Appendix Table 21). Schools between 5 and 20 percent are held out of the binary comparison. They enter only the dose-response specifications. These specifications estimate the effect at uptake thresholds of 10, 20, 30, and 40 percent.

Table 1 compares the two groups on pre-treatment observables. It also reports the two control sets used by the estimators. The matched sets nearly equal the treated schools on outcomes and affluence. The covariate-only donor sets equal them on affluence alone. 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 lower SEMH (28.4 versus 33.4 per 1,000). They also begin with markedly lower child DLA claiming (36.6 versus 60.5 per 1,000). Mean enrollment is balanced. The estimators in Section 4 account for these differences.

[Table 1 about here.]

[21] To 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. The low-uptake control pool is therefore unlikely to contain much unmeasured coordinated delay.

3.2 School Outcome: SEMH Needs

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

I assemble a panel of approximately 16,700 primary schools. Estimation uses the balanced subpanel of 14,260 schools with a complete SEMH series in the January censuses from 2019 to 2026. Group composition is therefore 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. These 329 schools form the matched estimation sample (sample selection in Appendix A). I begin in 2019 because the DfE suppressed small SEMH counts (below five) in earlier censuses. The DfE stopped this suppression 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). The autistic-spectrum category is recorded by the same school staff in the same census. A smartphone pledge is unlikely to affect it. I use it 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 SEMH identification is a school decision. Schools with engaged parents may record need differently. I therefore add child DLA claims as a second outcome. The DLA data also provide independent placebo outcomes unaffected by a phone delay. The DLA decision-maker is a DWP caseworker outside the school. The caseworker applies eligibility rules common across England and Wales (Section 2.4). A school cannot reclassify a DWP benefit. A similar fall in DLA claims would therefore corroborate the school result through

an independent decision-maker.

The DWP publishes quarterly counts of children receiving DLA with a recorded psychiatric main condition. The counts are split by disability component, age band, and sex. They are available for each Lower-layer Super Output Area (LSOA), a neighborhood unit of roughly 1,500 residents. Each count is the stock of current claimants at the snapshot date, not the flow of new benefits. At ages 5–15, about 60 percent of the psychiatric caseload is "learning difficulties," a developmental classification including autism. Behavioral disorder accounts for just under 30 percent and hyperkinetic/ADHD for about 10 percent. Psychoneurosis (the anxiety-and-depression category) and psychosis are effectively zero.[22] The ideal outcome would be behavioral claims among 5–11 year-olds. This outcome would target the age range and the emotional and behavioral conditions affected by a phone delay.

The DWP releases these counts through its Stat-Xplore portal with a short publication lag. Two privacy protections apply. First, any cell with fewer than five claims is replaced with a zero. Second, 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 the data to catchments raises most child psychiatric counts above the suppression threshold. I retrieve each catchment count from Stat-Xplore as a single custom-geography total (Appendix A). Each count is therefore perturbed only once. The catchments are built around all schools, whether or not they appear in the census panel. The DWP and school outcomes therefore use slightly different treated sets.[23] Second, I average the four quarterly snapshots within each calendar year. The stock of claimants changes little from one quarter to the next. Averaging therefore removes noise. It also aligns the series with the annual school census. Third, I use the 5–15 psychiatric

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

[23] The nearest-school catchment is an approximation. Families often attend a school other than the nearest one. This mismatch attenuates 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.

aggregate as the main outcome, keeping the estimation cells large. Sub-group cells at the catchment level fall below the suppression threshold. For heterogeneity, I therefore download the treated and control counts pooled over each matched set of catchments. Pooled counts are free of suppression at any split. Section 5.2 uses them to decompose the total effect by component,[24] age, sex, and claim duration. This decomposition identifies the groups that account for the main result.

After aggregation, zero cells are close to balanced between treated and matched control catchments. For the 5–15 psychiatric total, 8.9 percent of treated catchment-quarters are zero, against 7.7 percent of controls.[25] A cross-check against the near-unsuppressed MSOA marginals indicates most recorded zeros are close to an actual zero.[26] Imputing every zero with four claimants does not meaningfully change the results (Appendix Table 21).

I use two placebo outcomes that are decided outside the school. They complement the autistic-spectrum category recorded in the school census (Section 2.3). 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. Its conditions are structural and 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 that the effect is not specific to behavioral conditions. A fall in adult depression would indicate an area-level

[24] Appendix A.6 maps the correspondence between the DLA conditions and the school SEN categories.

[25] See Appendix Table 12 for the suppression rate of each cell. The ideal young behavioral cell is suppressed for over 70 percent of treated catchment-quarter observations at the catchment level. The pooled counts of Section 5.2 recover it without suppression.

[26] The MSOA marginals are a reliable benchmark. Only 0.1 percent of their cells are recorded as zero. I sum the LSOA counts in their native (2021) geography, leaving zero cells at zero. The resulting national caseload is 264,700 in May 2018 and 665,700 in November 2025. The MSOA benchmark is 262,600 and 663,900. The zero-fill therefore matches the benchmark to within 1 percent ($264{,}700/262{,}600 = 1.008$). The few cells with suppressed nonzero counts contain few claims. DWP's cell-key perturbation of the published marginals further offsets their omission. Filling every zero cell at four instead raises the May 2018 total to 293,300, about 12 percent above the benchmark ($293{,}300/262{,}600 = 1.117$). Figure 2 plots the England-wide catchment series. Its 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.

confound.

Panel B of Figure 2 plots the national series of the seven DLA outcomes. Panel A plots the school series. 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, when the SFC movement began (Section 2). Observations are labeled by their measurement date (Appendix A). The January 2024 census is therefore the last pre-treatment school observation. The January 2025 and 2026 censuses are the two post-treatment observations. The quarterly DLA series contributes eight post-treatment quarters, February 2024 through November 2025. Averaged to calendar years for estimation (Section 3), they form the 2024 and 2025 post-treatment years.

I compare high-uptake schools with low-uptake schools, as defined in Section 3.1. High- and low-uptake schools differ in level. They were also converging before the pact. My primary estimator is therefore a matched difference-in-differences (Dustmann et al., 2017). I match each treated unit to three low-uptake controls by nearest-neighbor matching on the Mahalanobis distance of the covariates, without replacement. The covariates are the same for every outcome. They are the outcome's own levels in each pre-treatment year from 2021 to 2023, neighborhood affluence, the free-school-meal share, and log size. Neighborhood affluence is the 2019 Index of Multiple Deprivation (IMD-2019).[27] On the matched sample

[27]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. Catchment size is the resident child population. IMD-2019 is England's official small-area deprivation index (McLennan et al., 2019). It combines seven domains with fixed weights. Income and employment each carry 22.5 percent of the weight. Education, skills and training, and health and disability each carry 13.5 percent. Crime, barriers to housing and services, and living environment each carry 9.3 percent. 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 assigns neighborhoods to catchments (Appendix A).

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. I report effects for each period and for the pooled post-treatment window. I express each effect as a percentage of the counterfactual level implied by the estimated fixed effects. This base is the treated units' predicted level under the control group's 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. First, I estimate synthetic difference-in-differences (Arkhangelsky et al., 2021). This estimator reweights the low-uptake donor pool so that the weighted controls track the treated schools' pre-treatment path. Its donor pool is the covariate-only match of the triple difference below. No outcome levels enter this match. The weights alone fit the pre-treatment trajectory. The estimator, its weights, and the numeric estimates are set out in Appendix B. Second, I estimate a triple difference that does not require matching on mental-health outcomes. It compares each mental-health outcome with a placebo outcome in the same unit. In schools, it compares SEMH with the autistic-spectrum category. In catchments, it compares the behavioral with the developmental DLA component. Unit-by-year fixed effects absorb shocks common to both categories (Section 6.1).

Treatment is spatially clustered. I therefore cluster standard errors by local authority district. Of these districts, 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 in unobserved trends as well as 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. First, treated communities could already be on a different trajectory. A continuing trajectory would look like an effect after treatment. Second, matching on the years just before treatment can select units at a temporarily high or low level. The level then reverts toward the mean. Daw and Hatfield (2018) describe this regression-to-the-mean bias. Four checks address these problems. First, the matched series track each other closely through the whole pre-period (Section 5). This period includes 2019 and 2020, which the match never uses. Controls selected on a transiently high or low 2021–2023 level would differ from the treated group in 2019 and 2020. They would then converge in the matching window. No such gap appears. Second, a placebo treatment date inside the pre-period shows no effect (Table 7). This placebo is the direct test of a continuing trend. Third, the placebo-in-space distribution bounds the bias that matching could produce under the null (Section 6.3). The placebo mean is less than a twentieth of the real estimate. It also has the opposite sign. Fourth, the triple difference compares the outcome with a placebo category inside the same school or catchment. Unit-by-year fixed effects absorb any trend common to the two categories (Section 6.1). Its pre-trends are flat without any matching on mental-health outcomes. Its estimates are similar to the main difference-in-differences estimates.

Within-district variation. With local-authority-by-year fixed effects, the comparison is within district and year. These fixed effects absorb every district-level time-varying confound. Such confounds are the most likely source of area-level selection. The effect is similar

with these fixed effects. The same comparison also tests for spillovers. The delay norm could spread from high-uptake schools to nearby low-uptake schools. The within-district comparison group would then be partly treated. 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. 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. Three facts limit this threat. 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 40 percent across schools. Uptake therefore does not measure school policy (Section 3.1). Second, a change in a school's recording practice would affect the developmental categories as well as SEMH. The developmental categories do not change (Section 6). Third, the DWP outcome is decided outside the school. It falls together with the school result (Section 5).

Placebo outcomes and the age pattern. A smartphone effect makes three testable predictions. First, it should reduce SEMH and child psychiatric claims. Second, it should not affect conditions unrelated to a phone. I test this prediction on the placebo outcomes of Section 3. They are the autistic-spectrum school category, the developmental DLA component, and recorded adult depression (QOF, 18+) in the same areas. Third, it should be larger among younger children than among older children (Section 5.3). The pact still postpones the first smartphone of younger children. Most older children already own a device.

5 Results

My primary estimator is the matched two-way fixed-effects event study of Section 4. It compares high-uptake schools with matched low-uptake controls. The match uses the outcome's own 2021–2023 levels, the IMD-2019, the free-school-meal share, and log size. Years before 2021 serve as a pre-trend check outside the matching window.

5.1 Effects on Mental Health

The school and DWP outcomes are complementary. The school census covers every school without suppression. It also spans several categories of need. It does not report children's ages. The school itself records the census data. The DLA benefit is determined outside the school. The DLA data report age and sex. They suppress small counts 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 double between the January 2019 and January 2024 censuses. National SEMH identification rose over the same period. After February 2024, the treated series falls below the control series. The differenced coefficients are flat and near zero in the pre-period. A statistically significant difference appears after the start of the movement. The average post-treatment SEMH reduction is 2.9 per 1,000 pupils (Table 2). This reduction is about 8 percent of the counterfactual. The counterfactual is the treated schools' predicted level on the control time path.The SEMH reduction is 2.4 per 1,000 in the first post-treatment census (2025). It rises to 3.4 in the second (2026). This increase is consistent with uptake growing over time.[28]

With district-by-year fixed effects, the reduction is at least as large (Table 2, column 2).

[28]My measure of uptake is a cumulative June 2026 snapshot. It dates partially treated schools as treated from the outset. The pooled estimate is therefore likely a lower bound.

These fixed effects absorb district-level trends. The pre-trend remains flat. The reduction therefore holds within districts, between high- and low-uptake schools in the same local authority and year. It does not arise from comparing affluent areas with poorer ones.

[Figure 3 about here.]

[Table 2 about here.]

DWP outcome. The second outcome, child psychiatric DLA claims, is decided outside the school (Figure 3). The SEMH measure is a school classification. School staff record it. Their engagement may also facilitate the pact's uptake. To test whether the reduction reflects children's well-being 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. I then run the same matched difference-in-differences on high-uptake and matched low-uptake neighborhoods (Figure 3, bottom row, and Table 2, columns 3–4).

In high-uptake areas, child psychiatric claims fall by 3.0 per 1,000 children. This reduction is about 6 percent of the counterfactual. The pre-trend is flat. The reduction is about 4 percent in the first post-treatment year. It rises to about 8 percent in the second. A school cannot reclassify a DWP benefit. The DLA result therefore corroborates the school result through an independent measurement system. The two reductions are of similar magnitude.

The school and DWP results could be linked through the EHC plan. The EHC plan is the single school-related document that the DWP includes in its decision (Section 2.4). The local authority issues EHC plans. They account for about 12 percent of the primary-school SEMH count.[29] I re-estimate the main SEMH event study on the school-decided SEN-support tier alone. The reduction on this tier agrees in sign with the full measure.[30]

[29] January 2024 census, state-funded primaries: 127,260 pupils on SEN support and 18,060 with an EHC plan under the SEMH primary need. The EHC tier is therefore 12.4 percent of the SEMH count ($18{,}060/(127{,}260 + 18{,}060) = 0.124$).

[30] The reduction is 6.4 percent of the counterfactual, against 8.2 percent for the full measure. Within districts, it is 9.9 percent. The pre-trends are flat.

The similarity of the school and DWP results therefore does not depend on the shared EHC document.

5.2 Decomposing the DLA Effect

With a fixed population, a difference-in-differences in rates is linear in the counts. The total effect on the caseload therefore equals the sum of the effects on its disjoint parts. This subsection uses that property to decompose the DLA result of Table 2. I re-estimate the effect on pooled counts, summing claims across the 330 treated catchments and, separately, across their 990 matched controls. The pooled counts are free of small-cell suppression. The pooled data allows me to decompose the results along a number of dimensions even if the cells are small: condition, sex, age, and claim duration. The estimate compares the average of the four 2023 quarters with the average of the four 2025 quarters. These are the last full pre-treatment year and the second post-treatment year. Each pool's 2023 child population is the fixed denominator, which keeps the decomposition additive. The percent effect divides the change in the rate by the control post-period rate. The two pools have closely matched pre-period rates (35.5 against 37.2 claims per 1,000). This base therefore approximates the treated counterfactual, the same base as in the main estimates. Standard errors in this subsection are approximate Poisson errors computed from the counts alone (see the notes to Table 3). The pooled total effect is −3.38 claims per 1,000 children (Table 3). This reduction of 6.1 percent is consistent with the matched TWFE estimate.

Table 3 shows the decomposition of the total effect. Behavioral claims account for 71 percent of the reduction, a within-group fall of 10.1 percent. Behavioral claims at ages 5 to 11 alone account for 56 percent of the reduction. Smartphone ownership is near-universal from age 11 (Section 2), so a delay of ownership binds mainly at primary ages. In line with this age pattern of exposure, behavioral claims fall by 12.9 percent at ages 5 to 11. I find no evidence of an effect on behavioral claims at ages 12 to 15. Among behavioral claims at ages 5 to 11, behavioral disorder falls by 15.1 percent. I find no evidence of an

effect on hyperkinetic syndrome (ADHD), about a quarter of this behavioral caseload. The split separates a mental-health category (behavioral disorder) from a neurodiversity category (ADHD). The largest within-group effect is a 21.5 percent fall in new behavioral-disorder claims at ages 5 to 11. New claims have a duration under twelve months, so all of them began after the start of the movement. I find no evidence of an effect on developmental claims, either in total or among new claims. The reduction in new claims is therefore specific to the behavioral component.

Claim duration separates prevented claims from claims that end early. Within behavioral disorder at ages 5 to 11, new claims, those under twelve months old, fall by 21.5 percent (standard error 5.5). Claims held for one to two years fall by 16.3 percent (5.2). I find no evidence of an effect on claims held for two years or more. In the 2025 quarters, every claim held for two years or more began before the start of the movement. The pact could reduce these claims only if children already on DLA improved and left the benefit. The pattern across durations therefore points to fewer new claims, not to earlier exits from existing ones.

Girls claim psychiatric DLA at less than half the rate of boys, 10.9 against 26.3 per 1,000 among matched controls in 2023. The girls' estimates are therefore less precise. Behavioral-disorder claims at ages 5 to 11 fall by 18.1 percent for boys (standard error 3.7) and by 9.1 percent for girls (5.4). The difference between the two estimates is within sampling error. Boys account for 42 percent of the total reduction and girls for 10 percent. A fall in new behavioral claims among young children is the pattern expected from a delay of smartphone ownership.

[Table 3 about here.]

5.3 Dose-Response and Affluence Heterogeneity

This section splits the effects by local adoption and by local affluence. The dose-response tests whether the effect scales with community uptake. The affluence splits compare the reductions across richer and poorer adopting communities.

Pact uptake. The effect strengthens with pact uptake (Table 4). The SEMH reduction is about 3 percent where a tenth signed. It is 8 percent at the one-fifth threshold. The DWP outcome shows the same pattern (Table 4, Panel B). Its reduction is about 4 percent where a tenth signed. It is 6 percent at the one-fifth threshold. The estimates at the two highest thresholds (30 and 40 percent) are consistent with this increase. They are imprecise because few schools reach these uptake levels. An effect that scales with local adoption is consistent with a community-level treatment.

[Table 4 about here.]

Affluence. Affluence splits are reported in Appendix Table 20. SEMH falls by 14 percent in the less affluent treated schools. I find no evidence of an effect in the more affluent treated schools. The DLA reduction is larger in the more affluent catchments (-8 against -4 percent). The smaller cells have wide confidence intervals. The opposite gradients are puzzling. I do not have an explanation for them. Section 7 discusses the heterogeneity of the effects, including how they depend on local affluence.

6 Robustness and Falsification

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

6.1 Alternative Estimators

The main estimates are robust to alternative estimators (Table 5). The estimators match treated and control units in different ways. 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. Each handles differential pre-treatment trends in its own way. Synthetic difference-in-differences (Arkhangelsky et al., 2021) uses the pre-treatment mental-health outcomes through reweighting. 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. Absent the pact, the weighted controls would then 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. In schools, it compares SEMH with the autistic-spectrum category. In catchments, it compares the behavioral with the developmental DLA component. Unit-by-year fixed effects absorb any shock or trend common to a unit's two categories. Identification therefore requires no matching on outcomes.

I compare the estimators in claims or pupils per 1,000 because their percent bases differ. The level designs measure the effect against the full psychiatric counterfactual. The triple difference measures the behavioral–developmental gap against the smaller behavioral base. The percent figures are therefore not comparable across estimators. Per 1,000 pupils, the SEMH reduction is -2.9 under the matched TWFE. It is -1.7 under SDID and -2.0 under the triple difference. Per 1,000 children, the DLA reduction is -3.0 under the matched TWFE. It is -2.4 under SDID and -1.7 under the triple difference. 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 is in behavioral conditions, the two should imply a similar reduction in claims. All three estimators agree in sign. The triple difference identifies the effect without using pre-treatment outcomes. Figure 4 plots the event studies under all three.

[Table 5 about here.]

[Figure 4 about here.]

6.2 Placebo Tests

I subject the result to placebo tests in outcomes and in time. Section 6.3 reports a third placebo, in space.

Placebo outcomes. Table 6 compares the two mental-health outcomes with two placebo outcomes. The mental-health outcomes are school SEMH and the behavioral DLA component. The placebo outcomes are the autistic-spectrum school category and the developmental DLA component. Each placebo comes from the same data system as its outcome. The DLA rows partition the caseload into disjoint components. Their sampling errors are therefore independent, unlike those of a total-versus-component comparison. Under both the matched and the district-by-year specifications, I find no evidence of an effect on either placebo. The moving outcomes fall by 8 to 10 percent. Autism is the natural school placebo. The same staff identify it in the same census as SEMH. 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, school relabeling alone cannot explain the triple difference results.

[Table 6 about here.]

Placebo in time. A placebo in time is an important falsification test for two of the designs. Both the baseline design and SDID use pre-treatment mental-health outcomes. The test checks whether treated communities diverged from their matched controls before the pact. Matching on pre-treatment outcomes aligns the two groups at the match date. It does not make their trends parallel. It can even induce a spurious trend through regression to the mean (Daw and Hatfield, 2018).

I re-date the treatment to February 2022. I then re-estimate on pre-treatment data alone. The placebo uses only years before the fake date for matching. The placebo effect is

measured on the last two pre-treatment years (2022 and 2023). For school SEMH, I find no evidence of a placebo effect under either matched TWFE or SDID (Table 7). The estimates at the true date are −2.9 and −1.7 per 1,000. For the DWP outcome, the matched TWFE placebo is marginally negative, at −1.2 per 1,000. Its 95 percent interval just includes zero. I find no evidence of a placebo effect under SDID. The true-date reductions are −3.0 and −2.4. The matched TWFE placebo may still reflect a small differential trend. With treatment assigned to 2022, the matched TWFE and SDID use the same three pre-treatment outcomes. Their placebo-in-time estimates are therefore similar.

The triple difference uses no pre-treatment outcomes. It compares the behavioral and developmental components within the same catchment. Any differential trend common to the two components is differenced out. At the true date, the estimate is −1.7 per 1,000. The drop is 9.4 percent of behavioral claims. At the fake date, I find no evidence of an effect. The school triple difference against autism shows the same pattern. It is −2.0 per 1,000 at the true date. At the fake date, I again find no evidence of an effect. The level-design placebos suggest a small differential trend. This trend does not explain the DLA result.

[Table 7 about here.]

6.3 Inference: A Calibrated Placebo-in-Space Permutation

Treated units cluster within districts. Analytic inference alone may therefore overstate 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 and Taber, 2011; Ferman and Pinto, 2019; MacKinnon and Webb, 2020). Each replication draws a placebo "treated" set of never-treated (<5 percent) schools. The set reproduces the real treated schools' count and per-district distribution. Each replication re-matches the set to the remaining low-uptake units on the same covariates. It then recomputes the matched difference-in-differences. Each placebo set is matched and estimated exactly as the real sample. The exercise therefore also bounds regression to the

mean. Any bias from matching on pre-treatment levels under the null appears in the placebo distribution. A real estimate in the tail of this distribution exceeds that bias. For the DWP outcome, the drawn schools define the placebo catchments. The DLA panel is built on school catchments. The school and DWP outcomes therefore 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. The re-centered comparison 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 lie in 118 districts, with a mean nearest-neighbor distance of 5.3 km. Each replication assigns the real per-district counts to a fresh sample of never-treated districts. The draws match both the number of treated districts and the within-district concentration. The placebo nearest-neighbor distance is 5.5 km. The DLA permutation applies the same school-level draw on the catchment panel. It assigns the real per-district counts of the 330 treated catchments to fresh never-treated districts. It draws placebo catchments from the full national pool of <5 percent catchments. Each draw is re-matched.

Both main estimates lie in the left tail of their placebo distributions (Figure 5). The DLA estimate is below all 200 placebo estimates. The three comparisons give $p \leq 0.005$ (panel B). For SEMH, the one-sided and mean-centered comparisons give $p = 0.010$ (panel A). The uncentered two-sided comparison of equation (2) gives $p = 0.015$. The placebo distributions are centered just above zero. Their means are +0.11 per 1,000 for SEMH and +0.60 for DLA. Any regression to the mean from the matching procedure therefore points toward positive estimates. It cannot account for the estimated reductions.

[Figure 5 about here.]

6.4 Specification Robustness

Both main effects are stable across specifications (Appendix Table 21). The specifications vary the low-uptake threshold (<5 versus <10 percent), the match ratio (1:3 versus 1:5), and replacement in matching. Both main reductions stay near 3 per 1,000 across the rows. Extending the SEMH pre-period back to 2015 restricts the panel to schools with complete early data. This restriction more than doubles the standard error. With this restriction, the estimate is too imprecise to provide evidence of an effect. Suppressed catchment-quarters, recorded as zero, have a true count of 0 to 4 under the disclosure rules. I impute each at its upper bound of 4. The main estimate changes from −3.39 to −3.15 per 1,000. The main catchment cell is rarely at the suppression threshold. Its zeros are nearly balanced between treated and control catchments. The upper bound therefore changes the estimate little. I also estimate the DLA event study on the raw quarterly claim snapshots (Figure 8). The pooled reduction is the same, −3.4 per 1,000. The quarterly estimates show weak evidence of anticipation in the quarters just before February 2024.[31] Dropping the three pledge towns with the most media attention leaves every estimate essentially unchanged. The effects therefore come from high-uptake communities across England.

7 Discussion

The evidence is consistent with coordinated ownership delay improving children's mental health. Two independent measurement systems record falls of similar size (8.2 and 6.0 percent of the contemporaneous counterfactual). Each fall grows across the post-treatment window. The effect strengthens with uptake. It is similar with district-by-year fixed effects. It is absent in the placebo outcomes, the other special-educational-need categories in the

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

same children and recorded depression among adults in the same communities.

The psychological literature proposes four mechanisms for the effect (Haidt, 2024). First, smartphones displace sleep. Second, they fragment attention through continual interruption. Third, algorithm design trains compulsive checking. Fourth, smartphones crowd out the independent play in which children learn to regulate themselves. Causal economic studies find harmful effects of social media on adults (Braghieri et al., 2022; Allcott et al., 2020). These mechanisms match the outcomes that fall in this paper. The SEMH category explicitly includes attention-deficit disorders alongside anxiety and attachment difficulties. The DLA category that falls is behavioral disorder, the conduct and dysregulation diagnosis. Teachers reported these behavioral changes when children first received smartphones (Haidt, 2024). They report the reverse changes when ownership is delayed today. Justine Elbourne-Cload, the headteacher of Cunningham Hill Junior School, gives an example (Section 2). The headteacher reports that once children no longer received smartphones, "we have seen less evidence of dysregulation for those children and better attention spans." The headteacher now counts "three, four children out of about four hundred that get very dysregulated, compared to a much higher number" before the change.[32] This paper provides causal evidence for the changes that teachers and psychologists describe.

The estimates are relevant to current fiscal debates. 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. A back-of-the-envelope calculation translates the estimates into the number of children helped and money saved. Applied nationally, the SEMH estimate implies about 13,000 fewer pupils with SEMH need, and the DLA estimate about 22,500 fewer claims each year. Because the same child can leave both registers, the number of distinct children lies between about 22,000 and 36,000.[33] At IFS unit costs, the savings come to about

[32]Interview with the author, July 2026.

[33]Cases: the matched TWFE effects at the 20 percent threshold are −2.9 SEMH pupils and −3.0 claims per 1,000 (Table 4). Locally, they apply to 92,675 pupils in the 329 high-uptake schools and 153,817 children aged 5 to 15 in the 330 treated catchments. Nationally, they apply to 4.55 million state primary pupils and

£200 million per year: roughly £80 million in notional SEN budgets (13,000 × £6,000) and £130 million in DLA awards (22,500 × £5,700). These figures assume that the effect extends beyond the adopting communities, and that fewer recorded cases mean fewer children with difficulties rather than difficulties going unrecorded. They also cover only the first two years of a growing effect. Finally, the figures leave out several costs. They do not count spending on NHS mental-health services. They do not count the well-being of the children, their parents, or their teachers. They miss anxiety and depression, which the two outcomes rarely record (Sections 2.3 and 2.5). They also omit longer-term costs, such as lower learning and a higher chance of being not in education, employment or training (NEET) as a young adult. If the pact reduces these costs as well, the figures understate the total savings.

8 Conclusion

This paper gives quasi-experimental evidence on coordinated smartphone delay. The Smartphone Free Childhood Parent Pact coordinates families within a school to withhold smartphones until age 14. Its take-up varies widely across England's primary schools. I identify its effect from that variation. High-uptake schools are those where at least a fifth of families signed. Their SEMH need falls by about 8 percent relative to matched controls. Child psychiatric DLA claims fall about 6 percent relative to matched communities. The DWP adjudicates these claims outside the school. The reductions strengthen with uptake and over time. They concentrate among younger children and in behavioral conditions. The pact still delays these children's first smartphone. A phone effect would be expected in these conditions.

Ownership delay may be more effective when paired with complementary policies. In

7.46 million children aged 5 to 15. Overlap: behavioral conditions at ages 5 to 11 account for 56 percent of the DLA effect (Table 3). Adding the developmental component raises the share to at most 84 percent. Near-complete overlap of those claimants with SEMH pupils gives about 22,000 distinct children. Zero overlap gives about 36,000 (13,000 + 22,500 = 35,500). Costs: £6,000 per SEMH pupil per year and £5,700 per child DLA claim per year, both from Latimer et al. (2025) in 2025–26 prices. The first is the notional SEN budget that a mainstream school funds before local-authority top-up. The second is the average award, with a range of £1,500 to £9,700.

the UK, the rise in mental-health need and smartphone ownership coincided with local-authority budget cuts. The cuts closed many outdoor clubs and public spaces. Children once developed self-regulation and social skills in these places. Children with the greatest need often lack these skills. Teachers and Parent Pact organizers alike stress that removing the phone is not sufficient on its own. A school governor told 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., Baron, E., Dee, T., Duckworth, A., Gentzkow, M., and Jacob, B. (2026). The Effects of School Phone Bans: National Evidence from Lockable Pouches. NBER Working Paper 35132.

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

Antalek, C., Dixon, F., Loyd, D., Peter, S., Roberts, A., Sharxhi, E., and Somerville, M. (2025). *Identifying and Supporting Children and Young People with Social, Emotional and Mental Health Needs: A Rapid Evidence Review*. Research report, September 2025. London: Department for Education and UCL Institute of Education.

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

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

Beneito, P. and 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., Bryson, A., and 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., and Makarin, A. (2022). Social Media and Mental Health. *American Economic Review*, 112(11), 3660–3693.

Burnell, K., Meter, D., Andrade, F., Slocum, A., and George, M. (2025). The effects of social media restriction: Meta-analytic evidence from randomized controlled trials. *SSM - Mental Health*, 7:100459.

Callaway, B. and Sant'Anna, P. (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics*, 225(2), 200–230.

Children and Families Act (2014).

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

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

Davis, C. G. and Goldfield, G. S. (2025). Limiting social media use decreases depression, anxiety, and fear of missing out in youth with emotional distress: A randomized controlled trial. *Psychology of Popular Media*, 14(1):1–11.

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

Department for Education (2025). *Education, Health and Care Plans*, reporting year 2025 (SEN2 collection). Explore Education Statistics. https://explore-education-statistics.service.gov.uk/find-statistics/education-health-and-care-plans.

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

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

Department for Work and Pensions (2026). *Disability Living Allowance (DLA) for Adults*. GOV.UK. https://www.gov.uk/dla-disability-living-allowance-benefit, accessed September 2026.

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

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

Ferguson, C. J. (2025). Do social media experiments prove a link with mental health: A methodological and meta-analytic review. *Psychology of Popular Media*, 14(2):201–206.

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

Girela-Serrano, B., Spiers, A., Ruotong, L., Gangadia, S., Toledano, M., and 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.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38(5):581–586.

Goodyear, V. (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.

Hunt, M. G., Marx, R., Lipson, C., and Young, J. (2018). No more FOMO: Limiting social media decreases loneliness and depression. *Journal of Social and Clinical Psychology*, 37(10):751–768.

Hutchinson, J. (2021). *Identifying Pupils with Special Educational Needs and Disabilities*. London: Education Policy Institute (funded by the Nuffield Foundation).

Hutchinson, J., Downs, J., and Ford, T. (2025). *Identifying SEND: Final Report on Special Educational Needs & Disabilities and Contact with CAMHS*. London: Education Policy Institute (funded by the Nuffield Foundation).

Jain, R. and 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., and Tyrefors, B. (2020). The Impact of Banning Mobile Phones in Swedish Secondary Schools. *Economics of Education Review*, 77, 102009.

Latimer, E., Sibieta, L., and 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.

Liu, Y., Mohamad, E. M. W., Azlan, A. A., and Tan, Y. (2025). Am I happier without you? Social media detox and well-being: A meta-analysis of randomized controlled trials. *Behavioral Sciences*, 15(3):290.

Locher, C., Koechlin, H., Zion, S. R., Werner, C., Pine, D. S., Kirsch, I., Kessler, R. C., and Kossowsky, J. (2017). Efficacy and safety of selective serotonin reuptake inhibitors, serotonin-norepinephrine reuptake inhibitors, and placebo for common psychiatric disorders among children and adolescents: A systematic review and meta-analysis. *JAMA Psychiatry*, 74(10):1011–1020.

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

May, W., Malouff, J. M., and Meynadier, J. (2025). Reducing social media use decreases depression symptoms: A meta-analysis of randomised controlled trials. *European Journal of Investigation in Health, Psychology and Education*, 15(11):222.

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

NHS Digital (2018). *Mental Health of Children and Young People in England, 2017*. Leeds: NHS Digital. Professional services, informal support and education tables, Table 8 (special educational needs by disorder, parent report).

NHS Digital (2022). *Mental Health of Children and Young People in England, 2021: Supplementary Data Tables on Special Educational Needs and Disability*. Leeds: NHS Digital.

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., Purba, A., Fokas, I., Gunschera, L., Yap, Z., 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.

Park, E., Lim, C., Rausch, Z., and Haidt, J. (2026). Are School Phone Restriction Policies Working? A Review of the Natural Experiments and Advice to Researchers on Expanding the Outcomes Measured. Working paper, Stern School of Business, New York University, SSRN 7443619.

Pieh, C., Humer, E., Hoenigl, A., Schwab, J., Mayerhofer, D., Dale, R., and Haider, K. (2025). Smartphone screen time reduction improves mental health: a randomized controlled trial. *BMC Medicine*, 23:107.

Rausch, Z. and Haidt, J. (2026). Why Ferguson's meta-analysis should not be used to inform debates on social media's impact on depression and anxiety: A commentary. *Psychology of Popular Media*. Forthcoming. SSRN 5224958.

Saffer, H. (2026). Youth Mental Health and School Smartphone Bans: Early Evidence. NBER Working Paper 35181.

Schmidt-Persson, J., Rasmussen, M. G. B., Sørensen, S. O., Mortensen, S. R., Olesen, L. G., Brage, S., Kristensen, P. L., Bilenberg, N., and Grøntved, A. (2024). Screen media use and mental health of children and adolescents: A secondary analysis of a randomized clinical trial. *JAMA Network Open*, 7(7):e2419881.

Sohn, S., Rees, P., Wildridge, B., Kalk, N., and 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.

University of Essex, Institute for Social and Economic Research (2026). *Understanding Society: Waves 1–15, 2009–2024 and Harmonised BHPS: Waves 1–18, 1991–2009*. Data collection, 21st edition. UK Data Service, SN 6614. http://doi.org/10.5255/UKDA-SN-6614-22.

Youthinmind (2016). *Scoring the SDQ*. sdqinfo.org, https://www.sdqinfo.org/py/sdqinfo/c0.py, accessed September 2026.

Table 1. Pre-treatment characteristics of high-uptake schools, low-uptake schools, and the two control sets.

	High uptake (≥20%)	Low uptake (all <5%)	Matched controls	Covariate-only controls
School census (January censuses)				
Number of schools	329	12,012	987	987
SEMH per 1,000	28.4	33.4	27.8	26.3
Free school meals (%)	10.5	26.0	11.8	10.8
IMD-2019 deprivation score	10.8	23.2	11.6	11.0
Mean enrollment	280.3	283.5	277.6	279.1
Urban (%)	64.7	73.3	64.0	61.4
Child DLA, school catchments (DWP Stat-Xplore)				
Number of catchments	330	11,003	990	990
Child psychiatric DLA per 1,000	36.3	60.5	37.8	42.7
Free school meals (%)	10.6	26.6	11.7	10.9
IMD-2019 deprivation score	10.6	23.5	11.3	10.8
Catchment child population (5–15)	478	513	477	478

Notes: Rows are grouped by data source, with the outcome measure first. Every value is measured in 2023, the last pre-treatment year. School census rows are from the January 2024 census (academic year 2023/24). Child DLA rows are the calendar-2023 mean of the quarterly stock, in claims per 1,000 children aged 5–15 in the school's catchment. School census rows cover the balanced school panel. Child DLA rows cover the catchment samples of the DLA design. The two samples overlap only partly. "Matched controls" are each outcome's own 1:3 matched set. For school rows, this set is the SEMH match on 2021–2023 levels, IMD-2019, FSM and log enrollment. For DLA rows, it is the catchment match of the main estimate. "Covariate-only controls" are the FSM, IMD-2019 and log-size donor sets of the SDID and triple-difference designs, with no outcome levels in the match. The catchment panel's free-school-meal and IMD-2019 rows are the anchor school's values, the covariates of the catchment match. The catchment child population is the denominator of the DLA rates, summed over the catchment's neighborhoods.

Table 2. Effect of high pact uptake on the two main 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 FE counterfactual is the treated units' predicted level on the control time path. Each outcome uses its own matched sample (treated ≥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) are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 3. Decomposition of the pooled DLA effect by component, age, sex, and claim duration.

	Change in rate (per 1,000)	Share of total	% change within group
Child psychiatric DLA, 5–15	−3.38 (0.86)	100%	−6.1 (1.6)
Behavioral	−2.41 (0.53)	71%	−10.1 (2.2)
Ages 5–11	−1.91 (0.41)	56%	−12.9 (2.8)
Behavioral disorder	−1.76 (0.36)	52%	−15.1 (3.1)
Boys	−1.41 (0.29)	42%	−18.1 (3.7)
Girls	−0.35 (0.21)	10%	−9.1 (5.4)
Claim duration 0–12m	−0.84 (0.21)	25%	−21.5 (5.5)
Claim duration 1–2y	−0.65 (0.21)	19%	−16.3 (5.2)
Claim duration 2y+	−0.25 (0.20)	8%	−6.7 (5.3)
Hyperkinetic (ADHD)	−0.15 (0.19)	4%	−4.9 (6.4)
Ages 12–15	−0.50 (0.33)	15%	−5.6 (3.7)
Developmental (placebo)	−0.92 (0.68)	27%	−2.9 (2.2)

Notes: The estimates are pooled additive differences-in-differences on suppression-free counts. The counts are summed over the 330 treated catchments and over their 990 matched controls, the matched set of Table 2. The estimates compare the 2023 average with the 2025 average. Each pool's 2023 child population is the fixed denominator. "Change in rate" is per 1,000 children aged 5–15. "Share of total" divides the row's change by the total change. "% change within group" divides the row's change by the control post-period rate of the row's own cells. Indented rows partition their parent row. Under behavioral disorder at ages 5–11, the sex rows and the duration rows are each a separate partition of that row. Psychoneurosis, the third behavioral sub-component, is omitted as near-empty (0.5 claims per 1,000, Table 1). The behavioral and developmental components exhaust the psychiatric caseload. The behavioral sub-components are queried singly from the same pooled counts (Appendix A.6). Shares do not sum exactly to 100 percent because each component is queried separately. Each query is subject to Stat-Xplore cell-key perturbation. The gap is 2 percent of the total effect. Standard errors (in parentheses) are approximate Poisson errors from the counts. Each pool-year count is treated as one Poisson observation with variance equal to its level. The four pool-year counts are treated as independent. They omit common area-level shocks. On the 5–15 total they coincide with the LAD-clustered standard error of the matched TWFE (0.86 claims per 1,000 in both cases).

Table 4. 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. Each comes from a separate matched TWFE event study of units above the stated pact-uptake threshold against 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. They are matched on their 2021–2023 claim levels, the anchor school's IMD-2019 and FSM share, and log child population. Controls come from all <5% catchments in the national catchment cube. The dose progression therefore holds the geographic design fixed across columns. The ≥20% DLA estimate is the main design of Table 2. Standard errors (LAD-clustered, in percent-of-counterfactual units) are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 5. The two main 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 are in claims or pupils per 1,000, averaged over the two post-treatment years. Standard errors are in parentheses. They are LAD-clustered for TWFE and the triple difference, and jackknife for SDID. 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 base is the FE counterfactual for TWFE and SDID. For the triple difference, it is the post-treatment mean of the treated moving category. 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]
Adult depression (QOF, 18+)†	0.8%	0.3%
	[−0.4, 1.9]	[−0.9, 1.4]

Notes: Each cell is the average post-treatment effect as a percent of the FE counterfactual, with its 95% confidence interval (LAD-clustered) in brackets. The estimates come 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. The main and placebo rows therefore partition the caseload into disjoint components with independent sampling error. Table 2 reports the 5–15 total. Table 3 decomposes it. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. †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. An area-wide confound, such as an engaged community expanding local services, would also affect adults.

Table 7. Placebo-in-time test for the two main 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 actual February 2024 treatment date. It averages the two post-treatment years. "Placebo" re-dates the treatment to February 2022. It uses pre-treatment data only. Its matched TWFE rows are 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.

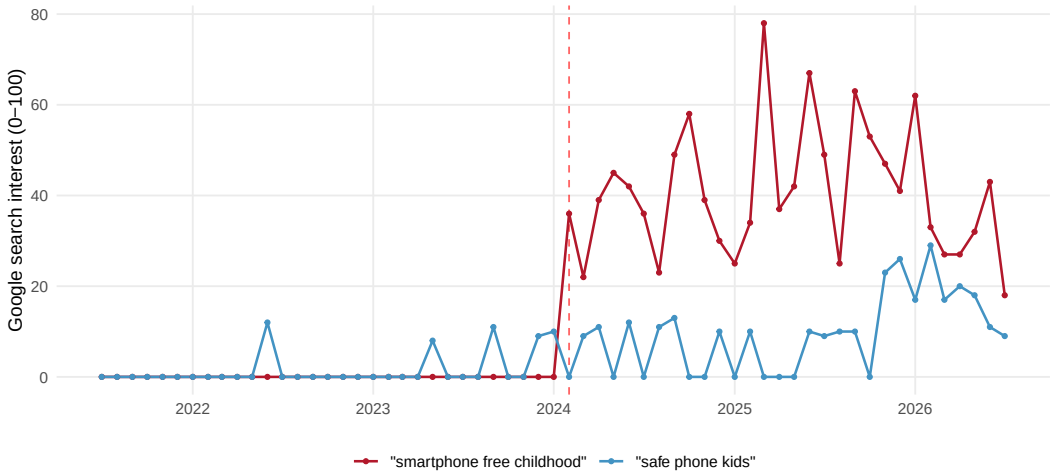


Figure 1. Google search trends on SFC and phone safety.

Notes: Monthly Google Trends search interest in the United Kingdom for "smartphone free childhood" and "safe phone kids." The two terms are 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.

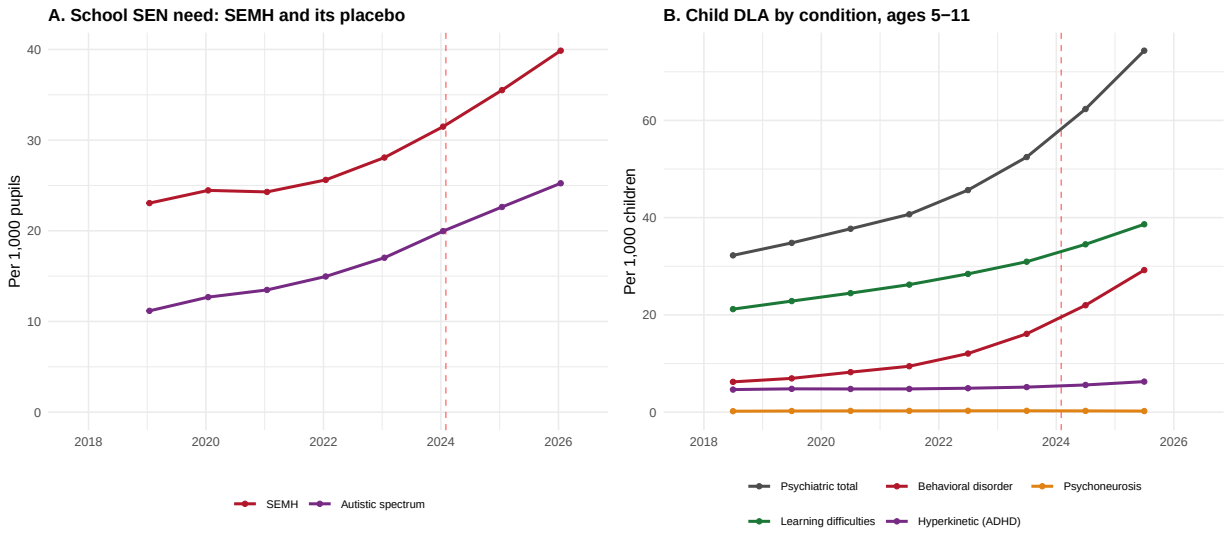


Figure 2. Evolution of the analysis outcomes, England.

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: child psychiatric DLA claims by main condition, per 1,000 children aged 5–11 in England. Each series is the annual calendar-year mean of the quarterly stock. The series are the psychiatric total and its components: learning difficulties (which record autism), behavioral disorder, hyperkinetic/ADHD, and psychoneurosis. Appendix Figure 7 shows the corresponding series for ages 12–15. The denominator is the 2023 child population. The dashed vertical line (February 2024) marks the start of the movement.

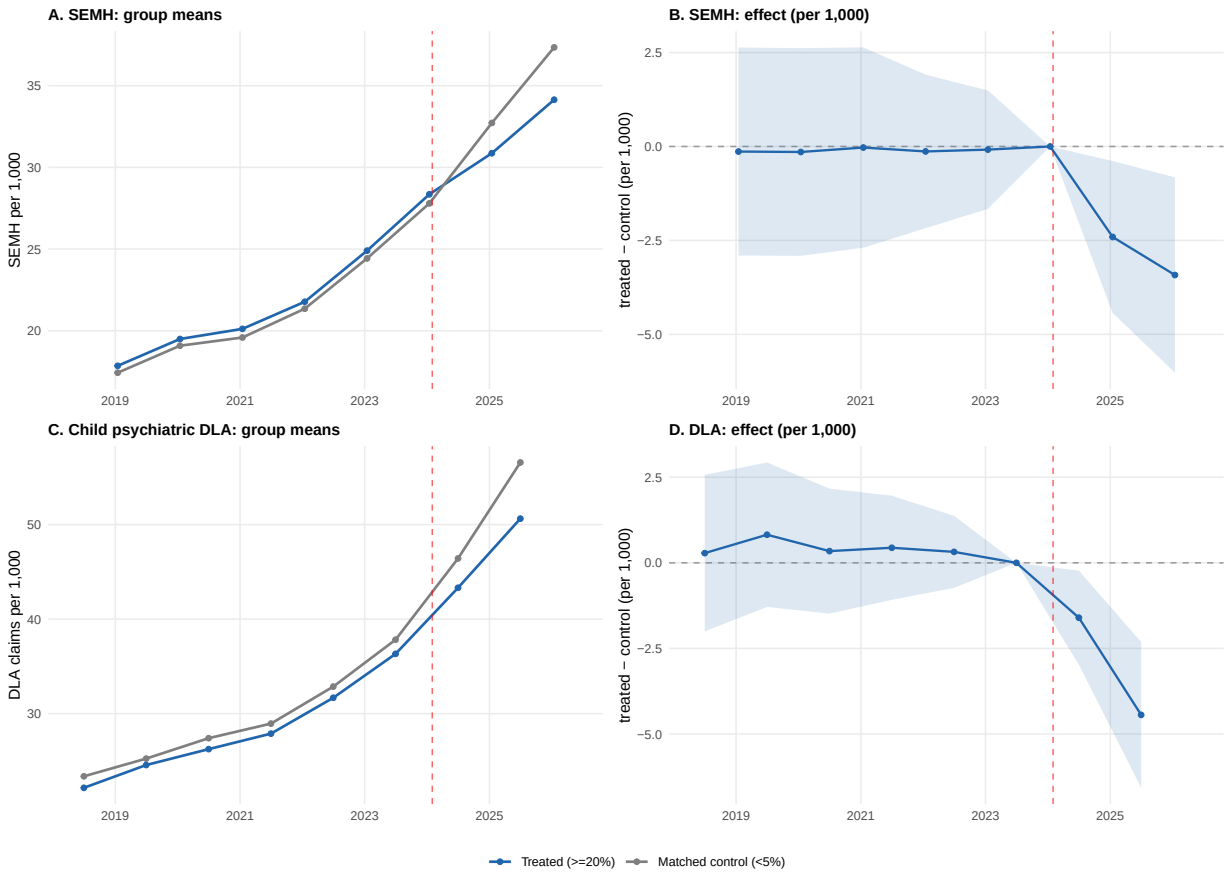


Figure 3. Matched event studies for the two main outcomes.

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 is by January census year. Child psychiatric DLA is the calendar-year mean of the quarterly DWP claim stock. Each neighborhood is 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.

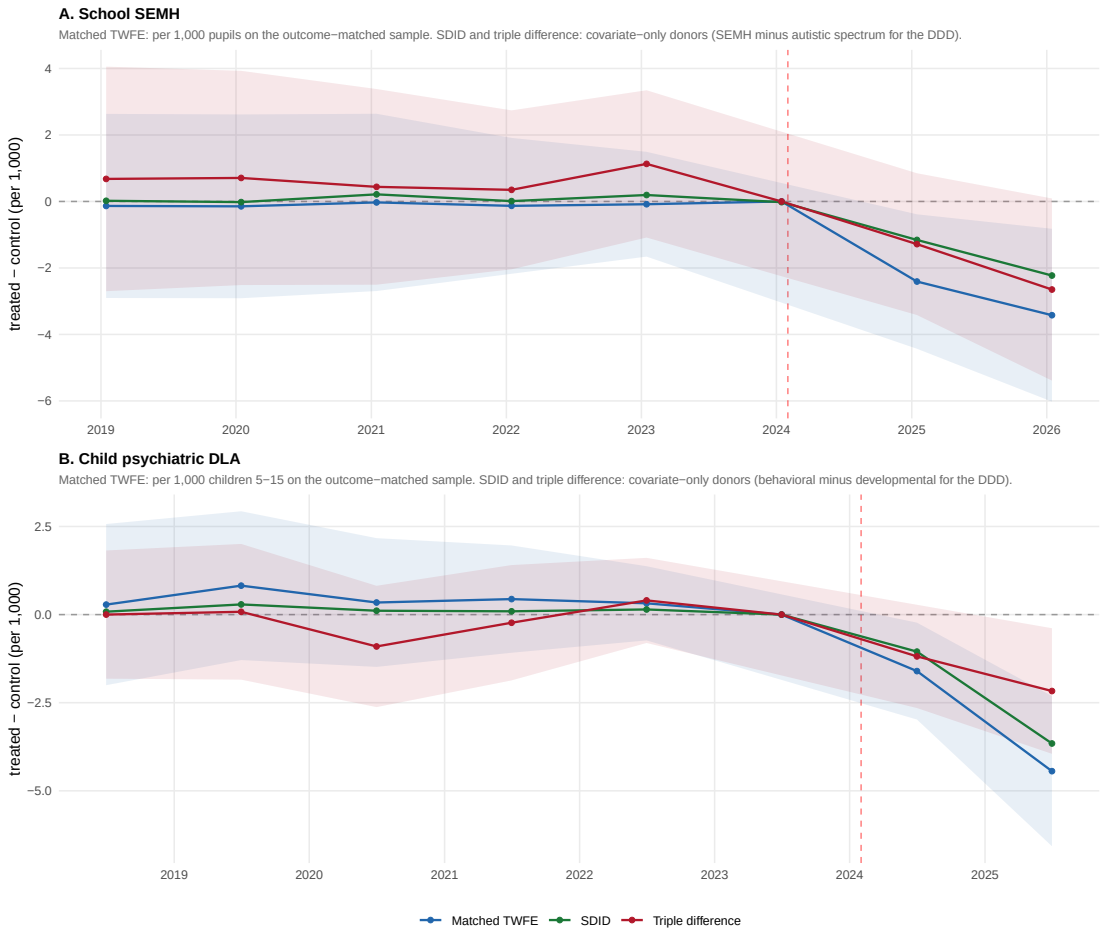


Figure 4. Event studies for the two main outcomes under the three estimators.

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, with 95 percent intervals. The moving category is SEMH minus autistic spectrum in Panel A and behavioral minus developmental in Panel B. The dashed vertical line (February 2024) marks the start of the movement.

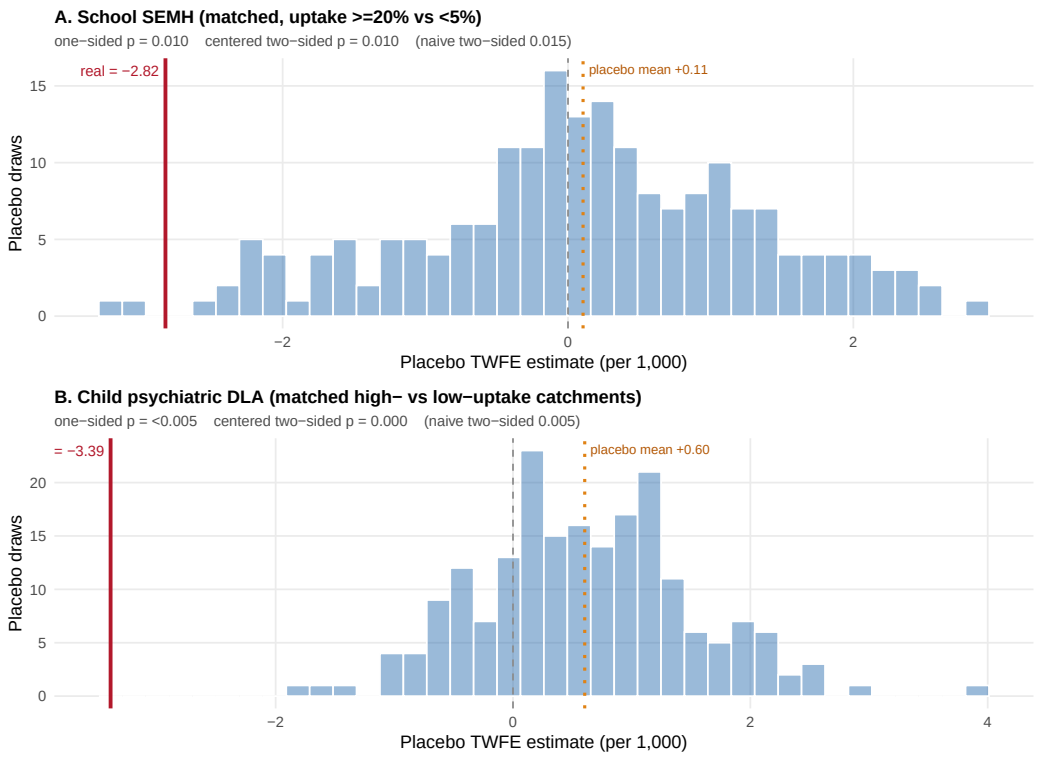


Figure 5. Placebo-in-space permutation distributions for the two main outcomes.

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. Each draw reassigns treatment to a same-size, same-clustering set of never-treated communities. The placebo set is re-matched in each draw. The solid red line is the actual estimate. The dashed gray line is zero. 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 initiatives over the study period. It distinguishes headteacher-led ownership-delay pledges, grassroots parent pacts, and in-school bans. In-school bans restrict use on premises but not ownership (Section 2).

[Table 9 about here.]

A.3 Additional Descriptive Tables

Table 10 gives the distribution of Parent Pact signup shares across schools (Section 3.1). It also shows the geographic clustering of the high-uptake group. 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). A pupil's FSM eligibility depends on whether a parent or guardian receives qualifying benefits. The main qualifying benefit is 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, it is the school-level socioeconomic covariate in the matched design.

Other school characteristics. Additional covariates include enrollment, school type, Ofsted rating, and other SEN rates. School types are community, academy, voluntary-aided, voluntary-controlled, and free school. The other SEN categories are communication and interaction, cognition and learning, and sensory and physical needs.

Census timing. The January census takes place in the third week of January. 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. For example, year 2024 denotes AY 2024/25, with a January 2025 census. The paper's text and exhibits label observations by their measurement date instead. That census therefore 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. Claimants transition to Personal Independence Payment at age 16. The under-16 caseload is therefore the complete count of child claims. The analysis excludes the small residual 16–17 band in the published data. 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 treated neighborhoods (25.6 against 10.0 percent of cells). These neighborhoods are more affluent, with lower caseloads. Any resulting bias has the opposite sign to the estimated reduction. As caseloads grow, more cells exceed the suppression threshold. The all-cell zero share falls from 20.1 percent in 2018 to 3.9 percent in 2025. This change mechanically raises recorded treated rates relative to control. The suppressed-cells-at-4 row of Table 21 bounds the estimate under the opposite extreme. In the catchment cube, pooling neighborhoods keeps the main total far above the suppression threshold. The treated and matched-control zero shares are close (8.9 against 7.7 percent). The sparser splits have more zeros in both groups. These splits are the behavioral component, the 11–15 band, and the sex split.

[Table 12 about here.]

Adult depression (QOF, placebo). As an adult placebo, I use recorded depression prevalence among GP-registered adults (18+). The data come from the Quality and Outcomes Framework (QOF). I map them from practice to LSOA level using patient-registration shares. No values are imputed. The COVID-disrupted FY 2020/21 register was not collected. The panel leaves it as a gap. FY 2023/24 is left out of the panel. That year, the QOF depression register changed from a prevalence to an incidence basis. The FY 2023/24 values are not comparable with earlier years. The event study therefore references the last observed pre-treatment year (FY 2022/23). The pre-versus-post reference is always an observed year. Matching uses the 2021 and 2022 levels. QOF runs annually from 2013, giving a long pre-treatment window. A primary-school smartphone-delay pact cannot affect adult depression. A change in adult depression would therefore indicate a confounding area-level shock.

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. The two taxonomies differ. This section describes the contents of each and their correspondence.

School SEN categories. The school census records a single primary type of need for each SEN pupil. SEN pupils are those on SEN support or with an EHC plan. The types follow 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) is the main moving outcome. The Code describes it as covering anxiety, depression, self-harm, substance misuse, eating disorders, and attachment difficulties (paragraph 6.32). It also explicitly covers 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, autistic spectrum (ASD), and physical and sensory need. Cognition and learning covers specific, moderate, severe and profound learning difficulty. Specific learning difficulty includes dyslexia, dyscalculia

and dyspraxia. Physical and sensory need covers physical disability and hearing, visual and multi-sensory impairment. I report these specific need types. The Code's 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).[34] I split these into two components. The behavioral component contains psychoneurosis, behavioral disorder and hyperkinetic syndrome. A smartphone-delay effect would plausibly affect these emotional and behavioral conditions. The developmental component contains learning difficulties and psychosis. These are structural conditions unaffected by a phone delay. Psychosis is negligible among children. Child claims 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 heavily suppressed at the catchment level. A matched estimate on this residual fails its pre-trend test (joint $p = 0.03$). Its standard error is near 7 percent of its counterfactual. It is too imprecise to inform the falsification. Unlike its working-age successor PIP, DLA has no separate autism category. Autistic-spectrum claims are recorded chiefly under learning difficulties. Autism is therefore in the developmental component. This placement is consistent with its placebo status in the school data.

Correspondence. Table 13 maps the more common conditions across the two systems. Conduct-type conditions are decreasing in both systems. The school outcome cannot separate specific conditions. In the DLA claims data, I find no decrease in ADHD (Section 5.2). Behavioral disorder corresponds to the conduct and oppositional disorders of ICD-10. DSM-5 groups these as disruptive behavior disorders. It classifies ADHD, autism, and learning difficulties as neurodevelopmental disorders. The estimated effects divide along the same boundary. 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. It appears only in the school data.

[Table 13 about here.]

A.7 Symptoms of Children with SEMH-Type Needs and Child DLA

This appendix describes the survey evidence in Section 2.5.

Data and groups. Understanding Society is a longitudinal survey of households in the United Kingdom (University of Essex, Institute for Social and Economic Research, 2026). Waves 13 and 15, with interviews from January 2021 to May 2025, are the only waves that ask about special educational needs. For each child aged 4 to 15, a parent reports any health or behavioral problem, special educational need, or disability. The parent also reports any EHC plan or statement, and the reasons from a checklist of conditions. I keep children living in England. The survey has no

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

SEMH category. I call a need SEMH-type when the parent names ADHD, behavioral problems or hyperactivity, or mental illness or depression. Of the checklist conditions, these come closest to the SEMH area of the SEND Code of Practice (Department for Education & Department of Health, 2015). Parents may name several reasons, while the census records one primary need. A narrower group therefore excludes children whose parents also name autism or learning difficulties.

The survey records a child's DLA as the income of the responsible adult. Only children under 16 can make a new DLA claim (Department for Work and Pensions, 2026). Adults below State Pension age claim Personal Independence Payment instead. I therefore treat DLA reported by an adult aged 65 or under as an award to a child in the household. Psychiatric-type DLA requires such a report and a parent-reported need for the child. The parent must name a learning, ADHD, autism, behavioral, or mental illness reason, and no sight, hearing, physical, or medical reason. This definition approximates the DWP's psychiatric conditions (Appendix A.6). Its attribution is ambiguous when a household has more than one child with a need.

Measures and estimation. The SDQ has 25 items on five scales (Goodman, 1997). Each scale has five items and a score from 0 to 10. Emotional symptoms, conduct problems, hyperactivity, and peer problems sum to the total difficulties score, from 0 to 40. The fifth scale measures prosocial behavior, a strength. Parents complete the questionnaire for children aged 5 and 8, with a catch-up at 6 and 9. Their questionnaire is part of a self-completion module, which telephone interviews skip. Youths aged 10 to 15 complete their own version. I use the parent report at ages 5 to 9 and the self-report at ages 10 to 11. The self-report row for psychiatric-type DLA uses ages 10 to 12, where the group is larger. The abnormal range follows the original three-band scoring (Youthinmind, 2016). It starts at a total of 17 on the parent report and 20 on the self-report.

All estimates use the survey's cross-sectional weights. These are the child development weight, the youth interview weight, and the enumeration weight for prevalence. Each gap is the coefficient from a weighted regression of a score on the group indicator. Children whose parents report no need form the comparison group. The regression pools the two waves, with fixed effects for single-year age by wave and for sex. Standard errors are clustered by primary sampling unit. Counts are rounded to the nearest 10. No statistic uses fewer than 10 children. A share is withheld when fewer than 10 children in the group have the attribute. It is also withheld when fewer than 10 lack it. In groups under 50 children, shares of 0 or 100 percent are withheld as well. The tables mark which side of a withheld share is small.

Results. Children without an EHC plan show nearly the full gap (Table 14). They make up most of the SEMH-type group (80 of 90). Their gap is 10.8 points (9.2 to 12.5), against 11.3 for the whole group. At ages 10 to 11, the self-reported gap is 7.6 points (5.9 to 9.4). This gap is 1.2 standard deviations, against 2.2 on the parent report. Figure 6 splits the parent-report gaps by subscale. For SEMH-type need, the hyperactivity gap is 4.6 points. The conduct-problems gap is 2.7 points. The gap on emotional symptoms is 2.0 points. Psychiatric-type DLA shows the same pattern. Its gaps are 4.7 points on hyperactivity and 1.3 on emotional symptoms. Autism or learning difficulties without an SEMH-type reason show a different profile. Their conduct gap is only 0.8 points. Their largest gaps are on hyperactivity (2.9), peer problems (2.4), and prosocial behavior (-2.3).

Table 16 compares survey prevalence with administrative rates. The survey's SEMH-type share is three times the census rate ($8.7/2.9 = 3.0$). Parents who also name autism or learning difficulties account for most of the difference ($8.7 - 3.9 = 4.8$ of 5.8 points). The survey EHC plan share is 6.6 percent. It is near the SEN2 rate for all settings, 5.2 percent (Department for

Education, 2025). It exceeds the state primary census rate of 2.6 percent. The survey records psychiatric-type DLA for 2.5 percent of children, against 5.1 percent in DWP counts. It therefore captures about half of psychiatric claims ($2.5/5.1 = 0.49$). The DLA rows describe only claims that parents report. Each such claim must also be attributable to a child.

[Table 14 about here.]

[Figure 6 about here.]

[Table 15 about here.]

[Table 16 about here.]

A.8 Parent Pact Registrations in Three Prominent Pledge Towns

The following tables report cumulative SFC Parent Pact registrations per school in three pledge towns. The registrations are as of 12 June 2026. The tables also include a Hertfordshire comparison benchmark. They document wide variation in signup rates within each town. This variation motivates measuring treatment by school-level uptake.

[Table 17 about here.]

[Table 18 about here.]

[Figure 7 about here.]

A.9 Software

The analysis uses R and 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). It 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. 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. Reweighting therefore handles a residual pre-trend in an unweighted event study. I apply SDID on the covariate-only donor sets of the triple difference. Their match uses FSM, IMD-2019, and log size, with no outcome levels. The weights therefore fit the treated pre-trend. Re-dating treatment tests the design directly (Table 7).

Table 19 reports the SDID estimates with unit-jackknife standard errors. The estimates use the covariate-only donor sets. Percent effects use the FE-counterfactual base of the matched TWFE tables. The estimates agree in sign with the matched TWFE estimates. They are smaller in magnitude. School SEMH is −1.69 per 1,000, against −2.92 under TWFE. The DLA total is −2.35, against −3.02. The behavioral component accounts for the DLA reduction. I find no evidence of an effect on the two placebos, the school autistic-spectrum category and the developmental DLA component.

[Table 19 about here.]

B.2 Additional Results Tables and Figures

This section collects exhibits summarized in the main text. Table 20 reports the affluence heterogeneity of Section 5.3. Table 21 reports the specification checks of Section 6. Figure 9 reports results from ongoing work on two further school outcomes, autumn-term illness absence and teacher turnover.

[Table 20 about here.]

[Table 21 about here.]

[Figure 8 about here.]

[Figure 9 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 ≥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 assigns each LSOA to its nearest-school catchment. The few schools without usable coordinates use their published census code. No school drops for missing matching data. The DLA catchments are built around the high-uptake schools directly, whether or not the school appears in the census panel (Section 3). The school and catchment counts therefore 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 signup 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 signup 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. Standard errors throughout the paper are clustered at this district level.

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, January 2024 census. Each pupil has one primary type of need. The counts combine the SEN-support and EHC-plan tiers. 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. ‡Placebo category. A smartphone pledge cannot affect this medically determined need.

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 returned as zero. A recorded zero is therefore either a true zero or a suppressed small count. Panel A covers the raw LSOA quarterly panel. Treated LSOAs have a nearest school with pact uptake $\geq$20%, and control LSOAs one with uptake $<$5%. Panel A also shows the all-cell zero share for the first and last panel years. Panel B covers the catchment cube on the main estimation sample (treated catchments against their matched 1:3 controls). Its counts are retrieved as single custom-geography totals. Suppression therefore 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)	placebo
Anxiety, depression	SEMH	Psychoneurosis (behavioral)	decreasing
Conduct, behavioral	SEMH	Behavioral disorder (behavioral)	decreasing
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. "Decreasing" denotes the categories in which recorded need falls under high pact uptake. "Placebo" denotes the categories it does not affect. ADHD is recorded within the school SEMH outcome. Its DLA claims are unchanged (Section 5.2). Inclusion in the school category does not by itself imply an expected effect. The "Anxiety, depression" row is a taxonomy correspondence only. Psychoneurosis is essentially absent among child DLA claims (Section 3). In the sample, the DLA side of that row has no cases. The behavioral component therefore consists of behavioral disorder and hyperkinetic/ADHD. †DLA, unlike its working-age successor PIP, has no separate autism category. Autistic-spectrum claims are therefore recorded chiefly under learning difficulties.

Table 14. Total difficulties by parent-reported need, Understanding Society, England.

		Total difficulties		Gap to no reported need	
	Children	Mean	Abnormal (%)	Points [95% CI]	SD
Panel A: parent report, ages 5–9					
No reported need	1,060	8.2	7.3		
Any reported need	200	17.0	51.6	8.5 [7.3, 9.8]	1.7
EHC plan (any reason)	60	17.3	52.7	8.9 [6.7, 11.2]	1.7
SEMH-type need	90	19.9	72.5	11.3 [9.9, 12.8]	2.2
no autism or learning reason	50	18.3	54.4	9.6 [7.6, 11.7]	1.9
no EHC plan	80	19.4	67.5	10.8 [9.2, 12.5]	2.1
Psychiatric-type DLA	30	19.2	‡	10.7 [8.2, 13.2]	2.1
Panel B: youth self-report, ages 10–11					
No reported need	600	11.0	9.6		
Any reported need	150	15.5	25.5	4.6 [3.1, 6.0]	0.7
EHC plan (any reason)	50	16.8	22.5	5.9 [3.6, 8.1]	1.0
SEMH-type need	60	18.7	39.7	7.6 [5.9, 9.4]	1.2
no autism or learning reason	20	19.4	†	8.3 [5.6, 11.0]	1.4
no EHC plan	40	17.8	36.9	6.9 [5.2, 8.6]	1.1
Psychiatric-type DLA, ages 10–12	30	16.1	26.0	5.2 [3.0, 7.4]	0.8

Notes: Understanding Society waves 13 and 15 (interviews January 2021 to May 2025), children in England. The groups come from the parent's report of a special educational need or disability and its reasons (Appendix A.7). SEMH-type need: the parent names ADHD, behavioral problems or hyperactivity, or mental illness or depression. The no EHC plan row includes children being assessed for a plan. Psychiatric-type DLA: an adult aged 65 or under in the household reports DLA. The parent also reports a need for the child. The parent names a learning, ADHD, autism, behavioral or mental illness reason and no sight, hearing, physical or medical reason. Children: children in the group with a positive survey weight, rounded to the nearest 10. Means, shares and gaps are weighted, with the child development weight in Panel A and the youth interview weight in Panel B. Abnormal: a total difficulties score of 17 or more (parent report) or 20 or more (self-report). Gap: coefficient from a weighted regression of the score on the group indicator, against children with no reported need. The regression includes fixed effects for single-year age by wave and for sex. Confidence intervals use standard errors clustered by primary sampling unit. SD: the gap divided by the weighted standard deviation of the score among children with no reported need. The last row of Panel B covers ages 10 to 12. Its comparison group is the 970 children of those ages with no reported need (mean 11.2). † Withheld: fewer than 10 children in the group score in the abnormal range. ‡ Withheld: fewer than 10 children in the group score below it.

Table 15. Parent-reported need, EHC plans and psychiatric-type DLA among children in each SDQ band, ages 5–9.

		Share with the label (%)			
	Children	Any need	SEMH-type need	EHC plan	Psychiatric-type DLA
All children	1,290	18.2	8.8	7.0	2.0
Abnormal total difficulties	180	60.9	41.7	24.1	9.5
Abnormal conduct or hyperactivity	340	43.6	29.3	16.4	6.3
Abnormal emotional symptoms	180	42.7	24.8	15.4	†
Combinations of abnormal ranges					
conduct or hyperactivity only	230	38.2	27.4	13.7	7.0
emotional only	70	19.5	†	†	0.0
both	110	53.4	32.7	21.4	†
neither	880	7.4	†	3.4	†

Notes: Understanding Society waves 13 and 15, children aged 5 to 9 in England with a parent-reported SDQ. Children: children in the band, rounded to the nearest 10. Shares are weighted with the child development weight. They use the children whose status on the label is known. Abnormal ranges follow the original three-band scoring of the parent questionnaire. The lower bounds are 17 for total difficulties, 4 for conduct problems, 7 for hyperactivity, and 5 for emotional symptoms. Table 14 defines the labels. † Withheld: fewer than 10 children in the band have the label.

Table 16. Reported need and DLA in Understanding Society against administrative rates, children aged 5–10 in England (percent).

	UKHLS	Administrative	Source
Panel A: special educational needs			
Any reported need	19.1	16.1 (18.7)	Census, any SEN
SEMH-type need	8.7	2.9 (3.2)	Census, SEMH
no autism or learning reason	3.9	2.9 (3.2)	Census, SEMH
EHC plan (any reason)	6.6	2.6 (5.6)	Census, EHC plan
		5.2	SEN2, all plans
SEMH-type need with an EHC plan	2.6	0.7	SEN2, SEMH plans
no autism or learning reason	0.5	0.7	SEN2, SEMH plans
Panel B: Disability Living Allowance			
Household reports DLA	7.3	6.4	DWP, all conditions
Attributed to a child with a need	3.8	6.4	DWP, all conditions
Psychiatric-type DLA	2.5	5.1	DWP, psychiatric
physical reasons allowed	3.4	5.1	DWP, psychiatric

Notes: UKHLS: weighted share of children aged 5 to 10 in England, waves 13 and 15 pooled, with the enumeration weight. The 3,770 to 3,840 children (rounded) are those whose status on the measure is known. Administrative rates are matched to each child's interview by census year and DWP quarter. They are averaged with the same weights. Census rates are shares of pupils in state-funded primary schools. The census measures are SEN support or an EHC plan (any SEN), SEMH as primary need, and an EHC plan. The rates in parentheses add state-funded special schools at all ages. SEN2 counts the EHC plans that local authorities maintain in every setting in January 2025, for all needs and for SEMH as primary need. DWP counts child DLA claims for all conditions and for psychiatric conditions. SEN2 and DWP counts cover ages 5 to 10. Each is divided by the population of England at those ages. The survey records DLA as the income of an adult. The household row therefore counts every child in a household where an adult aged 65 or under reports DLA. Table 14 defines the other survey measures. Restricting the survey to children at state schools changes no share by more than 0.3 percentage points.

Table 17. SFC Parent Pact registrations: St Albans primary 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. The share is pacts over enrollment as recorded by SFC. The Pact tool launched in September 2024, after the St Albans pledge. The rows are the state-funded primary schools in the St Albans district with postcodes AL1 to AL4. The paper classifies a school as treated if its own share is at least 20 percent (Section 3.1). Schools without an SFC listing are shown with zero pacts.

Table 18. SFC Parent Pact registrations: Hitchin, Cambridge CB4 and a Hertfordshire benchmark

School	Pacts	Share (%)
Panel A: Hitchin primary schools		
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 primary schools		
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 primaries (37 schools, 35 with pacts)	1380	10.8
Hitchin primaries (12 schools, 12 with pacts)	774	13.5
Cambridge CB4 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 17. The Hertfordshire 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. The benchmark therefore overstates typical uptake elsewhere in England. The benchmark has a mean share of 4.8% and a 90th-percentile share of 12.8%. Mean shares are 14.4% for St Albans, 19.2% for Hitchin, and 10.1% for Cambridge CB4.

Table 19. 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
Autistic spectrum (school placebo)	0.33	(0.75)	[−1.13, 1.79]	1.6%	0.10
Child psychiatric DLA (5–15)	−2.35	(0.74)	[−3.80, −0.90]	−4.7%	−3.02
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

Notes: Each row applies the synthetic difference-in-differences estimator of Arkhangelsky et al. (2021) to a covariate-only 1:3 donor set. Treated units have pact uptake ≥20% and controls <5%. The school outcomes, SEMH and the autistic-spectrum placebo, are matched on FSM, IMD-2019 and log enrollment. The child DLA rows use 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 fit the treated pre-treatment trajectory. The triple difference uses the same donor restriction. Re-dating treatment therefore tests the design directly (Table 7). School outcomes are per 1,000 pupils on the 2018–2025 panel years. Child DLA rows 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 20. 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. The estimates come from a matched TWFE event study (treated ≥20% pact uptake to three <5% controls). Estimates are shown 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) are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 21. Specification robustness for the two main outcomes.

Specification	SEMH est. (s.e.)	Child DLA psych. est. (s.e.)
Main: matched, ≥20% vs <5%	−2.82 (1.11)	−3.39 (1.15)
	[−5.01, −0.64]	[−5.64, −1.14]
Low-uptake pool <10%	−2.48 (1.12)	−2.86 (1.14)
	[−4.67, −0.29]	[−5.09, −0.63]
Match ratio 1:5	−2.98 (1.02)	−3.52 (1.08)
	[−4.98, −0.99]	[−5.63, −1.40]
Matching with replacement	−3.18 (1.17)	−3.75 (1.16)
	[−5.47, −0.88]	[−6.03, −1.47]
Excl. St Albans, Hitchin, Cambridge	−3.23 (1.14)	−3.24 (1.06)
	[−5.47, −0.99]	[−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) under the stated specification (main match 1:3). Its LAD-clustered standard error is in parentheses and its 95% confidence interval in brackets. 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. The match uses the outcome's 2021–2023 levels, the anchor school's IMD-2019 and FSM share, and log child population, as in the main design. The suppressed-cell row imputes every catchment-quarter recorded as zero at its upper bound of 4. Under the DWP disclosure rules, such a cell has a true count of 0–4. The quarterly-panel row estimates the same pooled treated × post difference-in-differences on the raw quarterly claim snapshots, on the same matched catchments. November 2023 is the last pre-treatment quarter and February 2024 the first post-treatment quarter. "Excl. St Albans, Hitchin, Cambridge" drops the three pledge towns with the most media attention from the treated group. The extended pre-period balances the school panel back to the 2015 census. It is unavailable for child DLA, whose series begins in 2018.

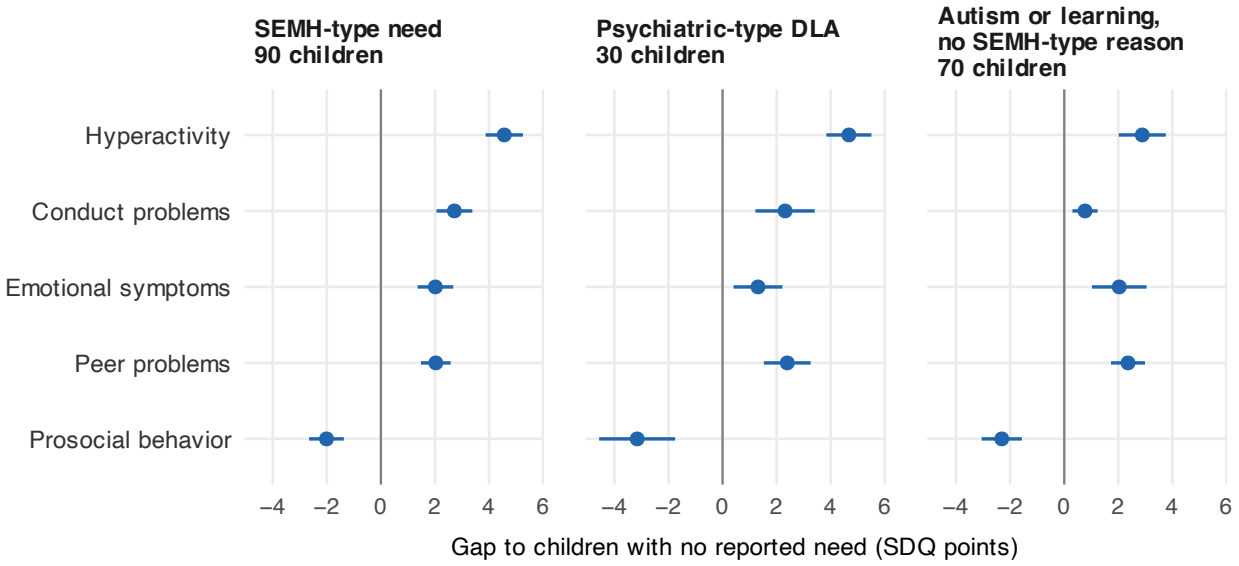


Figure 6. SDQ subscale gaps by parent-reported need, ages 5–9.

Notes: Understanding Society waves 13 and 15, children aged 5 to 9 in England with a parent-reported SDQ. Each point is the coefficient from a weighted regression of a subscale score (0 to 10) on the group indicator. The comparison group is children with no reported need. The regression includes fixed effects for single-year age by wave and for sex. Bars show 95 percent confidence intervals, with standard errors clustered by primary sampling unit. Prosocial behavior is a strength scale. A negative gap therefore means fewer prosocial behaviors. Panel titles give the number of children in each group, rounded to the nearest 10. The third panel covers children whose parents name autism or learning difficulties and no SEMH-type reason. Table 14 defines the other groups.

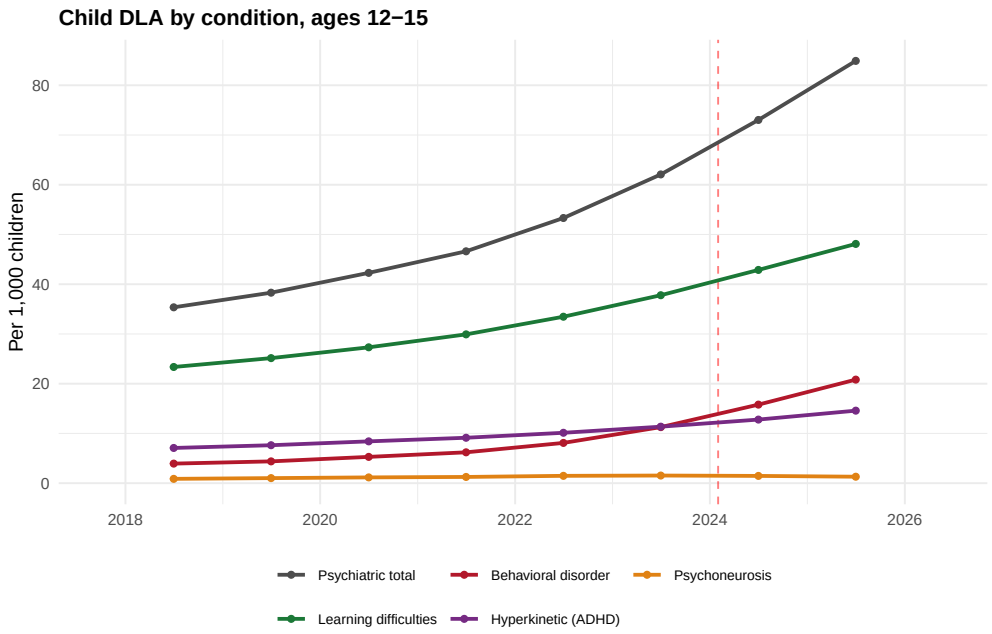


Figure 7. Child DLA claims by main condition, ages 12–15, England.

Notes: The ages 12–15 counterpart of Figure 2, Panel B. Each series is the annual calendar-year mean of the quarterly claim stock per 1,000 children aged 12–15 in England. The series are the psychiatric total and its components. The denominator is the 2023 child population. The dashed vertical line (February 2024) marks the start of the movement.

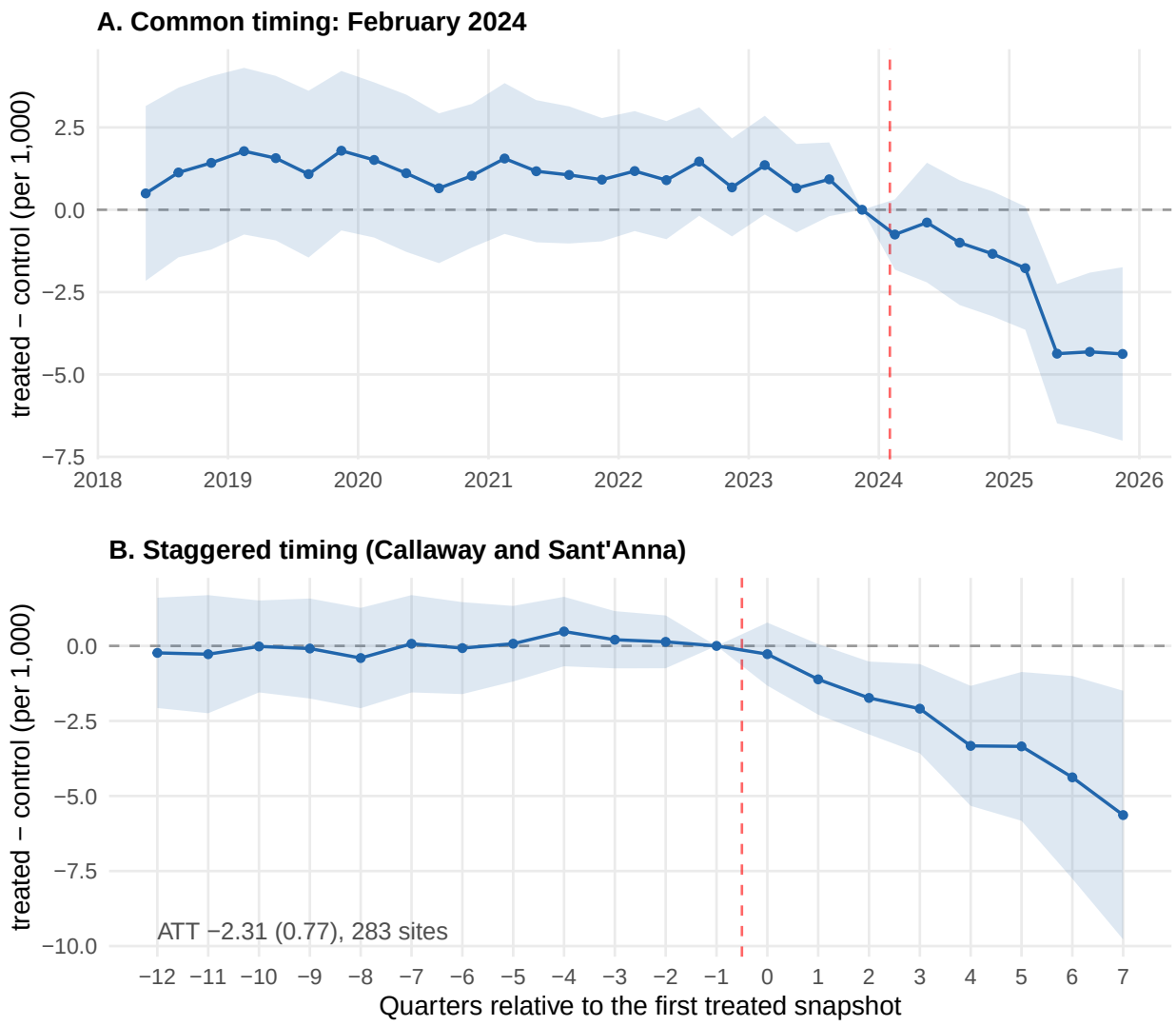


Figure 8. Quarterly child psychiatric DLA event studies with common and staggered timing.

Notes: Differenced coefficients (treated minus control, per 1,000 children aged 5–15) on the raw quarterly DLA snapshots. Panel A uses the main matched catchments (≥20% versus matched <5% pact uptake). It dates treatment to February 2024 for every catchment. Its snapshots are from February, May, August and November, from May 2018 to November 2025. Its 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 21). Panel B applies the staggered estimator of Callaway and Sant'Anna (2021) to school sites, which pool infant and junior school pairs. Its horizontal axis counts quarters relative to each site's first treated snapshot. Treated sites have pact uptake of at least 20 percent. Each enters at the first snapshot in or after the month its cumulative registrations reach 5 percent of enrollment. Registrations from the September 2024 launch month count from February 2024. Sites that cross then enter at the May 2024 snapshot. Controls are sites below 5 percent uptake, matched 1:3 on 2021–2023 claim levels, FSM share, IMD-2019 and log child population. The estimator adjusts for FSM share, IMD-2019 and log child population, with child-population weights. Its reference is the snapshot before treatment. Its 95 percent intervals come from a LAD-clustered bootstrap. Panel B includes the March 2026 snapshot. The average effect is −2.31 per 1,000 (standard error 0.77) across 283 sites.

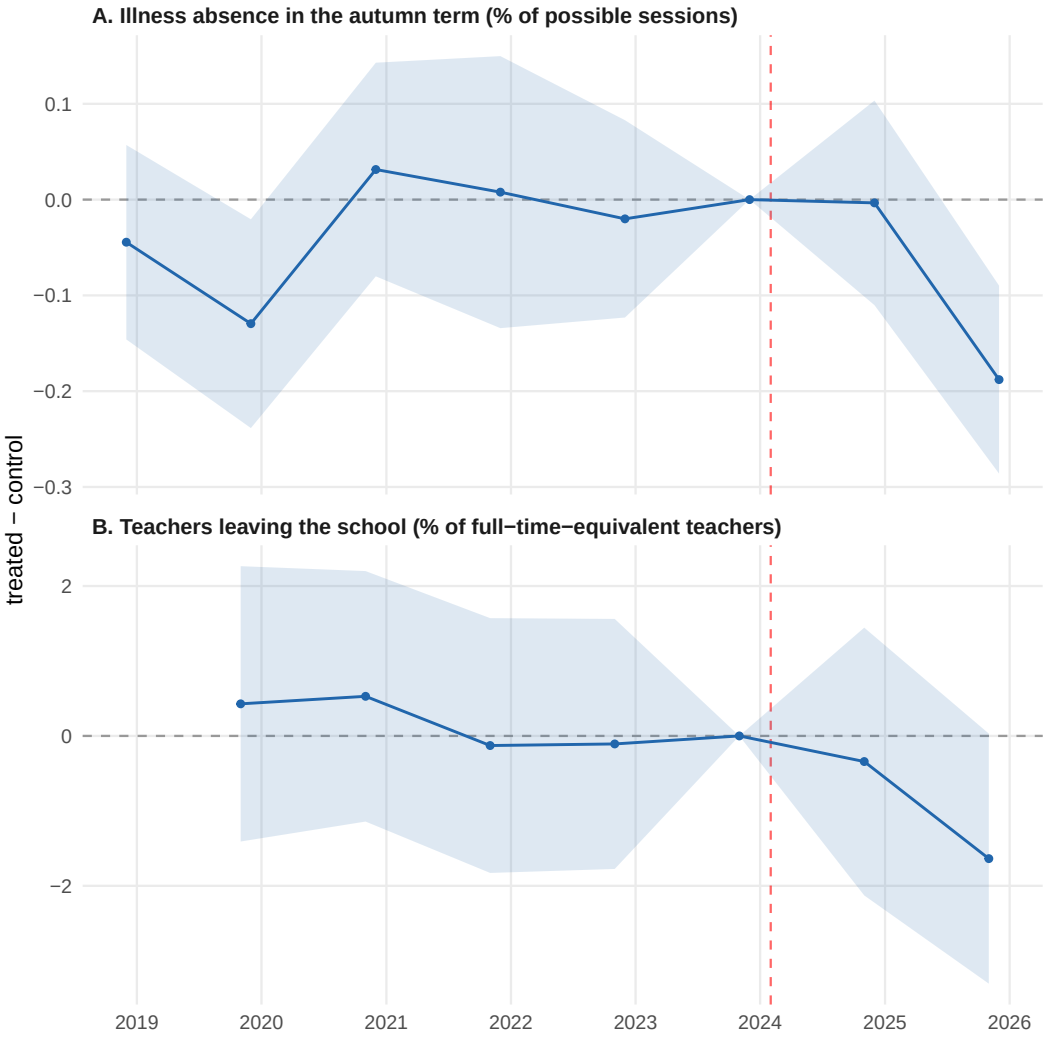


Figure 9. Matched event studies for autumn-term illness absence and teacher turnover.

Notes: Matched two-way fixed-effects event-study coefficients (treated minus control), with 95 percent LAD-clustered confidence intervals. Treated schools have pact uptake of at least 20 percent. Each is matched to three schools below 5 percent on the outcome's own pre-treatment levels, IMD-2019, FSM share and log enrollment. Pupils' mental health may affect both outcomes. Panel A shows authorized absence for illness in the autumn term, as a percentage of possible sessions. It uses only autumn terms, so every comparison is within the same season. Its matching levels are the autumn terms of 2021–2023. Its reference is autumn 2023. Each autumn term is dated at its end, in December. The pooled difference-in-differences over autumn 2024 and 2025 is −0.070 percentage points (standard error 0.038), 2.6 percent of the counterfactual mean. The autumn 2025 coefficient is −0.19 (0.05), 6.9 percent of that term's counterfactual mean. In autumn 2020, non-attendance for COVID reasons was recorded outside absence. Panel B shows the full-time-equivalent teachers who leave the school by the next November census, as a percentage of those at the first census. Its matching levels are the turnover years ending in November 2022 and 2023. Each turnover year is dated at its closing census. Its reference is the last year ending before February 2024. The pooled difference-in-differences over the years ending in November 2024 and 2025 is −1.13 percentage points (0.57), 8.4 percent of the counterfactual mean. The coefficient for the year ending in November 2025 is −1.64 (0.85), 12.4 percent of that year's counterfactual mean. The dashed line marks February 2024.