Australian National University

Crawford School of Public Policy

TTPI

Tax and Transfer Policy Institute

The effect of school closures on standardized test scores: Evidence from Australia

TTPI - Working Paper 6/2023
June 2023

Christian Gillitzer
The University of Sydney

Nalini Prasad
UNSW Sydney

Abstract

We estimate the causal effect of pandemic school closures on standardized test scores in Australia using variation in the length of school closures across states. States independently implemented school closures as part of Australia's successful COVID elimination strategy. School closures in 2020-21 ranged from 4 to 112 school days. We measure achievement using student level panel test score data from a common compulsory standardized test with high participation rates. We find no evidence of large declines in test scores, including for low socioeconomic groups. The variation in test scores we observe post school closures was like that observed in previous years.

Keywords: student test scores, COVID-19, pandemic, NAPLAN, Australia

* *We would like to thank Kadir Atalay, Bob Breunig, Colin Cameron, Daniel Carr, Ash Craig, Gigi Foster, Richard Holden, Greg Kaplan, Gregor Pfeifer, Stefanie Schurer, Mathias Sinning and seminar participants at the University of Sydney and the Australian National University for helpful comments. The project received ethics approval by the University of New South Wales and institutional exemption from ethics review by the University of Sydney. The NAPLAN data used in this study are available to academics, researchers and higher degree students at Australian tertiary institutions and authorized employees of Australian governments by application to the Australian Curriculum and Reporting Authority. This paper also uses data from the Household Income and Labour Dynamics in Australia (HILDA) Survey. See the disclaimer in the Supplementary Appendix for the HILDA survey.*
Authors: Christian Gillitzer, christian.gillitzer@sydney.edu.au; Nalini Prasad, nalini.prasad@unsw.edu.au

Tax and Transfer Policy Institute

Crawford School of Public Policy

College of Asia and the Pacific

+61 2 6125 9318

tax.policy@anu.edu.au

The Australian National University

Canberra ACT 0200 Australia

www.anu.edu.au

The Tax and Transfer Policy Institute (TTPI) is an independent policy institute that was established in 2013 with seed funding from the federal government. It is supported by the Crawford School of Public Policy of the Australian National University.

TTPI contributes to public policy by improving understanding, building the evidence base, and promoting the study, discussion and debate of the economic and social impacts of the tax and transfer system.

The Crawford School of Public Policy is the Australian National University's public policy school, serving and influencing Australia, Asia and the Pacific through advanced policy research, graduate and executive education, and policy impact.

1 Introduction

The COVID-19 pandemic caused large disruptions to education systems around the world. It is estimated that around 94 percent of students globally were affected by school closures as governments sought to contain the spread of the COVID-19 virus (United Nations, 2020). A narrative has developed from the literature that school closures during the pandemic caused substantial overall learning losses, with larger losses for students from low socioeconomic backgrounds (Betthäuser, Bach-Mortensen and Engzell, 2023). We show that in Australia learning losses were small and indistinguishable from typical variation in tests scores across states and time. This is despite the most affected students spending more than half the 2020 school year learning from home. Quantifying the effect of school closures on student achievement is necessary for policymakers to target post-pandemic educational funding and to guide choices in future pandemics. Our results demonstrate that learning losses from even long school closures need not be large.

A credible study setting to estimate learning loss requires exogenous contemporaneous variation in the duration of school closures within a testing jurisdiction in which students sit a common test with high participation (Werner and Woessmann, 2021). The Australian setting has all these features. While there have been many studies on the effect of school closures on learning loss, none of the 42 high-quality studies surveyed in Betthäuser, Bach-Mortensen and Engzell (2023) can exploit a contemporaneous control group. Studies in the U.S. have used district-level variation in schooling mode but differ from ours in analyzing district-level pass-rates rather than individual test scores (Jack et al., 2021) or results from non-mandated tests where the students taking the test may not be representative of the population (Goldhaber et al., 2022).

Most studies rely on before after comparisons because school closures affected almost all students equally within a jurisdiction. Studies that compare achievement of the cohort affected by school closures to previous cohorts suffer from the problem that the cohorts may have performed differently absent school closures (Werner and Woessmann, 2021). A smaller number of studies use student-level panel data and compare the growth in achievement between tests taken before and after school closures, compared with earlier cohorts. This research design controls for pre-pandemic differences in the level of achievement between cohorts. But it assumes that, absent school closures, learning trajectories of the affected cohort would have been the same as earlier cohorts. Differences over time in the difficulty or design of tests, changes in the time of the year tests are administered, changes in educational systems and effects of the pandemic on household incomes and health could have affected measured learning gains independently of school closures. With few exceptions, test participation was lower in the pandemic, leading to concerns that students participating in tests during the pandemic were not representative of the student population.

We address these issues using quasi-experimental variation in the duration of school closures across Australian states. Australia successfully pursued a COVID elimination strategy until mid-2021. The international border was closed, with states operating hotel quarantine systems to facilitate limited exemptions to the closed international border. States independently implemented lockdowns and school closures to maintain a zero-COVID environment. Unlike in many other countries, lockdowns and school closures were implemented preemptively to stop the spread of COVID cases rather than a last-resort policy measure (Schurer et al., 2022). Individual schools had no power to choose instructional mode. All states had periods of lockdowns and school closures in the first half of 2020 during the first COVID wave, but of differing length. Lockdowns during this first COVID wave were successful in bringing national COVID infections down close to zero. Following this students returned to the classroom. Then, in June 2020, lapses in the hotel quarantine system saw COVID case numbers rise in Victoria, Australia's second largest state. This led to the re-imposition of lockdowns and school closures in Victoria. Victoria's border to other states was closed. Victoria was locked down with schools closed between July and October 2020 to successfully regain zero-COVID, while other states remained open. School students in Victoria (*N=185,648*) spent an additional 68 days (Grade 5) to 78 days (Grade 9) learning from home in 2020 compared with students in other states (*N=557,657*). This provides us with contemporaneous variation in school closures that is plausibly exogenous to the school system. The other states provide a relatively homogeneous comparison group (Tables S1 and S2).

Our main results compare achievement of primary and secondary students in Victoria to those in other states on standardized National Assessment Program — Literacy and Numeracy (NAPLAN) tests taken in May 2021 after students in all states had returned to the classroom. We also analyze the relatively smaller variation in the duration of school closures among the other states. Several features of our study setting enable us to clearly isolate the effect of school closures on student achievement. First, each year the same NAPLAN test is administered to all students nationwide. NAPLAN test scores have been shown to be a strong predictor of a student's college entrance exam test score (Houng and Justman, 2014) and hence their future earnings. Secondly, tests in 2021 were taken in-person under normal testing conditions. Test participation remained high at 95 percent in Victoria and in the rest of Australia in 2021, minimizing selection issues. Third, the zero-COVID strategy with associated lockdowns meant health effects were moderate by international standards, including for Victoria during its second COVID wave. Fourth, the effect of the pandemic on household incomes was mild in Australia due to generous federal income support. This feature of the Australian environment is important as previous work has shown that family income losses during school closures are associated with declines in test scores (Kogan and Lavertu, 2021).

We analyze the universe of student-level longitudinal NAPLAN data between 2013 and 2021 in a difference-in-difference design. We find no evidence of large learning losses associated with school closures: estimated learning losses are –0.03, –0.01 and –0.02 standard deviations for Grades 5, 7 and 9 students in Victoria compared with students in other states. These estimates imply that a student in Victoria lost between 6 to 14 days of learning, which is considerably shorter than the additional 68 to 78 days of school closures that a student in Victoria experienced relative to their peers in other states. Results are similar for the literacy and numeracy components of the test. In all cases, we cannot reject the null hypothesis of no learning loss at the 5 percent significance level. There is no evidence of large learning losses even for students from disadvantaged backgrounds. The variation in test scores we observe post school closures is similar to that observed in previous years.

We consider the available evidence on mechanisms contributing to limited learning loss. First, we document that there was widespread access to reliable internet, electronic devices and study space to facilitate learning from home. Second, household survey data show increased parental inputs. Parents of children experiencing longer durations of school closures were more likely to reduce hours of paid work than parents of children experiencing short durations of school closures. Third, we look for evidence that there were temporary learning losses that faded by the time of the May 2021 NAPLAN tests. We find suggestive evidence of this from "check-in" (progress) tests administered during 2020.

Our paper is related to a sizable literature estimating the effect of school closures on test scores. Betthäuser, Bach-Mortensen and Engzell (2023) and Hammerstein et al. (2021) provide reviews and meta-analyses of the literature. We relate our findings to the literature after presentation of our results. Our main contribution is to show that learning losses from lengthy pandemic school closures are not necessarily large. We do this in a study setting that more closely approximates the ideal than existing work, providing confidence in our results.

2 Background

The power to suspend in-person learning rests with the states in Australia. Key to our study design are differences both across (and within) states in the amount of time schools were closed. The school year in Australia starts in late January. All states closed schools at the start of the pandemic in late March or early April 2020, but the return to in person learning and subsequent school closures varied (Figure S1). Table 1 reports the number of days that schools were closed between the start of 2020 to when national standardized tests took place took place, on 11 May 2021. The amount of time schools were closed for ranged from 4 days in South Australia to 112 days for a Grade 9 student located in Melbourne, Victoria. The second COVID wave in Victoria resulted in schools being closed for between 71 (Grade 3 student in primary school outside of Melbourne) to 112 days (Grade 9). In contrast, the next longest duration of school closure was 42 days. On average, students in states outside Victoria experienced 27 days of school closures.

All school closures occurred during broader regional lockdowns which imposed stay at home orders and closed non-essential retail stores and workplaces. In contrast to the international experience, COVID cases numbers were low during the period in which schools were closed; all lockdowns started when new daily infections were less than 100 and often with fewer than 10. The peak in new daily infections in any individual state was 788 (16 per 100,000 people) and the vast majority of COVID deaths were among those aged over 70 years, indicating that the health burden for school aged children and their families was low (Schurer et al., 2022). School closures occurred in the absence of widespread financial difficulty for households (Schurer et al., 2022). The provision of a generous federal wage subsidy program (JobKeeper) mitigated profit losses for businesses and kept workers employed. The government also provided a substantial income supplement to individuals relying on income-support payments.

We measure student achievement post school closures from standardized national tests that were conducted in person under normal testing conditions in May 2021. Low COVID–19 case numbers across all states in 2021 – national daily infections averaged 11 (0.04 per 100,000 people) from the start of 2021 until the date of the standardized exams – meant that students spent the vast majority of the year leading up the 2021 standardized tests in the classroom (Figure S1). Three states did close schools in the 2021 school year, however, these closures were brief lasting from between three to five days (Figure S1). This ensured that participation in national standardized tests remained high and mitigates concerns that our results reflect a change in test conditions or students being unaccustomed to being in the classroom.

The vast majority of students had access to sufficient technological resources to undertake studying from home. At least 96 percent of students reported that they had access to reliable internet, electronic devices and space for study. There was little difference in access to these resources across states and socio-demographic groups (see Appendix B.1 for more details). Hence, we can rule out that our results are being driven by differences in access to the internet, electronic devices and study space.

3 Data

3.1 School closures

We collect data on school closure dates from state government press releases and newspaper articles. We define a school as being closed for in-person learning if the government asked students to learn from home or if the duration of school holidays was extended. The latter lasted no longer than five days and primarily occurred at the start of the COVID-19 pandemic in March 2020 to allow teachers to prepare for remote learning. During the school closures, teachers assigned students lessons to complete at home and checked in on students using online platforms. Teachers used a combination of worksheets, pre-recorded videos, synchronous online lessons and online learning platforms to deliver lessons (Ziebell et al., 2020).

3.2 Standardized test scores

We measure achievement using student-level test scores from the National Assessment Program – Literacy and Numeracy (NAPLAN) tests. All students in Grades 3, 5, 7 and 9 in Australia are required to sit the NAPLAN tests. All schools in Australia, regardless of whether they are public or private, receive funding from the government. A condition of that funding is that students have to sit the NAPLAN test and school-level average NAPLAN test scores are required to be published. The NAPLAN test has taken place in the second full week of May since 2008. The only exception was in 2020 when the NAPLAN test was canceled owing to the COVID-19 pandemic.

Students sit standardized tests in reading, writing, spelling, grammar & punctuation and numeracy. The tests are calibrated to a constant level of difficulty. Within a given subject area a particular score represents the same level of achievement over time. The psychometric and scaling methods used to produce NAPLAN scores are similar to that used by the Programme for International Student Assessment. The writing test is graded by a person. Questions for all other tests have a specific answer and are graded by a computer. A transition from paper-based to online adaptive testing began in 2018 (ACARA, 2021). Scores for paper-based and online tests have been equated by the Australian Curriculum, Assessment and Reporting Authority (ACARA, 2021).

We have access to de-identified student-level NAPLAN test score data for each year from 2013 to 2019 and 2021 from the Australian Curriculum and Reporting Authority (ACARA). The dataset contains the population of students in Grades 3, 5, 7 and 9. We analyze students in Grades 5, 7 and 9 in order to make use of prior test score data, to control for cohort effects. The full dataset comprises over 9 million observations. This includes 743,305 Grade 5, 7 and 9 students who sat the NAPLAN test in 2021, of which 185,648 (or 25 percent of) students were in Victoria. We standardize test scores by the grade-level national standard deviation of test scores between 2013 and 2019.

The dataset contains information on student-level demographics including age (to one decimal place), gender, indigenous status (yes or no), language background other than English (yes or no), highest level of school education for each parent (Grade 9, Grade 10, Grade 11 or Grade 12), the level of post-school education for each parent (none, certificate, diploma, or bachelor's degree or above) and parental occupation (senior management and qualified professionals; other business managers and associate professionals; tradespeople, clerks, skilled office and sales & service staff; machine operators, hospitality staff, assistants and labourers; not in paid work in previous 12 months). We combine education and occupation categories across parents by taking the higher group of either parent. The dataset records the test mode (paper or online) for each student and we control for this in our regression analyses. There is also information on school characteristics including the state and remoteness area classification (metropolitan, inner regional or outer regional/remote) of the school, school sector (public or private) and a random school identifier. The data contains the record of all students including those who did not sit or abandoned the test; this allows us to calculate test participation rates. There are minimum standards of achievement defined for each grade level; students below minimum standard are considered at risk of being unable to progress without additional support (ACARA, 2021).

A key advantage of our setting is that participation rates for the NAPLAN test have averaged 95 percent in each year including in 2021. Table 2 reports NAPLAN participation rates for 2013-2019 and 2021 by state and grade level. There was no discernible change in participation in any state for the 2021 NAPLAN test, which took place following COVID-19 related school closures. This mitigates concerns regarding selection.

4 Empirical methodology

We use individual test score data from the 2013–2019 and 2021 NAPLAN tests and information on the number of days a student's school was closed owing to the COVID-19 pandemic to quantify the effect of school closures on student achievement. Using a difference-in-difference approach we compare the test scores of students in regions where schools were closed for a significant period of time to the test scores for students in regions where schools were closed for a shorter period of time. Our dataset allows us to use a student's previous test score (e.g. from the 2019 NAPLAN for 2021 tests) to control for differences in the ability of students across time (cohort effects).

Our baseline analysis uses two different treatment classifications. The first classifies students in Victoria as treated and students in other states as controls. The second is a continuous treatment variable equal to the number of days of learning from home in a state or state sub-region between January 2020 and the NAPLAN tests in May 2021. The mean duration of learning from home in Victoria between January 2020 and the 2021 NAPLAN tests was 95, 98 and 106 days for Grade 5, 7 and 9 students, respectively; the mean duration of learning from home in the rest of the country over the same period was 27, 27 and 28 days for Grades 5, 7 and 9 students, respectively. We also estimate separate treatment effects for all states relative to South Australia, which had only 4 days of learning from home. We look for evidence of heterogeneous effects by socio-demographic characteristics by interacting the treatment indicator with each characteristic sequentially. An event study model is used to validate our difference-in-difference design.

4.1 Estimation models

Baseline specification———We estimate the mean effect of learning from home on test scores using the difference-in-difference regression

$$s^g_{i,j,t} = \delta^g_j + \theta^g_t + \sum_g \sum_{j \neq control} \beta^g_j \left(T_j \times post\right) + \rho^g s^g_{i,j,t-2} + \sum_g \sum_k \gamma^g_k X^g_{i,j,t} + \varepsilon^g_{i,j,t} \tag{1}$$

where $s^g_{i,j,t}$ is the standardized score for student i in state j, year t and grade level g; δ^g_j is a state-by-grade fixed effect; and θ^g_t is a year-by-grade fixed effect. The coefficients of interest are β^g_j, on the interaction between the treatment variable T_j and the dummy variable *post*, which takes the value one in 2021 and zero otherwise. We estimate results for three different treatment classifications: (i) first, we classify students in Victoria in 2021 as treated and all other states as controls, in which case $T_{VIC} = 1$, $T_{j \neq VIC} = 0$ and $\beta^g_j = \beta^g$; (ii) second, we specify a continuous treatment variable, in which case T_j is equal to the number of days of learning form home between January 2020 and May 2021 and $\beta^g_j = \beta^g$; (iii) third, we estimate separate treatment effects for each state relative to South Australia (which had only 4 days learning from home), in which case $T_{j \neq SA} = 1$ and $T_{j=SA} = 0$. The term $s^g_{i,j,t-2}$ is student i's score on the same test two years prior, and is included to control for ability. The term $\sum_g \sum_k \gamma^g_k X^g_{i,j,t}$ is the set of covariates, with the coefficient on each covariate γ^g_k entering separately by grade level. The included covariates are the student-level demographic variables described earlier in Section 3.2 and test mode. $\varepsilon^g_{i,j,t}$ is an error term. We cluster standard errors at the region × school sector level. There are two regions (capital city and non-capital city) per state and school sector is either government or non-government.[1] This provides the most conservative standard errors among feasible alternative. Because the number of clusters is small, we report standard errors using the Wild Cluster Bootstrap.[2] We pool the data across grade levels, rather than estimating separate regressions for each grade level, to allow for cross correlations of errors across grade levels and cohorts. We estimate the regression separately

[1] The Australian Capital Territory (ACT) and the Northern Territory (NT) are a single region.

[2] We impose the null of no effect. We have confirmed that p-values are similar for the Wild Cluster Restricted (null imposed) and Wild Cluster Unrestricted (null not imposed) bootstrap procedures, as required if standard errors are valid with few treated clusters (MacKinnon and Webb, 2018).

for each test $s \in \{$Reading, Writing, Spelling and Grammar, Numeracy$\}$, the composite (average) score and the average of the literacy components.

Heterogeneous effects specification———We estimate heterogeneous treatment effects by characteristic k by re-specifying Equation (1) as follows:

$$s^g_{i,j,t} = \delta^{g,k}_j + \theta^{g,k}_t + \sum_k \sum_g \sum_{j \neq control} \beta^{g,k}_j \left(T_j \times post \times I_k\right) + \rho^{g,k} s^g_{i,j,t-2} + \varepsilon^g_{i,j,t} \quad (2)$$

where k is the characteristic heterogeneity of interest (e.g. male/female) and I_k is a dummy variable taking the value for characteristic k. All other terminology is the same as in Equation (1).

Event study specification———We also estimate results using an event study design, which includes lags of the treatment variable. The regression specification is

$$s^g_{i,j,t} = \delta^g_j + \theta^g_t + \sum_{t \neq 2019} \sum_g \sum_{j \neq control} \beta^g_{j,t} \left(T_j \times I_t\right) + \rho^g s^g_{i,j,t-2} + \sum_g \sum_k \gamma^g_k X^g_{i,j,t} + \varepsilon^g_{i,j,t} \quad (3)$$

where I_t is a dummy variable equal to one for year t, and 2019 is the base year. All other terms are the same as in Equation (1). Finding the estimated pre-treatment effects $\beta^g_{t \neq 2019}$ to be insignificantly different from zero provides evidence that there were no confounding pre-trends prior to the pandemic.

5 Results

We present cross-tabulation results in Section 5.1, density plots in Section 5.2 and the causal effects estimated using regression analysis in Section 5.3.

5.1 Tabulation results

For each student sitting the NAPLAN test in 2021, we calculate the test score gain as the change in test scores between the 2019 and 2021 tests. We also calculate this gain for all previous years. Table 2 shows the mean composite standardized test score gain by state and grade level. Differences in the standardized gain between the 2021 cohort and previous cohorts by state are shown in column (6) of Table 2. For Grade 5 students, the difference in state-level gains for the 2021 cohort relative to the average of previous cohorts ranged from -0.104 to -0.008 of a standard deviation, with four states experiencing smaller relative gains than Victoria despite schools being closed for a noticeably shorter period of time. For Grade 7 students, this difference ranged from -0.080 to 0.028, with five states experiencing smaller relative gains than Victoria. For Grade 9 students, the difference ranged from -0.080 to -0.029, with Victorian students experiencing the smallest gain in test scores relative to previous cohorts.

These before-after comparisons do not control for other common factors affecting all states in 2021. We calculate a difference-in-difference estimate by computing the test score gain for Victoria relative to all other states. The difference in gains between students in Victoria relative to the rest of the country was small, ranging from -0.022 to 0.001 of a standard deviation (column (8) of of Table 2). We can also compute a difference-in-difference estimate for each state relative to South Australia, which experienced the shortest duration of school closures of 4 days (column (7) of of Table 2). For Grade 5 students, the state-level difference ranged from -0.074 to 0.022, with four states reporting smaller gains than Victoria. For Grade 7 students, this difference ranged from -0.061 to 0.048, with four states reporting smaller gains than Victoria. For Grade 9 students, the difference range from -0.049 to 0.002 with Victorian students reporting the smallest gains.

5.2 Density plots

Mean effects can mask offsetting effects at different parts of the test score distribution. We use kernel density plots to look at what occurs at other parts of the distribution. Figure 1 plots the z-score of test score gains in Victoria and the other states in 2021 and an average of the years 2013-2019, by grade level. For each grade level, the distribution of test score gains in 2021 shifted left relative to earlier years by a similar amount for Victoria and other states. For Grades 5 and 7, there is little evidence that Victoria experienced a larger decline in test score gains than other states between 2021 and earlier years. For Grade 9, there is suggestive graphical evidence of a decline in Victoria compared with other states around the middle of the test score distribution.

5.3 Causal effect of school closures on test scores

Baseline results———The left panel of Figure 2 shows results when students in Victoria are classified as treated and students in other states are classified as controls. Across all grade levels, Victorian students' composite test score in 2021 was 0.02 standard deviations (se 0.02) lower than students in other states. Results are similar by grade level. The confidence internals include no learning loss at each grade level. Estimated learning loss is similar for literacy and numeracy (Figure 2). However, there is some offsetting variation within the components of literacy (Figure S2). The estimates in the right panel of Figure 2 report estimates where treatment is defined to be equal to the number of learning from home days, expressed per 100 days. The estimates are very similar, indicating modest learning losses for students in Victoria of between 0.04 (se 0.04) and 0.00 (se 0.05) standard deviations.

As a robustness check, we re-estimated our baseline results (Figure 2) including school fixed effects. The inclusion of school fixed effects controls for school-level time-invariant biases not captured by our control variables. There is negligible difference between the fixed effects model and our baseline estimates.

Event study———To look at how students in Victoria in 2021 performed relative to students in other states and relative to previous cohorts we use an event study design. The event study augments the baseline difference-in-difference model with lags of the treatment variable, that allows for placebo treatments in all prior years. The difference-in-difference model requires outcomes in the control states to provide a valid counterfactual. Figure S3 shows treatment effects relative to 2019. The estimated placebo pre-treatment effects are statistically insignificantly from the base year 2019 in all but one case. This indicates that test scores in Victoria and other states would likely have moved in parallel in 2021 absent school closures.

We find no statistically significant difference in Victorian students' test scores relative to other states in 2021 (Figure S3). Further the variation in test scores seen in Victoria in 2021 is similar to that seen in previous years (Figure S3).

State-by-state results———South Australia provides a close approximation to outcomes under no school closures because there were only 4 learning from home days in South Australia. We estimate separate treatment effects for each state relative to South Australia. Figure 3 presents the results. The treatment effect for Victoria relative to South Australia is similar to the baseline results (Figures 2 and 3). All states show some evidence of learning loss relative to South Australia, except New South Wales in Grades 5 and 7. However, there is no evidence of a dose-response relationship: Victoria is estimated to have had smaller learning loss than some other states with shorter duration of learning from home. The absence of significant pre-treatment effects for each state in an event study provides support for the study design (Figure S4).

Because students in all states experienced at least some length of school closures, a causal interpretation to our baseline estimates requires the assumption of strong parallel trends (Callaway, Goodman-Bacon and Sant'Anna, 2021). That is, Victoria would have had the same average learning loss as other states under a short duration of school closures. However, South Australia provides an untreated baseline (abstracting from the 4 days of school closures). This means we need assume only standard rather than strong parallel trends for pairwise comparisons to South Australia (Callaway, Goodman-Bacon and Sant'Anna, 2021). The similarity of the results comparing Victoria to either South Australia or to all other states provides confidence in our baseline results.

Heterogeneous effects———Learning loss could have been substantially larger for particular socio-demographic groups than the mean. We investigate this possibility by estimating separate treatment effects for different socio-demographic groups. Figure 4 shows estimates where students in Victoria are classified as treated and students in other states controls. A separate regression is estimated for each group of characteristics, including only those covariates and students' previous score as controls. There is little variation in learning loss across socio-demographic characteristics (Figures 4 and S5). Results are similar where the treatment variable is continuous and equal to the number of days of learning from home (Figure S6).

Students at risk of meeting minimum national benchmarks could have been most negatively affected by learning from home restrictions (Jack et al., 2021). Figure S7 shows the percentage point change in the share of students meeting minimum standards by test and grade level caused by learning from home restrictions. Writing for Grade 9 students is the only test for which we see evidence of a noticeable decline in the share of students meeting minimum standards. We find little effect for other grades.

5.4 Discussion

We find no evidence of large learning losses associated with school closures. Our baseline point estimates of learning loss are between 0.01 to 0.03 standard deviations for students in Victoria, who experienced considerably longer school closures, relative to those in the rest of the country. To put that in perspective, the average gain over two years between two consecutive standardized tests for a student in Grade 5, 7 and 9 are 1.2, 0.7 and 0.6 standard deviations, respectively. Our point estimates imply that the number of days of lost learning for a Grade 5, 7 and 9 student in Victoria are 10, 6 and 14 days, respectively. This is similar to the 18 days per year that a student on average was absent from school in the period prior to COVID. It is also noticeably shorter than the additional 68, 71 and 78 days a Grade 5, 7 and 9 students from Victoria spent learning from home relative to their peers in other states.

We can also benchmark our estimates against learning losses from summer holidays. Estimates at the bottom of the range suggest that during summer holidays student achievement falls by 0.001 standard deviations per day (Downey, von Hippel and Broh (2004); Kuhfeld (2019); von Hippel and Hamrock (2019)). These estimates would imply that students in Victoria had a learning loss relative to their peers in other states of between 0.07 to 0.08 standard deviations, which is around 2-3 times larger than what we observe.

Our estimated learning loss from school closures during the pandemic is towards the lower end of estimates documented in the meta-analysis of Betthäuser, Bach-Mortensen and Engzell (2023). A key difference between the Australian setting and the papers cited in the meta-analysis is that COVID case numbers were considerably lower in Australia. Over the period of our study new daily COVID cases averaged 2.31 cases per million in Australia. In contrast, the next lowest number of cases for the countries cited in the meta-analysis of Betthäuser, Bach-Mortensen and Engzell (2023) was 38.1 cases per million. For these other countries new daily COVID cases were above the maximum daily cases numbers for Australia between 48 and 75 per cent of the time (Mathieu et al., 2020).

Our result that the variation in test scores we observe post school closures is similar to previous years is consistent with results from Gore et al. (2021). Using data from NSW, the largest state of Australia, Gore et al. (2021) find no significant difference in the learning gains between the 2019 and 2020 cohorts as measured by tests administered in the beginning and end of each year.

Relationship with other methodologies———There is a growing literature assessing the effects of school closures on student learning (Werner and Woessmann, 2021; Patrinos, Vegas and Carter-Rau, 2022). One issue the literature has faced is disentangling the effects of school closures from typical cohort effects. Some studies such as Engzell, Frey and Verhagen (2021) have used panel data when comparing the test scores of the affected cohort to previous cohorts to address the bias from cohort effects. These studies do not exploit contemporaneous variation in the duration of school closures like we do but instead do a before after comparison of learning trajectories. We replicate this methodology using our dataset, comparing the performance of the affected cohort in Victoria to previous Victorian cohorts, while still controlling for previous test scores. The composite learning loss estimates are similar for each grade level (Figure S10). However, this methodology indicates substantial learning losses in numeracy and gains in literacy, in contrast to our results exploiting contemporaneous and longitudinal variation in the duration of school closures.

6 Mechanism Exploration

Why don't we find large learning losses associated with school closures? There is evidence that lost instructional time owing to extreme weather (Marcotte and Hemelt, 2008; Hansen, 2011; Miller and Hui, 2022), teacher strikes (Jaume and Willén, 2019; Belot and Webbink, 2010) and shortening the school year (Fitzpatrick, Grissmer and Hastedt, 2011; Hansen, 2011; Pischke, 2007) lowers student achievement. In these studies, students stopped learning while schools were closed. In our case students were expected to learn at home. Surveys indicate that students spent on average 4 hours a day on schooling activities while learning from home, which is only slightly less than the approximately 5 hours per day students spend in the classroom during a regular year (Bower, Lai and Van Bergen 2021; Australian Institute of Family Studies 2021).

Nevertheless, it could still be the case that learning from home is less effective than learning in a classroom. In particular there could have been learning losses associated with school closures, but these loss were temporary and unwound with the resumption of in-person schooling. Students across all states had returned to the classroom by 23 October 2020 and spent the vast majority of the 2021 schooling year prior to the May 2021 NAPLAN tests in school.

To provide evidence that any learning losses were temporary we utilize data from "check-in" tests administered to students in NSW once students were back in the classroom following the end of school closures. These check-in tests had similar format and psychometric properties to NAPLAN tests. Tests were administered to Grade 5 students for reading and numeracy and Grade 9 students for numeracy. These tests were conducted between 53 to 67 days after students would have sat a regular NAPLAN exam. We compare the standardized mean scores of Grade 5 and Grade 9 students in these check-in tests in 2020 to previous cohorts taking the NAPLAN exam (Figure S8). NAPLAN scores are categorized into six different bands. Figure S9 compares the proportion of students achieving scores in the bottom 2, middle 2 and top 2 bands from the check-in tests to that from previous NAPLAN exams. A caveat to this analysis that we are comparing different cohorts. However, mean NAPLAN scores have been relatively stable through time. The mean 2020 check-in scores were within the range of scores seen in the past. The distribution of scores was also similar (except for the Grade 9 numeracy exam which had a smaller proportion of students in the top band). But students in 2020 took the exam around 60 days later than when students in previous cohorts sat the NAPLAN exam. This provides suggestive evidence of some learning loss during school closures. However, as we have shown in Section 5, this learning loss was unwound by the time of the 2021 NAPLAN exam—which took place 128 days later. This is consistent with evidence in the literature that learning losses recede over time (Cattan et al. 2017; Sacerdote 2012;

Singh, Romero and Muralidharan 2022).[3]

Another explanation of why we don't find large learning losses is that learning from home could have become more effective the longer the duration of school closures. Models of human capital accumulation indicate that increased parental teaching effort can mitigate the negative effect of school closures (Fuchs-Schündeln et al., 2022). We find evidence of an increased substitution towards parental supervision of learning activities the longer the duration of school closures. In households where both parents worked, the probability that at least one parent reduced the amount of paid work they undertook (by reducing their hours, taking leave or quitting their job) increased by close to 10 percentage points when the duration of learning from home increased from 30 days (the median amount of time for school closures among survey respondents) to 70 days (the 90th percentile). Most of these parents chose to reduce their hours of work (see Appendix B.3 for more details). Survey evidence suggests that teaching activities supervised by parents produced work similar to that supervised by teachers (Ziebell et al., 2020).

7 Conclusion

Australia provides an excellent study setting to estimate the effect of pandemic school closures on standardized test scores. Australia successfully pursued a COVID elimination strategy until mid 2021, with states independently implementing school closures are part of broader lockdowns to suppress community transmission of COVID cases. This provides contemporaneous variation in the duration of school closures that is plausibly exogenous to the school system. Students in Victoria, Australia's second largest state, spent an additional 68 days (Grade 5) to 78 days (Grade 9) learning from home compared with students in the rest of Australia. We measured student achievement from NAPLAN tests—a common compulsory test with high participation.

Our main finding is that students in Victoria experienced small and statistically insignificant leaning losses relative to their peers in other states. The variation in test scores we observed between states post-pandemic is similar to that observed in years prior to the pandemic. Perhaps most

[3]Though others have found that these losses are persistent (Ichino and Winter-Ebmer 2004, Jaume and Willén 2019).

surprisingly, we find no evidence of large learning losses even for disadvantaged socio-economic groups. Our results demonstrate that learning losses are not necessarily large. Potential mechanisms include parental substitution for teachers in schools and catch-up in the classroom prior to standardized testing. Further work should seek to provide a deeper understanding of these mechanisms.

This paper should not be taken to imply that school closures do not have costs. It is possible that school closures lowered student achievement in domains not tested by NAPLAN. Testing took place at least one full term after students returned to schools in all states and teachers may have substituted instructional time or effort from non-tested to tested domains. Lack of face-to-face schooling may have had negative effects on students' wellbeing and social skills. Furthermore, research has found that supervision of children during school closures had costs to mothers, in the form of reduced work hours and lower mental health (Schurer et al., 2022).

References

ACARA. 2021. "NAPLAN Achievement in Reading, Writing, Language Conventions and Numeracy: National Report for 2021." Australian Curriculum, Assessment and Reporting Authority.

Australian Institute of Family Studies. 2021. "The Longitudinal Study of Australian Children." Australian Institute of Family Studies.

Belot, Michèle, and Dinand Webbink. 2010. "Do teacher strikes harm educational attainment of students?" *Labour*, 24(4): 391–406.

Betthäuser, Bastian A, Anders M Bach-Mortensen, and Per Engzell. 2023. "A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic." *Nature Human Behaviour*, 1–11.

Bower, M, J Lai, and P Van Bergen. 2021. "Parent perceptions of COVID19 Schooling from home." Macquarie School of Education.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro HC Sant'Anna. 2021. "Difference-in-differences with a continuous treatment." *arXiv preprint arXiv:2107.02637*.

Cattan, Sarah, Daniel Kamhöfer, Martin Karlsson, and Therese Nilsson. 2017. "The short-and long-term effects of student absence: evidence from Sweden."

Downey, Douglas B., Paul T. von Hippel, and Beckett A. Broh. 2004. "Are Schools the Great Equalizer? Cognitive Inequality during the Summer Months and the School Year." *American Sociological Review*, 69(5): 613–635.

Engzell, Per, Arun Frey, and Mark D Verhagen. 2021. "Learning loss due to school closures during the COVID-19 pandemic." *Proceedings of the National Academy of Sciences*, 118(17): e2022376118.

Fitzpatrick, Maria D, David Grissmer, and Sarah Hastedt. 2011. "What a difference a day makes: Estimating daily learning gains during kindergarten and first grade using a natural experiment." *Economics of Education Review*, 30(2): 269–279.

Fuchs-Schündeln, Nicola, Dirk Krueger, Alexander Ludwig, and Irina Popova. 2022. "The long-term distributional and welfare effects of Covid-19 school closures." *The Economic Journal*, 132(645): 1647–1683.

Goldhaber, Dan, Thomas J Kane, Andrew McEachin, Emily Morton, Tyler Patterson, and Douglas O Staiger. 2022. "The consequences of remote and hybrid instruction during the pandemic." National Bureau of Economic Research.

Gore, Jennifer, Leanne Fray, Andrew Miller, Jess Harris, and Wendy Taggart. 2021. "The impact of COVID-19 on student learning in New South Wales primary schools: an empirical study." *The Australian Educational Researcher*, 48(4): 605–637.

Hammerstein, Svenja, Christoph König, Thomas Dreisörner, and Andreas Frey. 2021. "Effects of COVID-19-related school closures on student achievement-a systematic review." *Frontiers in psychology*, 12: 746289.

Hansen, Benjamin. 2011. "School year length and student performance: Quasi-experimental evidence." *Available at SSRN 2269846*.

Houng, B, and M Justman. 2014. "NAPLAN Scores as Predictors of Access to Higher Education in Victoria." Melbourne Institute Working Paper 22/14.

Ichino, Andrea, and Rudolf Winter-Ebmer. 2004. "The long-run educational cost of World War II." *Journal of Labor Economics*, 22(1): 57–87.

Jack, Rebecca, Clare Halloran, James C Okun, and Emily Oster. 2021. "Pandemic schooling mode and student test scores: Evidence from US states." National Bureau of Economic Research.

Jaume, David, and Alexander Willén. 2019. "The long-run effects of teacher strikes: evidence from Argentina." *Journal of Labor Economics*, 37(4): 1097–1139.

Kogan, V, and S Lavertu. 2021. "The COVID-19 pandemic and student achiement on Ohio's Third Grade English Language Arts Assessment." Ohio State University Applied Report.

Kuhfeld, Megan. 2019. "Surprising new evidence on summer learning loss." *Phi Delta Kappan*, 101(1): 25–29.

MacKinnon, James G, and Matthew D Webb. 2018. "The wild bootstrap for few (treated) clusters." *The Econometrics Journal*, 21(2): 114–135.

Marcotte, Dave E, and Steven W Hemelt. 2008. "Unscheduled school closings and student performance." *Education Finance and Policy*, 3(3): 316–338.

Mathieu, Edouard, Hannah Ritchie, Lucas Rodes-Guirao, Cameron Appel, Charlie Giattino, Joe Hasell, Bobbie Macdonald, Saloni Dattani, Diana Beltekian, Esteban Ortiz-Ospina, and Max Roser. 2020. "Coronavirus Pandemic (COVID-19)." *Our World in Data*. https://ourworldindata.org/coronavirus.

Miller, Rebecca K, and Iris Hui. 2022. "Impact of short school closures (1–5 days) on overall academic performance of schools in California." *Scientific reports*, 12(1): 1–13.

Patrinos, Harry Anthony, Emiliana Vegas, and Rohan Carter-Rau. 2022. "An Analysis of COVID-19 Student Learning Loss."

Pischke, Jörn-Steffen. 2007. "The impact of length of the school year on student performance and earnings: Evidence from the German short school years." *The Economic Journal*, 117(523): 1216–1242.

Sacerdote, Bruce. 2012. "When the saints go marching out: Long-term outcomes for student evacuees from Hurricanes Katrina and Rita." *American Economic Journal: Applied Economics*, 4(1): 109–35.

Schurer, Stefanie, Kadir Atalay, Nick Glozier, Esperanza Vera-Toscano, and Mark Wooden. 2022. "Zero-COVID Policies: Melbourne's 112-Day Hard Lockdown Experiment Harmed Mostly Mothers." *medRxiv*.

Singh, Abhijeet, Mauricio Romero, and Karthik Muralidharan. 2022. "Covid-19 Learning Loss and Recovery: Panel Data Evidence from India." National Bureau of Economic Research.

United Nations. 2020. "Education during COVID-19 and beyond." United Nations.

von Hippel, Paul T., and Caitlin Hamrock. 2019. "Do Test Score Gaps Grow Before, During, or Between the School Years? Measurement Artifacts and What We Can Know in Spite of Them." *Sociological Science*, 6(3): 43–80.

Werner, Katharina, and Ludger Woessmann. 2021. "The legacy of COVID-19 in education." IZA Institute of Labor Economics Discussion Paper 14796.

Ziebell, Natasha, Daniela Acquaro, Cath Pearn, and Wee Tiong Sean. 2020. "Examining the impact of COVID-19 Report Summary." Melbourne Graduate School of Education.

Table 1: Number of Days Students were Required to Learn From Home: January 2020 to 11 May 2021

	Grade 5	Grade 7	Grade 9
NSW	30	30	30
Victoria: Melbourne and Mitchell Shire	103	103	112
Victoria: Rest of state, primary schools	80	80	-
Victoria: Rest of state, other school types	83	83	85
Queensland: Greater Brisbane	32	32	32
Queensland: Rest of state	29	29	29
Western Australia: Perth, Peel and South-west	27	27	27
Western Australia: Rest of state	22	22	22
South Australia	4	4	4
Tasmania: Southern Tasmania	32	32	42
Tasmania: Rest of state	31	31	41
ACT	32	37	37
Northern Territory	13	13	13

Notes: The table shows the number of days students were required to learn from home from January 2020 until the date of the 2021 NAPLAN tests by state and where applicable by region and school type within a state.

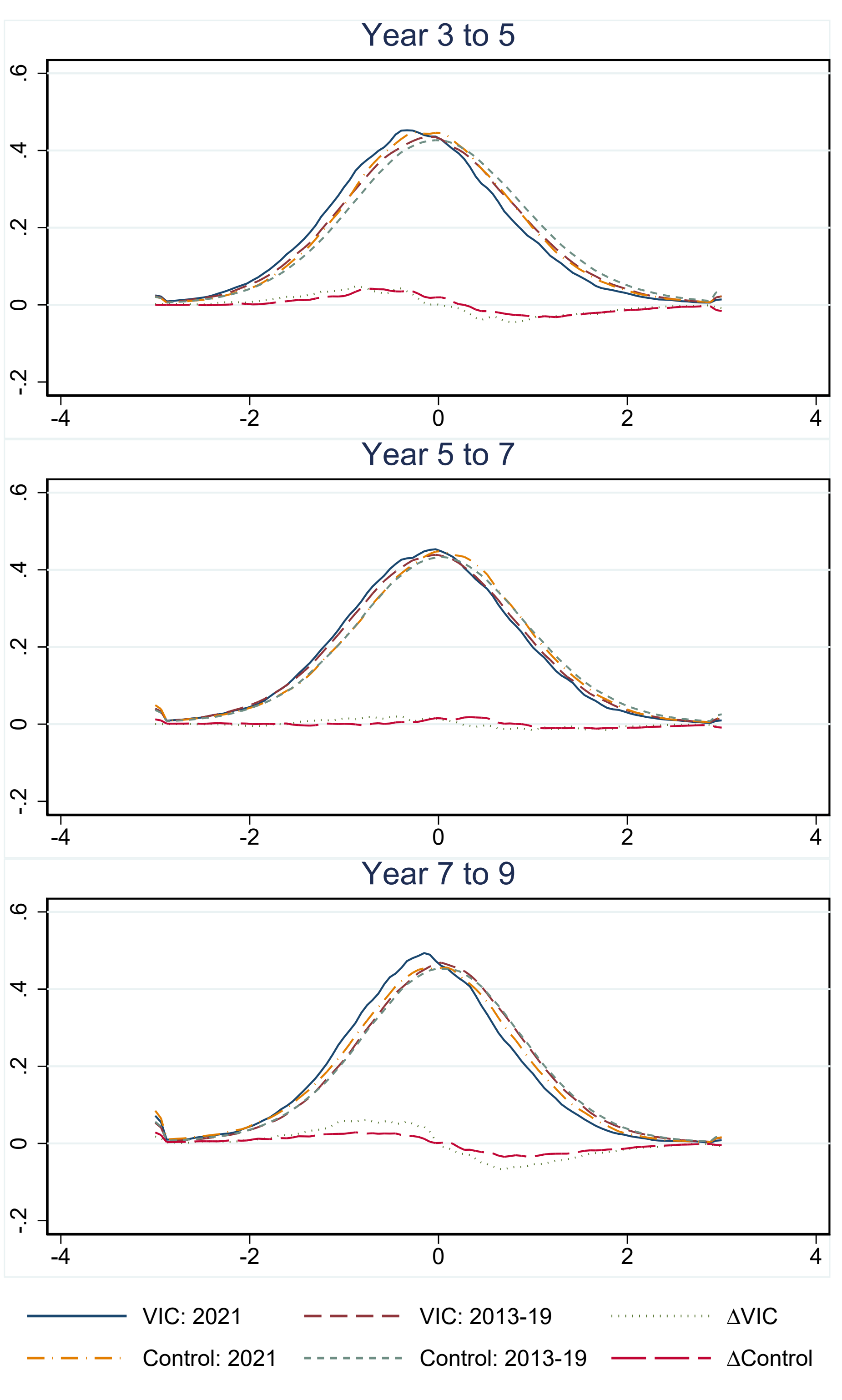
Table 2: Participation Rates and Standardized Composite NAPLAN Score

	N: 2021	Participation 2013-2019	Participation 2021	Δ Score: 2021	Δ Score: Avg. 2013-2019	Time difference	Difference rel. to SA	Difference rel.to ex. VIC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Grade 5								
VIC: 95d	64,523	0.96	0.96	1.097	1.160	-0.063	-0.033	-0.011
ACT: 32d	4,637	0.96	0.96	1.137	1.156	-0.020	0.010	
TAS: 32d	5,271	0.96	0.97	1.061	1.164	-0.103	-0.073	
QLD: 31d	54,762	0.95	0.95	1.150	1.255	-0.104	-0.074	
NSW: 30d	84,867	0.98	0.98	1.168	1.176	-0.008	0.022	
WA: 26d	29,632	0.97	0.97	1.178	1.269	-0.091	-0.061	
NT: 13d	1,656	0.91	0.91	1.232	1.326	-0.094	-0.064	
SA: 4d	15,516	0.95	0.95	1.195	1.225	-0.030		
Total ex. VIC: 27d	196,341	0.96	0.96	1.164	1.216	-0.052		
Total	260,864	0.96	0.96	1.147	1.202	-0.056		
Grade 7								
VIC: 98d	62,694	0.97	0.97	0.666	0.684	-0.018	0.002	0.001
ACT: 37d	4,359	0.96	0.96	0.731	0.704	0.028	0.047	
TAS: 32d	5,348	0.96	0.96	0.683	0.709	-0.026	-0.007	
QLD: 31d	51,900	0.93	0.93	0.672	0.752	-0.080	-0.061	
NSW: 30d	81,716	0.98	0.98	0.706	0.678	0.028	0.048	
WA: 26d	28,333	0.97	0.97	0.754	0.803	-0.049	-0.029	
NT: 13d	1,702	0.89	0.89	0.716	0.755	-0.039	-0.019	
SA: 4d	13,862	0.95	0.95	0.833	0.852	-0.020		
Total ex. VIC: 27d	187,220	0.96	0.96	0.713	0.732	-0.019		
Total	249,914	0.96	0.96	0.701	0.720	-0.020		
Grade 9								
VIC: 106d	58,431	0.94	0.94	0.489	0.568	-0.080	-0.049	-0.022
ACT: 37d	4,786	0.92	0.92	0.507	0.576	-0.069	-0.038	
TAS: 42d	4,178	0.93	0.93	0.461	0.538	-0.077	-0.046	
QLD: 31d	45,962	0.90	0.89	0.465	0.523	-0.058	-0.027	
NSW: 30d	76,770	0.96	0.95	0.513	0.581	-0.069	-0.038	
WA: 26d	27,465	0.96	0.96	0.634	0.672	-0.039	-0.008	
NT: 13d	1,458	0.84	0.83	0.552	0.581	-0.029	0.002	
SA: 4d	13,477	0.92	0.92	0.505	0.536	-0.031		
Total ex. VIC: 28d	174,096	0.93	0.93	0.517	0.575	-0.058		
Total	232,527	0.93	0.93	0.510	0.574	-0.063		

Notes: *N: 2021* is the number of students taking all NAPLAN tests in 2021. Participation is the fraction of students taking a NAPLAN test, by state and grade level; exempt students (those with a significant disability and migrants within the past year from non-English speaking country) are excluded from calculation of participation rates. Δ *Score 2021* is the student-level mean test score gain in standard deviation units between the 2019 and 2021 NAPLAN tests, by state and grade level; Δ *Score Avg. 2013-2019* is the student-level mean gain in standard deviation units between 2013 and 2019. *Time difference* is equal to Δ *Score 2021* minus Δ *Score Avg. 2013-2019*. *Difference rel. to SA* is the *Time Difference* for each state/region minus the *Time difference* for South Australia (SA); it is the difference-in-difference estimate for each state shown where South Australia is the control state. *Difference rel. to ex-VIC* is the *Time Difference* for Victoria minus the *Time difference* for Australia excluding Victoria; it is the difference-in-difference estimate where the control group is all states ex-Victoria. Number of learning from home days is shown next to each state by grade.

Summary: Participation rates remained high in 2021. The difference-in-difference estimate of the learning gain in Victoria relative to other states is -0.011 standard deviations (s.d.) for Grade 5, 0.001 s.d. for Grade 7 and -0.022 s.d. for Grade 9.

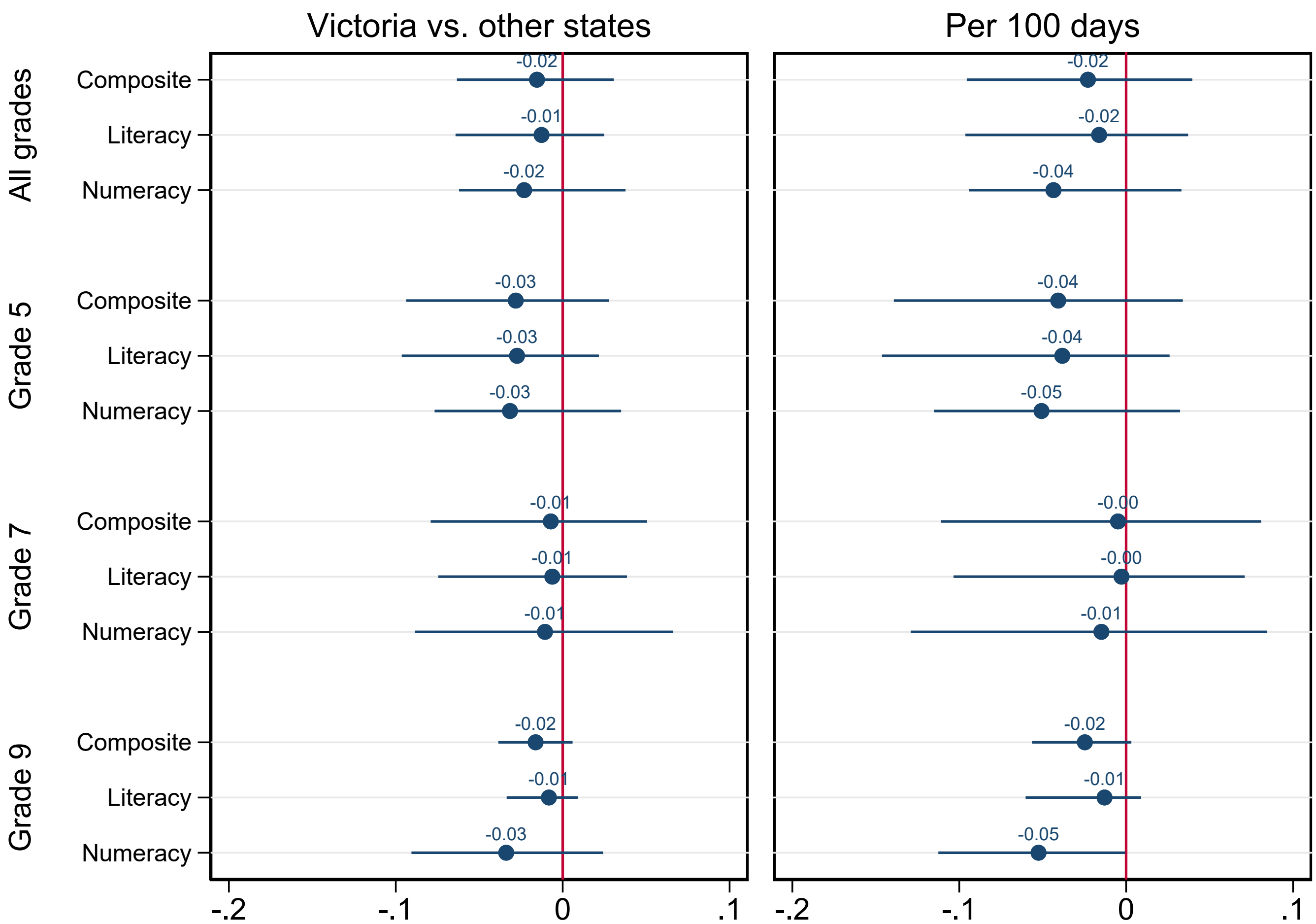
Figure 1: Kernel Density: z-score of Change in Composite NAPLAN Test



Notes: Each figure shows the kernel density of z-scores for student test score gains. *VIC 2021* is the density of test score gains between 2019 and 2021; *VIC 2013-19* is the density of test score gains for tests taken between 2013 and 2019 and Δ *VIC* is the difference between the densities: *VIC: 2021* minus *VIC: 2013-19*. *Control: 2021, Control: 2013-19* and Δ *Control* are analogous densities for students in all other states.

Summary: There is little graphical evidence of greater learning loss in Victoria across the distribution of test score gain z-scores.

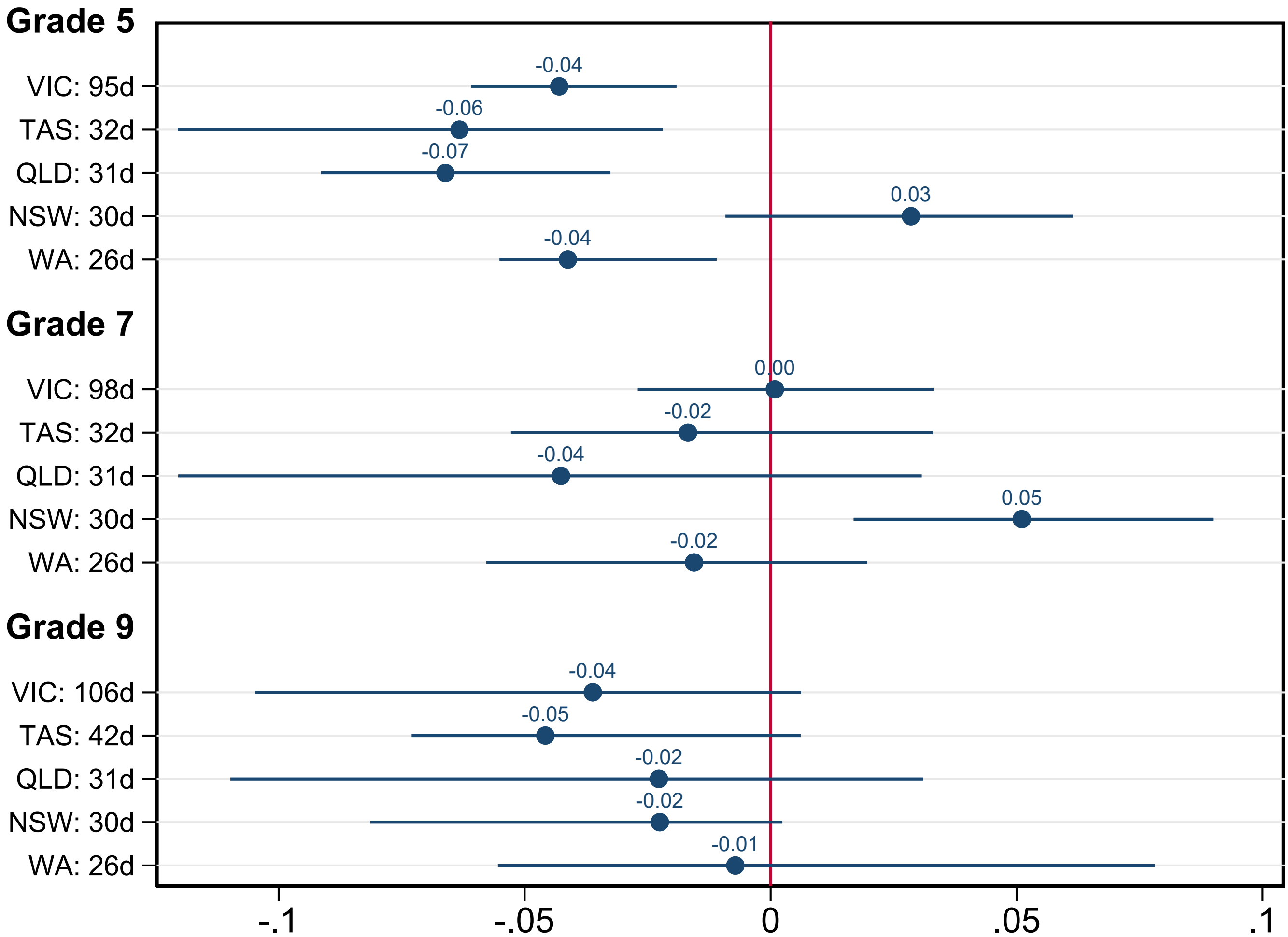
Figure 2: Baseline results



Notes: The figure shows estimated β coefficients for Equation (1). The left panel shows results where Victoria is classified as treated and other states classified as controls. The right panel reports results with a continuous treatment variable equal to the number of days of learning from home between January 2020 and May 2021; results in the right panel are scaled by 100.

Summary: There is evidence of only modest learning losses from Victoria's learning from home restrictions. The learning from home period in Victoria was 68, 71 and 78 days longer than the average for the other states for Grade 5, 7 and 9 students respectively.

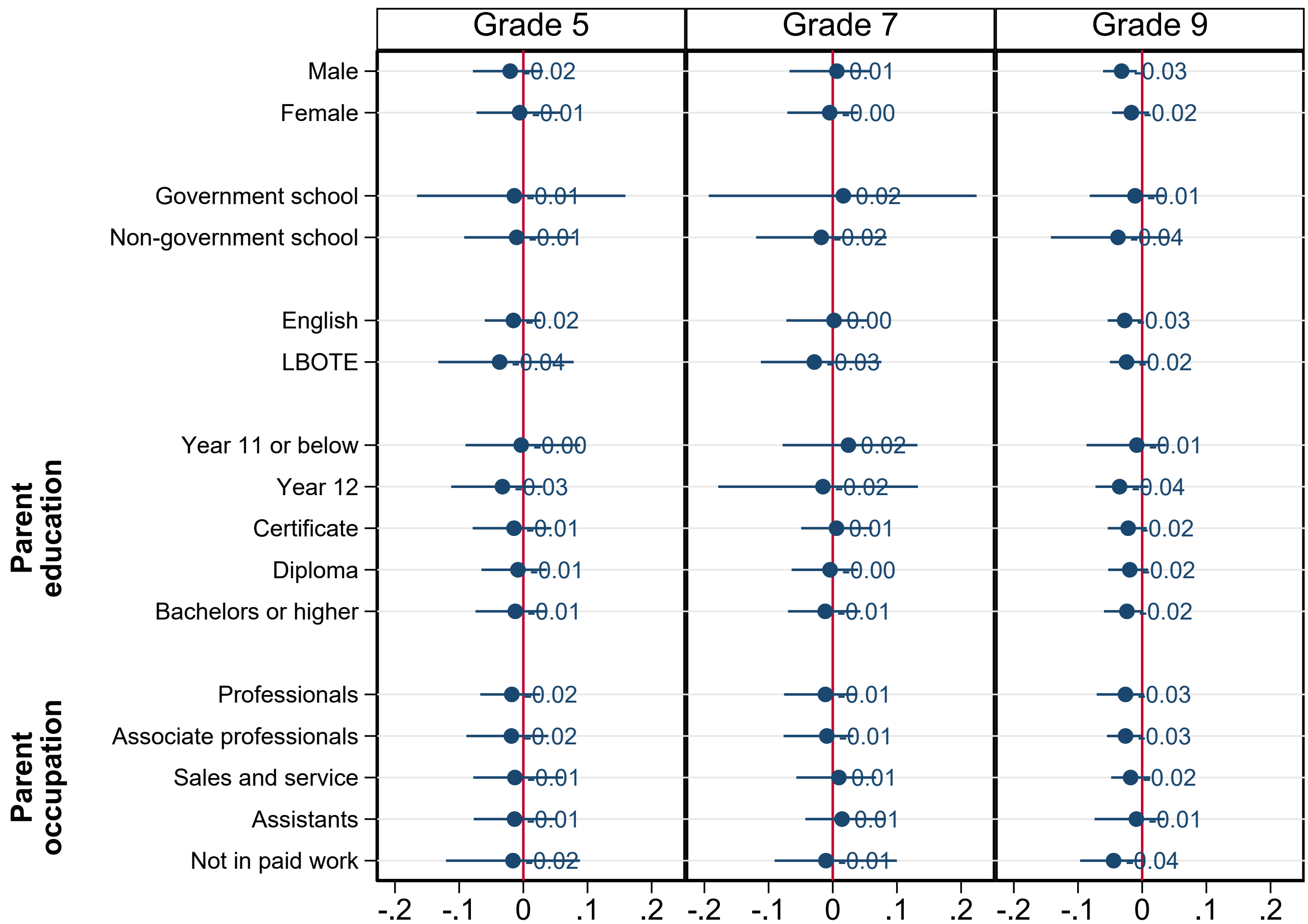
Figure 3: Results by State



Notes: The figure shows estimated β coefficients for Equation (1). Separate treatment effects are estimated for each state relative to South Australia.

Summary: There is little evidence that Victoria experienced larger learning losses than other states with much shorter periods of learning from home.

Figure 4: Heterogeneity



Notes: The figure shows estimated β coefficients for Equation (2). Victoria is classified as treated and other states classified as controls. LBOTE is students with Language Background Other Than English.

Summary: There is little evidence of differences in learning loss by student and parent socio-demographic characteristics.

Supplement

A Appendix A: HILDA Survey

This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey, conducted by the Melbourne Institute of Applied Economic and Social Research on behalf of the Australian Government Department of Social Services (DSS) (Wave 20, ADA Dataverse.) The findings and views reported in this paper, however, are those of the authors and should not be attributed to the Australian Government, the DSS, or the Melbourne Institute. The data used are available free of charge to researchers through the National Centre for Longitudinal Data Dataverse at the Australian Data Archive (https://dataverse.ada.edu.au/dataverse/ncld). Access is subject to approval by the Australian Government Department of Social Services and is conditional on signing a license specifying terms of use.

B Appendix B: Additional Results

B.1 Access to technology and study space while learning from home

We measure student's access to technology and study space during school closures using data from the Longitudinal Study of Australian Children (LSAC). The LSAC data contains responses from a representative sample of 1,300 individuals aged 16 to 17 when schools were closed. We use responses to questions in wave 9C1 which asked: "Please think about the period when restrictions were first at their peak. For most people this would have been between March and May 2020. During the coronavirus restriction period, how often did you have the following? Reliable internet access for all my needs". Similar questions were asked about access to sufficient electronic devices and study space. The five available responses were: never, rarely, sometimes, often and always. We classify an individual as having sufficient access to the internet, electronic devices or space if they answered sometimes, often or always. 96, 98 and 98 per cent of respondents reported having sufficient access to reliable internet, electronic devices and space, respectively.[4] There was very little difference in access to these resources for students in Victoria compared to the other states (Table S3).

We also look at access across different socio-demographic groups by running the following regression separately for access to each resource:

$$r_i = \sum_k \beta_k I_{i,k} + \varepsilon_i \tag{B.1}$$

where r_i is a dummy variable that is equal to one if individual i has sufficient access to either reliable internet, electronic devices or study space and k is the characteristic heterogeneity of interest (e.g. highest level of parental education) and I_k is a dummy variable taking the value for characteristic k, where characteristics are defined the same way as in Section 4. For each socio-demographic characteristics we test whether all the β_k coefficients are equal. Table S4 reports the associated F-statistics and p-values. We find no difference in access to reliable internet, electronic devices or study space across any of the socio-demographic characteristics or by geographic location.

The LSAC surveyed children aged 16 to 17. One concern in using the responses from this survey could be that these children might have access to better resources because they were closer to finishing school compared to younger children. However, we believe that the results from this study can generalize to younger children. Firstly, around half of all survey respondents had younger siblings living in the same house. Secondly, internet penetration rates in Australia are high—close to 90 per cent of the population has access to the internet and internet speeds are fast (https://datareportal.com/reports/digital-2020-australia). Thirdly, data from Programme for International Student Assessment (PISA) is consistent with estimates from the LSAC. The PISA study which was conducted in 2018 found that 98, 98 and 88 per cent of 15 year old students in Australia had access to internet, electronic devices and a quiet place to study at home respectively. Two-thirds of students reported having access to three or more electronic devices at home.

[4] If we classify individuals as having access to sufficient resources if they answered often or always then 88, 95 and 90 per cent of respondents reported having sufficient access to reliable internet, electronic devices and space.

B.2 Check-in assessments in NSW

To assess whether there any temporary learning losses associated with school closures we utilize data from "check-in" tests administered in NSW. These check-in tests were conducted in the classroom between 50 to 64 days after school closures ended. Tests were given to Grade 5 students for reading and numeracy and Grade 9 students for numeracy. The test had a similar format and psychometric properties to NAPLAN tests. For each test we have access to average check-in scores and the distribution of scores across bands from New South Wales Department of Education. The check-in tests were administered to students in government schools. 59,618 and 34,560 Grade 5 and 9 students respectively sat the check-in exam. Participation rates for Grade 5 and 9 students were 86 and 61 per cent respectively. When calculating an average check-in test score individual student scores were weighted by student's performance in prior NAPLAN tests and by geographic location to arrive at population estimates for government school students.

B.3 Parental supervision of learning activities

We measure parental supervision of learning activities during school closures using data from Household, Income and Labour Dynamics in Australia (HILDA) Survey, a nationally representative survey of Australian households. We use responses from wave 20 of the survey and make use of two questions, the first, "Did children staying home from school have any impact on your ability to undertake paid work?" and the second "And what about other members of this household? Did children staying home from school have any impact on their ability to undertake paid work?" Impact on parent's ability to undertake paid work is measured by the parent having to take either paid or unpaid leave, reducing their work hours or quitting their job. Our sample consists of responses from 4,127 parents. Table S5 shows the breakdown of responses by household members for duration of school closures of less than 30 days (the median duration of school closures) and more than 70 days (the 90th percentile) for households where both parents worked. We find that the probability of an effect on at least one parent's ability to undertake paid work increases by 10 percentage points from 34 to 44 per cent as the length of school closure increases from less than 30 days to more than 70 days.

Table S1: Test Score Summary Statistics: 2013-2019

	N	Mean	p10	p50	p90
Grade 5					
NSW	625,900	7.73	6.43	7.72	9.03
VIC	460,983	7.80	6.66	7.78	8.98
QLD	404,372	7.57	6.32	7.59	8.81
WA	211,199	7.58	6.29	7.61	8.84
SA	125,838	7.46	6.27	7.48	8.65
TAS	41,356	7.47	6.17	7.49	8.73
ACT	32,408	7.78	6.60	7.79	8.94
NT	19,335	6.75	4.72	6.99	8.42
Ex. VIC	1,460,408	7.62	6.34	7.64	8.90
Grade 7					
NSW	595,420	8.38	7.09	8.36	9.71
VIC	440,645	8.42	7.25	8.40	9.63
QLD	364,652	8.26	7.03	8.26	9.49
WA	201,149	8.32	7.04	8.33	9.58
SA	122,637	8.25	7.07	8.25	9.45
TAS	39,835	8.16	6.90	8.17	9.41
ACT	32,492	8.49	7.30	8.50	9.68
NT	17,548	7.46	5.48	7.68	9.15
Ex. VIC	1,373,733	8.31	7.05	8.31	9.60
Grade 9					
NSW	563,262	8.92	7.64	8.92	10.25
VIC	404,518	8.95	7.78	8.94	10.16
QLD	335,199	8.76	7.55	8.77	9.99
WA	191,669	8.94	7.71	8.97	10.17
SA	115,004	8.75	7.56	8.77	9.96
TAS	37,746	8.67	7.42	8.7	9.93
ACT	29,474	9.05	7.85	9.08	10.26
NT	14,831	8.15	6.25	8.35	9.77
Ex. VIC	1,287,185	8.86	7.60	8.86	10.14

Notes: The table shows sandardized NAPLAN scores for the period 2013-2019. Scores are standardized by the grade-level national standard deviation over the period 2013-2019. *Ex. VIC* is all states other than Victoria.

Table S2: Means by State

	Female	Non-Govt.	LBOTE	Bachelors	Professional	Indigenous	Metro
NSW	0.49	0.38	0.29	0.38	0.50	0.06	0.75
VIC	0.49	0.40	0.28	0.40	0.50	0.02	0.76
QLD	0.49	0.35	0.13	0.32	0.45	0.08	0.66
WA	0.49	0.37	0.20	0.34	0.48	0.07	0.76
SA	0.49	0.37	0.15	0.31	0.45	0.04	0.73
TAS	0.49	0.34	0.06	0.25	0.42	0.09	0.82
ACT	0.49	0.46	0.22	0.53	0.63	0.03	1.00
NT	0.49	0.31	0.42	0.23	0.33	0.43	0.00
Ex. VIC	0.49	0.37	0.21	0.35	0.48	0.07	0.72

Notes: The table shows means over the period 2013-2019 for students in Grades 5, 7 and 9. *Ex. VIC* is all states other than Victoria. *Non-Govt.* is non-government school, *LBOTE* is language background other than English, *Bachelors* is an indicator for either parent having a Bachelors degree or higher, *Professional* is an indicator for either parent's occupation being professional or associate professional, *Indigenous* is Aboriginal or Torres Strait Islander, and *Metro* is metropolitan region.

Table S3: Percentage of Respondents Reporting Sufficient Access to Resources to Learn From Home

	Reliable Internet	Electronic Devices	Study Space
Whole sample	95.8	97.8	97.0
Victoria	95.4	98.4	97.6
Other states (ex. VIC)	96.0	97.6	96.7

Notes: The table shows the percentage of respondents reporting that they had sufficient access to different resources while learning from home. The data is sourced from the LSAC.

Summary: The vast majority of students had access to sufficient resources to study from home. There was little difference in access to these resources between students in Victoria and those in other states.

Table S4: Heterogeneity: Sufficient Access to Resources to Learn From Home
F-statistics for test that coefficients are equal, *p*-values in brackets

	Reliable Internet	Electronic Devices	Study Space
Gender	0.534	0.000	0.024
	(0.465)	(0.996)	(0.876)
Govt/Non-govt school	0.040	1.470	0.225
	(0.842)	(0.226)	(0.635)
Language background	0.134	0.003	0.001
	(0.714)	(0.957)	(0.978)
Parent education	0.223	1.318	0.966
	(0.926)	(0.261)	(0.425)
Parent occupation	0.593	1.381	0.453
	(0.620)	(0.247)	(0.715)
Indigenous status	0.613	0.313	0.440
	(0.434)	(0.576)	(0.507)
Metro/Non-metro location	1.698	2.276	2.046
	(0.193)	(0.132)	(0.153)
State	0.852	0.525	0.880
	(0.530)	(0.790)	(0.509)

Notes: We test whether there are difference in access to reliable internet, electronic devices and space to learn from home by socio-demographic characteristics. The table reports F-statistics and p-values in parentheses associated with test that the β_k coefficients in Equation (B.1) are jointly equal.

Summary: There was little difference in access to resources across states and socio-demographic groups.

Table S5: Did School Closures Affect Parent's Ability to Undertake Paid Work
Percentage of Response by Category

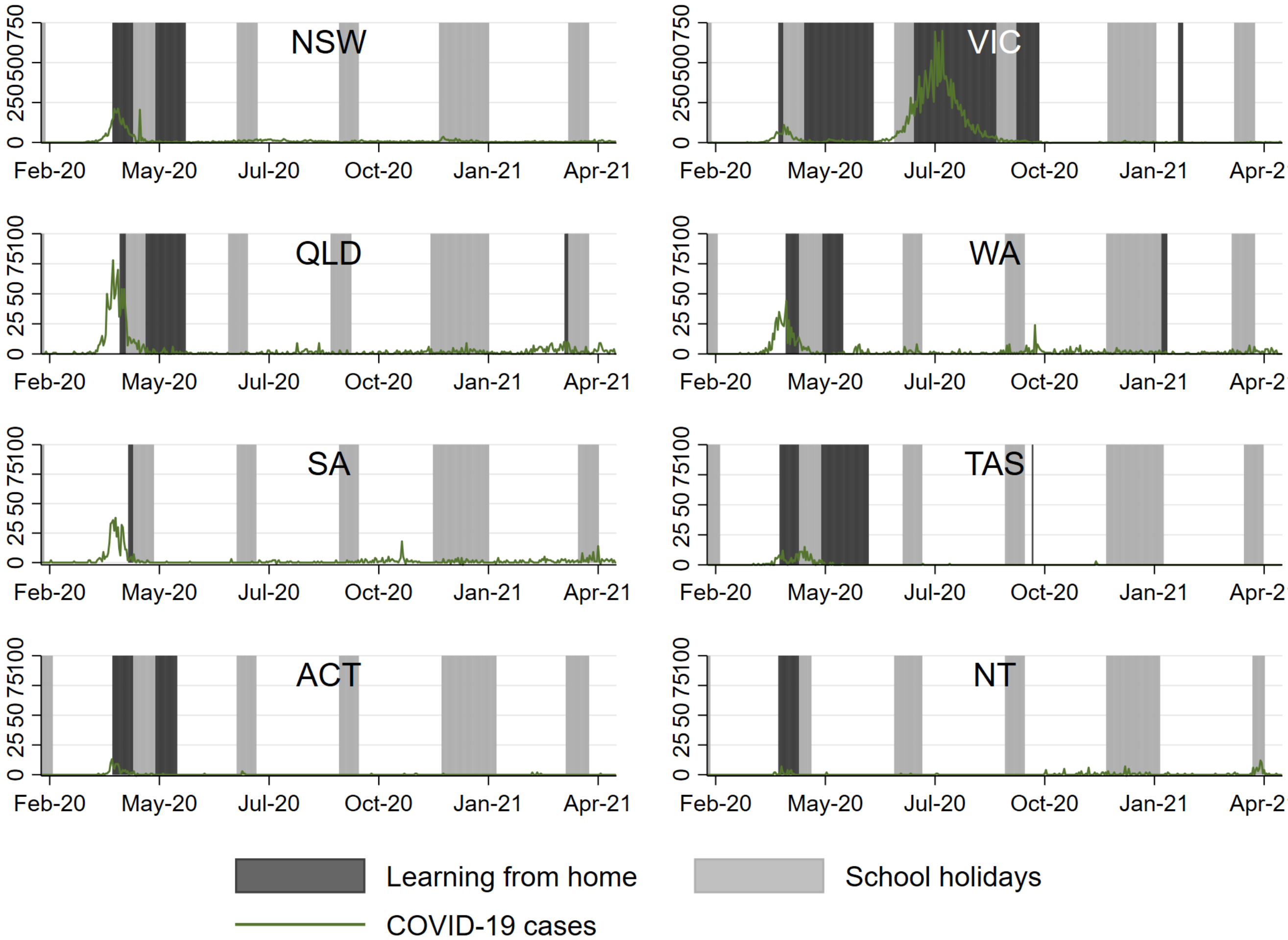
		School closure length: <30 days (Median length of school closures)	
		Parent 1	
		Yes	No
Parent 2	Yes	10.0	19.6
	No	3.9	66.5

		School closure length: >70 days (90th percentile of school closure length)	
		Parent 1	
		Yes	No
Parent 2	Yes	14.9	23.6
	No	4.1	57.3

Notes: The percentage of responses from parents by category to the question: "Did children staying home from school have any impact on your ability to undertake paid work?". Conditional on both parents initially being in paid employment. Impact on the ability of parents to undertake paid work is measured by parents having to take paid or unpaid leave, reducing their work hours or quitting their job. The data is sourced from HILDA Wave 20 Household Questions 33, 35-38.

Summary: The probability that there was a reduction in the ability of at least one parent to undertake paid work increased by close to 10 percentage points when the duration of learning from home increased from 30 days (the median amount of time schools were closed for) to 70 days (the 90th percentile).

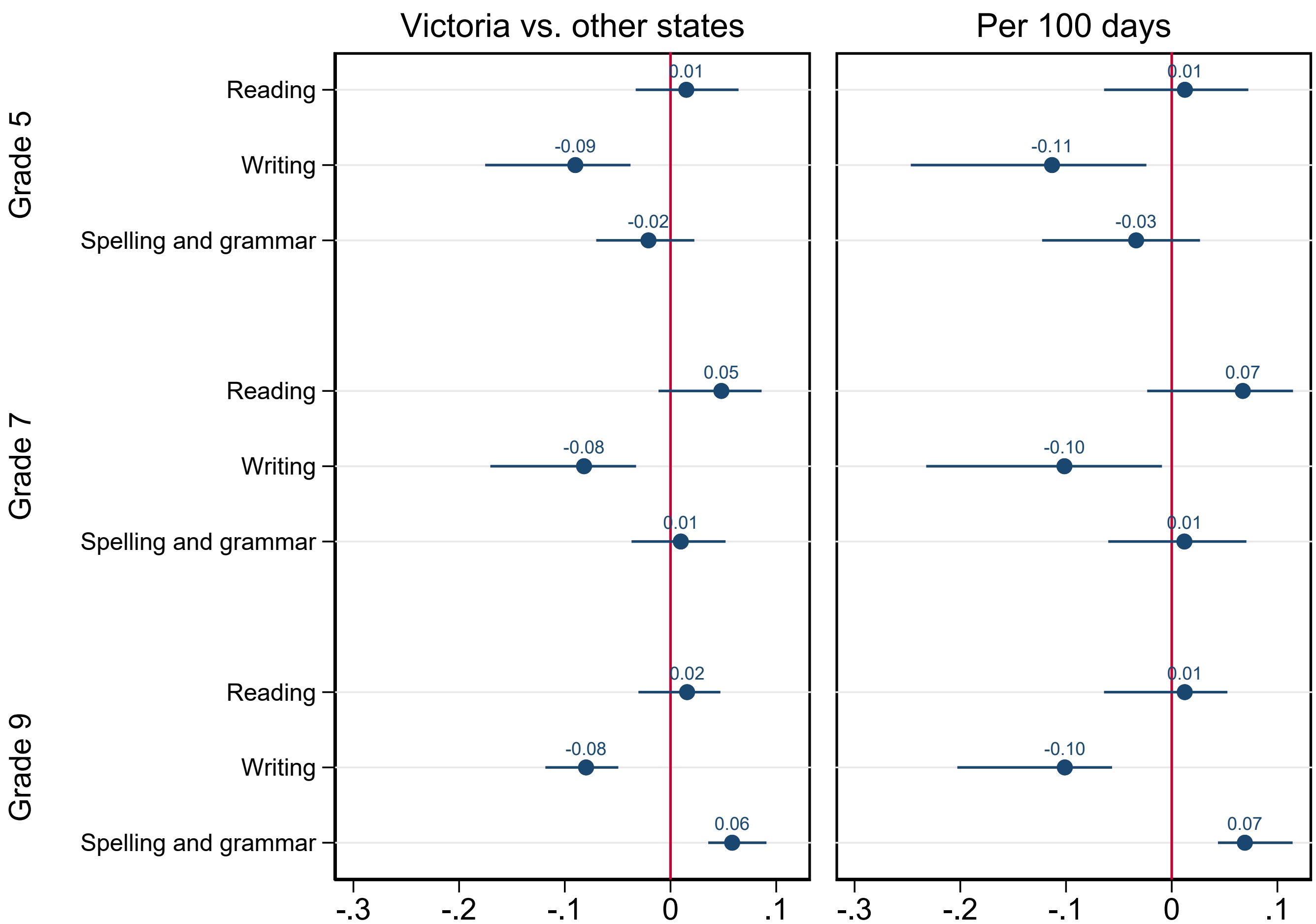
Figure S1: COVID-19 Cases and Learning from Home



Notes: This figure shows state level daily COVID-19 case numbers from January 2020 to 11 May 2021, the date when the 2021 NAPLAN exams took place. The shaded regions indicate when students were required to learn from home and school holidays.

Summary: COVID case numbers were extremely low, even during learning from home periods, because Australia pursued a zero-COVID policy over the period shown.

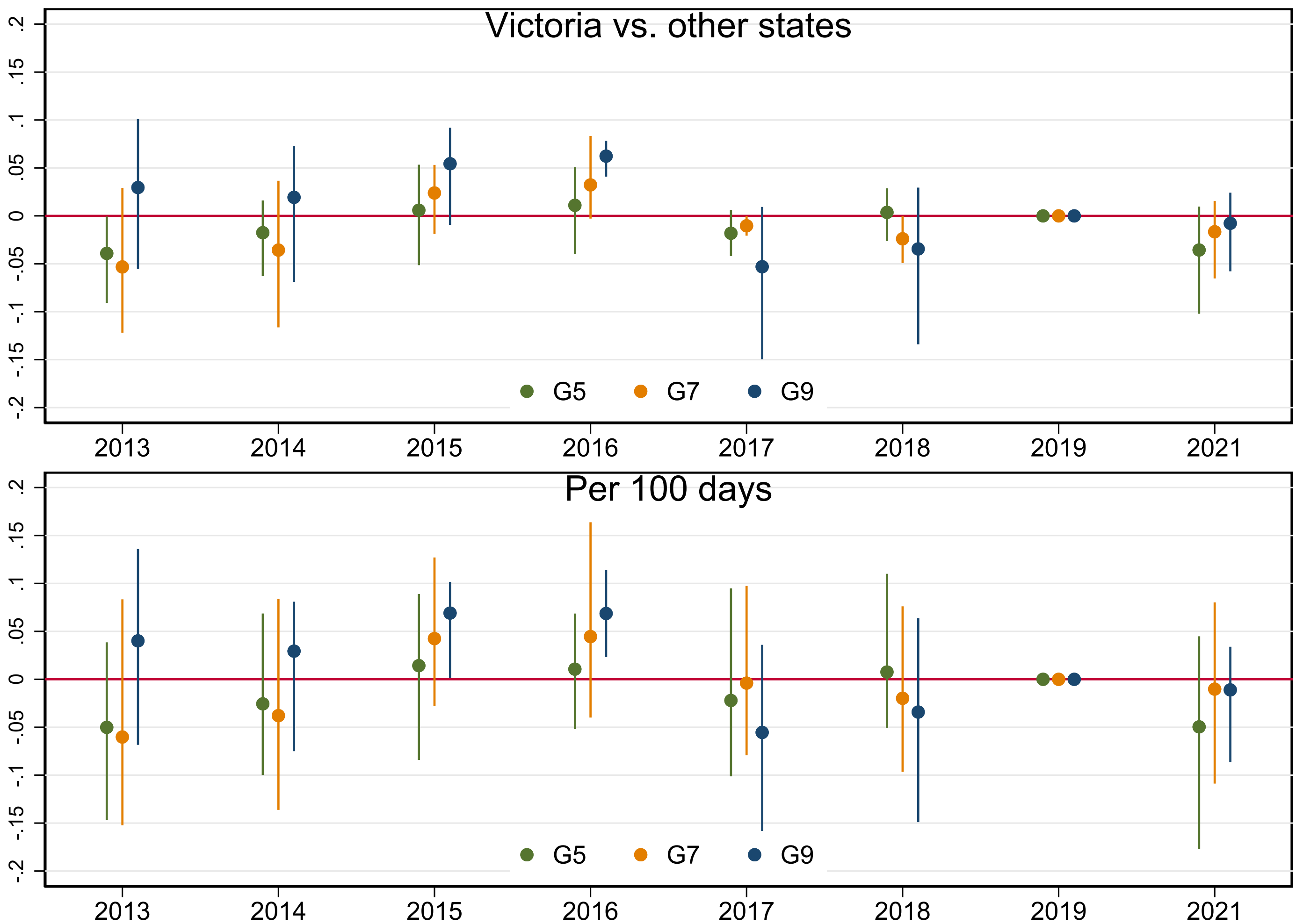
Figure S2: Components of Literacy



Notes: The figure shows estimated β coefficients for Equation (1) for each literacy test. The left panel shows results where Victoria is classified as treated and other states classified as controls. The right panel reports results with a continuous treatment variable equal to the number of days of learning from home between January 2020 and May 2021; results in the right panel are scaled by 100.

Summary: There is some offsetting variation within the components of literacy.

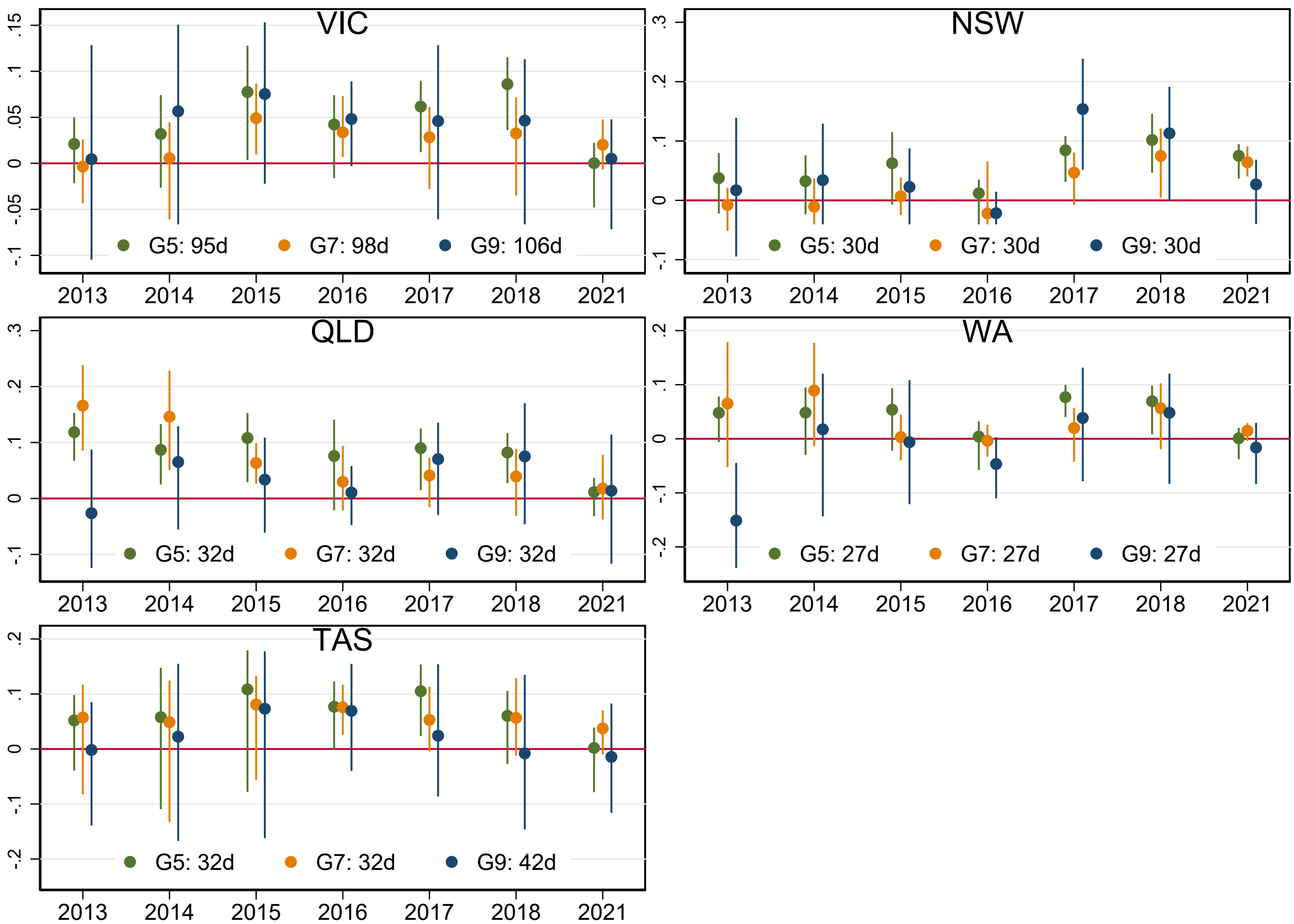
Figure S3: Event Study: Victoria vs. Other States



Notes: The figure shows estimated β_t coefficients for Equation (3) for the composite test score. The top panel shows results where Victoria is classified as treated and other states classified as controls. The bottom panel reports results with a continuous treatment variable equal to the number of days of learning from home between January 2020 and May 2021; results in the bottom panel are scaled by 100.

Summary: There is not any statistically significant difference in Victorian student test scores relative to other states in 2021. The variation in test scores seen in Victoria in 2021 is similar to that seen in previous years.

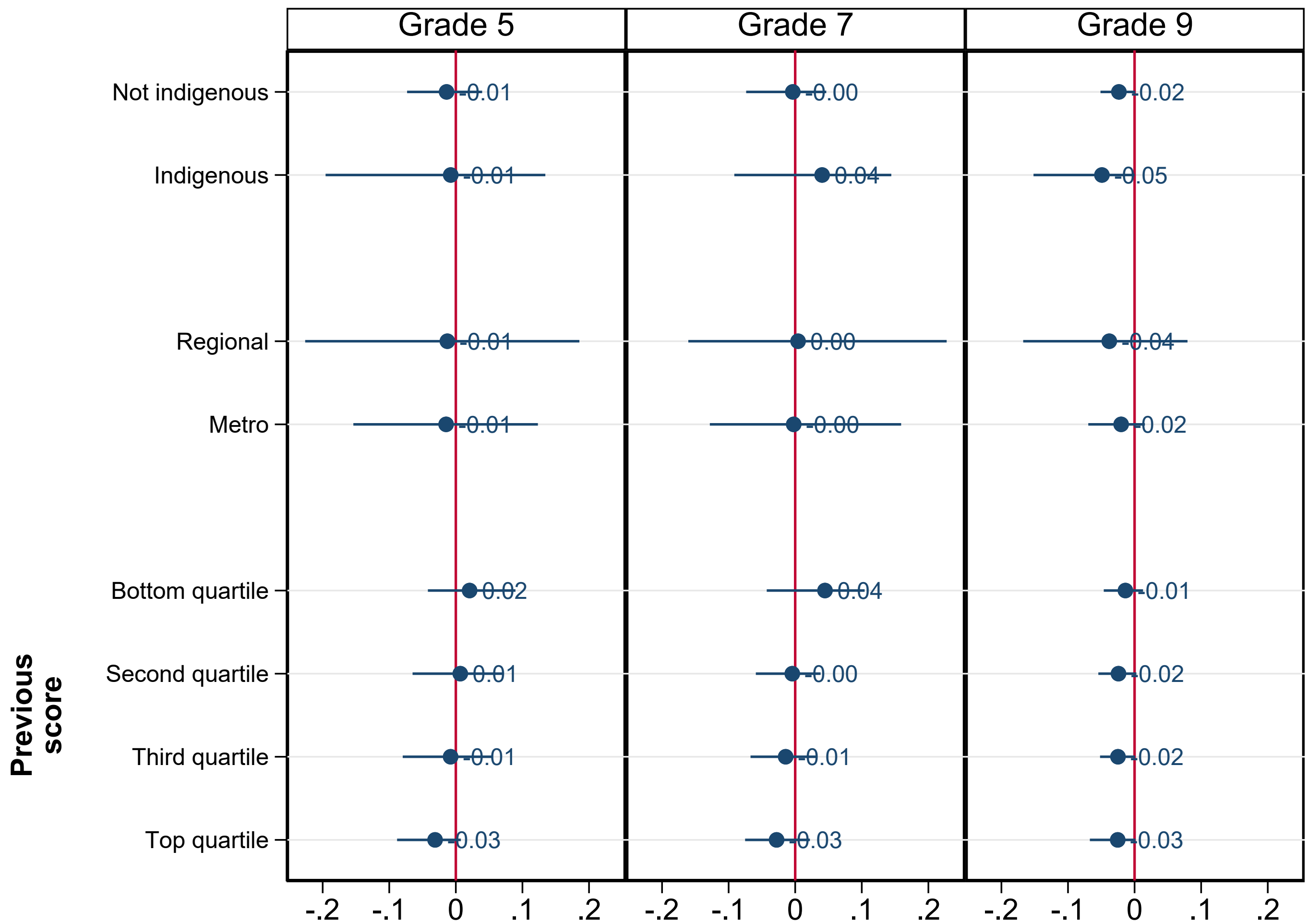
Figure S4: Event Study: Relative to South Australia



Notes: The figure shows estimated $\beta_{j,t}$ coefficients for Equation (3) for the composite test score. Separate treatment effects are estimated for each state relative to South Australia.

Summary: There is little evidence of significant pre-treatment effects.

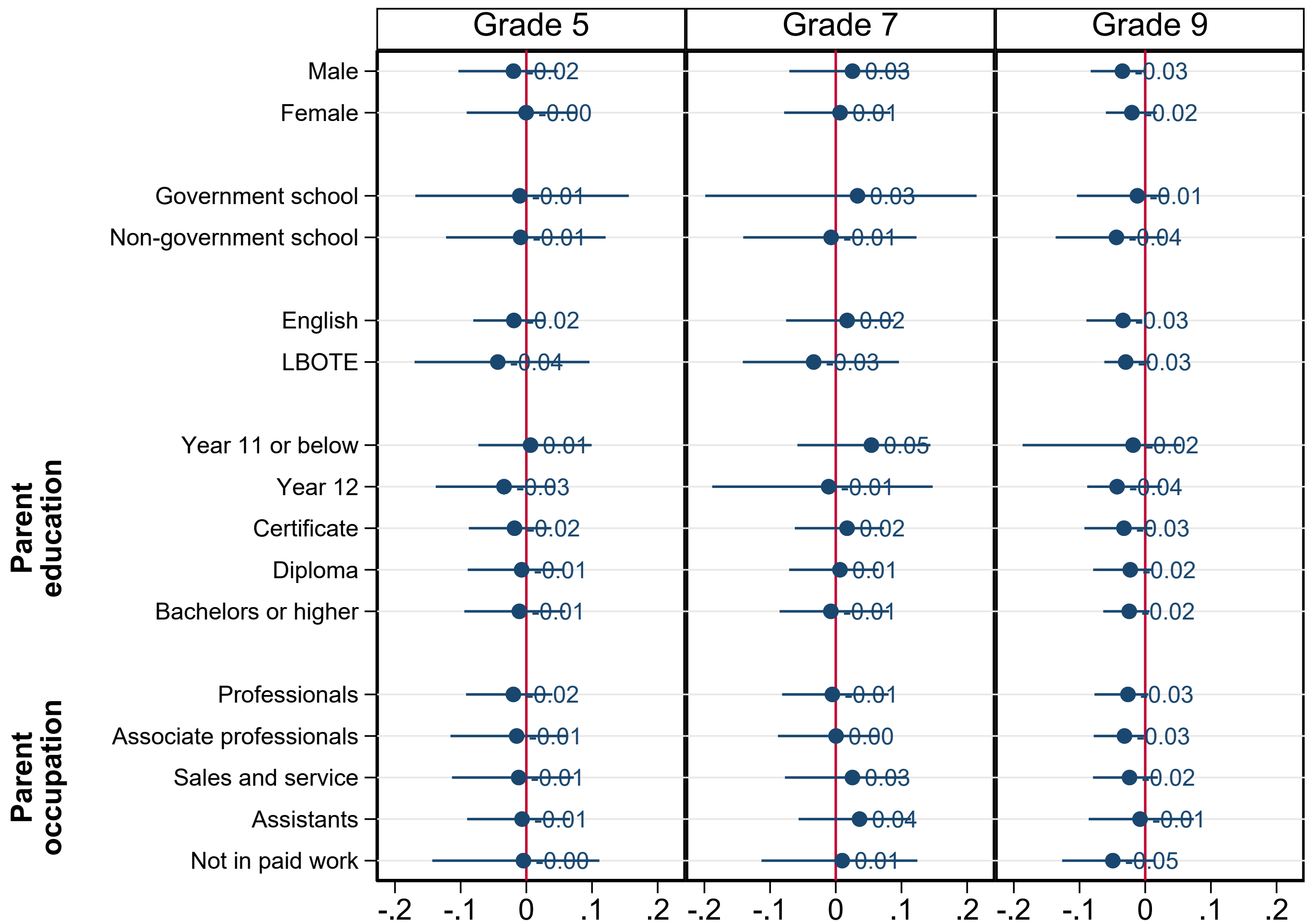
Figure S5: Additional Heterogeneity



Notes: The figure shows estimated β coefficients for Equation (2). Victoria is classified as treated and other states classified as controls.

Summary: There is little evidence of differences in learning loss by indigenous status, between regional and metropolitan areas and by previous NAPLAN score.

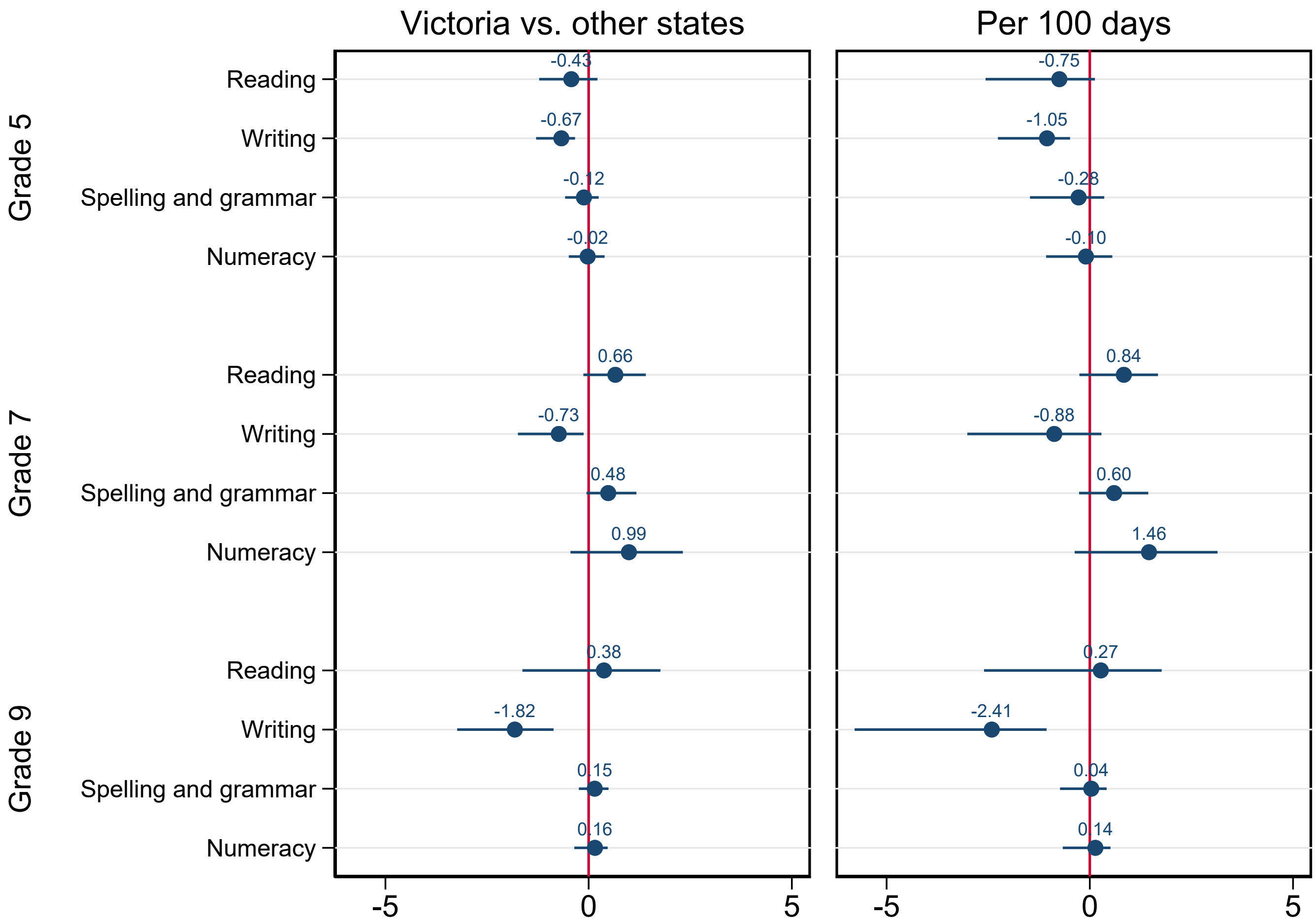
Figure S6: Heterogeneity: Per 100 days of Learning from Home



Notes: The figure shows estimated β coefficients for Equation (2). The treatment variable is equal to the number of days of learning from home between January 2020 and May 2021; results are scaled by 100.

Summary: There is little evidence of differences in learning loss by student and parent socio-demographic characteristics.

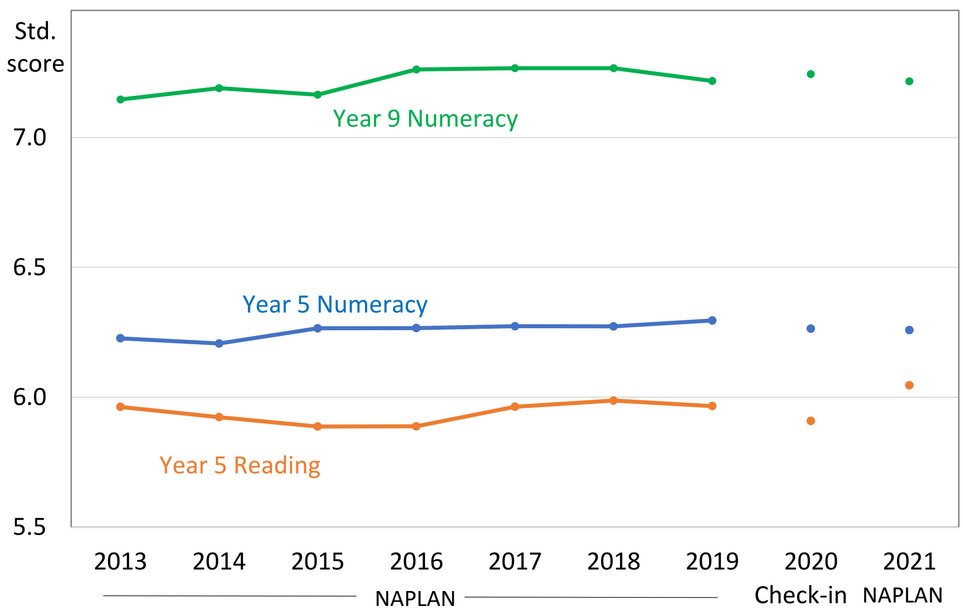
Figure S7: Meeting Minimum Standards



Notes: The figure shows estimated β coefficients for Equation (1) where the dependent variable is replaced by an indicator equal to 1 if the student is meeting minimum national standards at their grade level. The left panel shows results where Victoria is classified as treated and other states classified as controls. The right panel reports results with a continuous treatment variable equal to the number of days of learning from home between January 2020 and May 2021; results in the right panel are scaled by 100.

Summary: The extended learning from home period in Victoria did not cause a decline in the share of students meeting minimum learning standards, with the exception of Grade 9 writing.

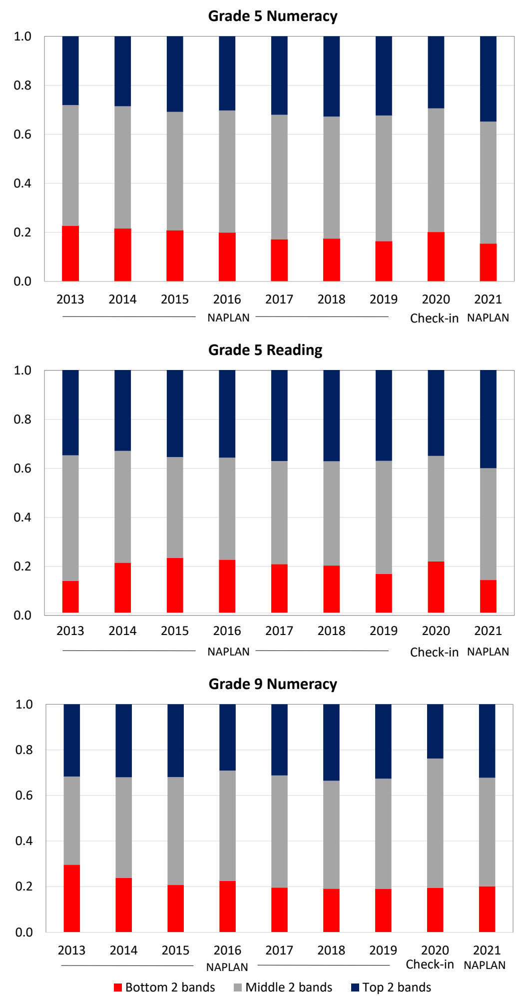
Figure S8: Check-in and NAPLAN Standardized Test Scores



Notes: The figure shows mean scores from the 2020 Check-in assessment conducted in NSW compared with the mean scores from prior and subsequent cohorts on the same type of NAPLAN test.

Summary: NSW students' performance on 2020 Check-in assessments was similar to that of prior cohorts taking the NAPLAN but the 2020 cohort took the test 60 days later than when students would typically sit a NAPLAN exam.

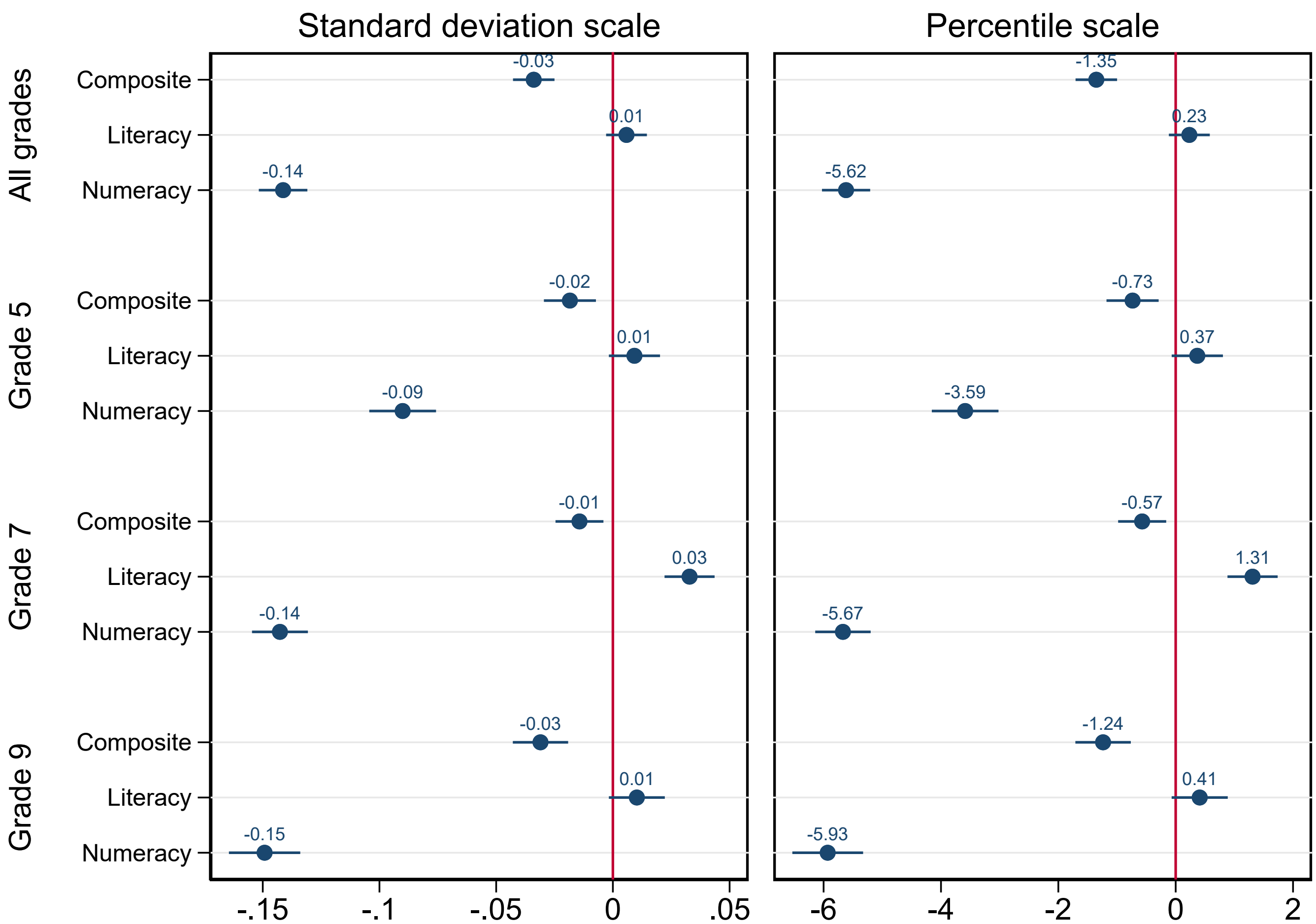
Figure S9: Check-in and NAPLAN Standardized Test Scores
Distribution of Scores



Notes: The figure shows distribution of scores on the 2020 Check-in assessment conducted in NSW compared with the distribution of scores from prior and subsequent cohorts on the same type of NAPLAN test.

Summary: The distribution of test scores were similar between the 2020 Check-in assessments and prior cohorts taking the NAPLAN exam, except for Grade 9 numeracy.

Figure S10: Baseline results: Before-After Learning Trajectories Methodology



Notes: The figure reports results using the difference-in-difference model specification from Equation (1) in Engzell, Frey and Verhagen (2021). Standard errors are clustered at the school level following Engzell, Frey and Verhagen (2021). Percentile estimates in the right panel are computed by transforming the estimates using Equation (3) in Engzell, Frey and Verhagen (2021).

Summary: Not being able to exploit contemporaneous variation in the duration of school closures leads to larger estimated learning losses in numeracy and gains in literacy.