

School Resources, Autonomy and Student Achievement: Evidence from a Regression Discontinuity Design

Mike Helal*

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne

October 17, 2012

Abstract

Contrasting results characterize research into the effectiveness of resources in improving student achievement. Recent findings of substantial principal and teacher effects increasingly suggest that what matters is how school resources are managed and spent. This paper presents quasi-experimental evidence on the impact of resourcing and autonomy using the largest government-funded direct intervention in Australian education. The Smarter Schools National Partnership provided \$2.5 billion in funding to disadvantaged schools with the spending decisions left largely up to schools within broadly defined goals. The program was found to have had the greatest effect on growth in cognitive skills for secondary school students compared to primary school students. The analysis reveals substantial heterogeneity in treatment effects and identifies mechanisms by which this may have occurred.¹

*Email: m.helal@unimelb.edu.au

¹This paper has benefitted from discussions with John Micklewright, Victor Lavy, Chris Ryan, Deborah Cobb-Clark and Michael Coelli as well as seminar participants at the University of Melbourne and Department of Education and Early Child Development.

Acknowledgments

This research uses data from the Victorian Department of Education and Early Childhood Development (DEECD) under the Research and Evaluation Partnerships Agreement with the Melbourne Institute of Applied Economic and Social Research. The views expressed in this report are those of the authors alone and do not represent those of DEECD.

I Introduction

The ability of schools to affect academic achievement after controlling for students' background has been questioned since the Coleman et al. (1966) report finding that school resources had little impact on educational outcomes once family background was accounted for. This finding had a major policy impact as it proposed shifting governments' focus on student outcomes from schools to the home. Over four decades later, the debate around whether schools matter and if so, how and to what extent, continues to be an issue of fundamental concern to policy-makers and researchers. That enhanced educational outcomes should be targeted is rarely challenged yet no consensus can be found on how this is achieved.

Contrasting views of the effectiveness of resources in improving student achievement are best summarised in Hanushek (1996; 2007) and Hedges and Greenwald(1996). Examples of resource allocations that have been studied include class-sizes (Angrist and Lavy (1999), Krueger (1999), and Hoxby (2000)), teacher qualifications and remuneration. While results are mixed across resource categories, more recent research has argued that the aforementioned studies have confused difficult measurability with a lack of true effects. The combination of substantial teacher effects and newly emerging results around principal effects increasingly suggest that how school resources are managed and spent is what matters².

This paper contributes to the debate by presenting quasi-experimental evidence on the impact of school resourcing and management using a large-scale national intervention in Australian schools. The Smarter Schools National Partnership (SSNP) represented one of the largest Commonwealth-funded school initiatives in Australian education (SSNP 2009). Aimed at addressing disadvantage, supporting teachers and school leaders and improving literacy and numeracy, the SSNP include approximately \$2.5 billion in funding provided to schools through state and territory education departments. Crucially, the program is grounded in flexibility, with states deciding how to implement specific reforms.

We analyse the impact of these reforms on Victorian government-schools using a

²see for example Aaronson et al. (2007), Clark et al. (2004), Koedel (2008) and Konstantopoulou (2007) for teacher effects as well as Coelli and Green (2012), Leithwood et al. (2004) and Clark et al. (2009) for principal effects

Fuzzy Regression Discontinuity design (FRD). Discontinuities in program assignment by a forcing variable are exploited to evaluate the causal impact of the SSNP on growth in students' numeracy and literacy outcomes. Unlike other Australian states, Victorian schools operate in a largely autonomous environment with little direction from central authorities. This provides an ideal setting to analyse the impact of the SSNP which can be considered as additional funding received by schools with nearly unrestricted sovereignty in allocating these resources within broadly specified goals.

Previewing the results, I find the program had the greatest effect on growth in cognitive skills for secondary school students compared to primary school students. Differential effects were found by cognitive domain with significant positive impact on students' numeracy skills one year after the program while effects on reading skills were less pronounced. Heterogeneous program effects were additionally found suggesting particular mechanisms by which the SSNP affected outcomes. The efficiency of the program is additionally analysed through detailed financial data on SSNP schools' revenue and expenditure. Finally, we explore the program's impact on factors relating to the education environment and find it has significantly enhanced growth in both students' and teachers' reports of the quality of the teaching and learning in SSNP schools compared to otherwise identical schools.

This paper is organised as follows: The following section summarises the relevant literature on the effect of school governance and resources on student achievement. The SSNP is detailed in section three while the fourth section describes key features of the data. Section five outlines the FRD empirical strategy used in the evaluation and the results are presented in the sixth section. Evidence of heterogeneity in program impact and robustness checks are discussed in the seventh and final section.

II Relevant Literature

2.1 School quality

Policy and research interest in education derive from established findings on the link between human capital and individual market and non-market outcomes. Using individ-

ual quantity of schooling as a proxy for stock of human capital an extensive literature demonstrates its theoretical and empirical relationship with income, productivity, and economic growth. As rates of school completion converged in developed countries concern has shifted to the quality of schooling received. To motivate the background behind the SSNP as a program targeting student achievement and to consider anticipated pathways for its effects, the relevant literature on school quality and school resourcing is highlighted here.

Why should we be concerned with quality and believe it to contribute to some of the differences in rates of return observed for individuals with similar qualifications? The concept of quality is difficult to measure and is thus often represented by achievement in test scores or against explicit developmental criteria. Blackburn and Neumark (1993, 1995) Murnane et al. (1995), Currie and Thomas (2000) all find strong evidence of earnings advantages to cognitive skills measured by higher achievement on standardised tests. Hanushek and Kimko (2000) along with Woessman et al (2008) show productivity and national growth rates increasing in similar measures of school quality. Despite acknowledgement of the imperfect measurement provided by any standardised test Murnane et al. (2002) demonstrate increased returns to measured skills across simple analysis and various error-corrected models.

One potential pathway by which achievement causes improved labour market outcomes may be in its positive reinforcement of the value of education to individuals. Manski and Wise (1983) first highlighted that students who do better in earlier school years tend to study further. Rivkin (1995) reasons that variation in test scores explains most of the black-white differences in attainment. While these effects may be capturing dynamic features of the rate of skill accumulation which may vary at different points of the achievement distribution, they nonetheless highlight the potential for subsequent positive spillover when raising achievement.

Given these effects, enhancing school quality forms a desirable yet elusive objective of most governments. This has led to significant growth in resources allocated to education which now represents 25% of social spending on average across OECD budgets (OECD 2010).

2.2 Resources and Quality of schooling - Peers, Teachers, Management and Resourcing

The disparate findings in research on quality returns to resource investment in education are well-documented elsewhere. Hanushek (1986; 1997; 2003) declares the failure of inputs commonly advocated in the education sphere to achieve meaningful impact on quality of education as measured by student test scores. The author's meta-analysis of existing studies on classroom resources, financial aggregates and facilities investment shows mixed evidence of statistical significance and the direction of effects. This contrasts with Greenwald et al. (1996) Angrist and Lavy (1999) for instance who find elements of the above inputs positively related to achievement. Experimental evidence from Project STAR, a randomised control trial placing students in different class sizes, also produced mixed evidence. Findings range from the highly positive relationship found in Word et al. (1990) and Krueger (1999; 2003) to half the effects which decline to insignificance in later grade levels in analysis by Rivkin et al. (2005). Other seemingly important factors in schooling such as teacher qualifications (Kane, Rockoff and Staiger 2006), teacher training (Boyd et al. 2005) and teacher salaries fail to provide consistent evidence of positively impacting student achievement.

2.3 Peers

How can these counter-intuitive findings be explained? The Coleman report (1966) and other similar studies would justify this insignificance as true ineffectiveness of the above inputs in the education production function. Instead, they would argue that the family background effect dominates all others in education, rendering school interventions second-best if not completely ineffective compared to redistributive assistance for disadvantaged households for instance. Another strand of the literature that could explain the apparent orthogonality of student outcomes and school inputs would be peer effects literature. If what matters most for a student's own achievement is the socio-economic background, the ability or the motivation of their peers - in class, at school or in the neighbourhood - additional resources for schools may not be as effective as particular streaming or sorting and integration policies within and across schools. More recent re-

search that exploits advances in data collection and linkage has yielded evidence which refutes the claim that school inputs such as staff provide no additional benefit to students after controlling for background. In earlier work using value-added modelling of school effects, Helal et al. (2012) find that while peer effects may matter, there is a small yet significant effect of schools as a whole on growth in student achievement. Without decomposing the school effect into its component productive parts, measurement-error adjusted estimates of school effectiveness suggest a difference in student growth equivalent to one quarter of a year between effective and ineffective schools on average. These effects were additionally found to be robust to alternative, less restrictive modelling (Helal 2012).

2.4 Teachers

Teacher effects studies exploit variation of student outcomes by different cohorts for the same teacher or the same cohort with different teachers. Sizeable effects are found most notably by Balou et al. (2004), Hanushek et al. (2005), Kane et al. (2006) and Lavy (2011). Estimates suggest moving one standard deviation up the distribution of teacher fixed effects raised students' reading and mathematics test scores between one-tenth and one-fifth of a standard deviation on the national scale. Leigh (2011) uses biennial test scores to estimate teacher effects that reveal a difference of approximately half a year of progress within the same academic year between an effective and ineffective teacher. The convincing empirical evidence on teacher effects, though it may not reveal the characteristics of effective teachers, implies that resources which develop effective teachers can have an impact on student achievement.

2.5 Management and Resourcing

School principals are another input component of schools that has more recently been found to affect student achievement. Here again, observable productive traits are not found to cause improvement in achievement yet principals around found to have significant effects in Branch et al. (2009), Coelli and Green (2009) and Dhuey and Smith (2011). Methods and effect sizes are similar among the studies and are as high as 0.3 standard deviations in student test scores for a a one standard deviation improvement

in principal quality. Leithwood et al. (2004) note that successful principals are those who lead improvement, innovation and change while efficiently managing the school. Yet principals' impact on student achievement is dependent upon the legislative environment that schools operate in. The importance of school leaders is further heightened where a greater range of responsibilities are exercised at the school level. In cross-country comparisons, Hanushek, Link and Woessmann (2011) and OECD (2010) find a wide range of decision-making powers vested in principals across and within countries. The conclusion from these studies is that the effect of principal autonomy is largely determined by a combination of factors including the types of devolved responsibilities, the school's capacity to assume these responsibilities and the extent of accountability in the system.

The nontrivial effect of teachers and principals on student achievement in the above studies show that analyses of aggregate school resourcing effects may conceal substantial variation. Reconciling the apparently contradictory findings, one can see that what matters for student achievement may not be aggregated expenditure or average allocation per pupil. Indeed with most school resourcing comprised of funds that are *ex ante* earmarked for essential operational costs it may not be surprising that overall funding analysis fails to pick up effects.

Instead, there are likely to be some inputs that matter more than others for student achievement. The decision-making context governing resource allocation is another confounding factor that needs to be separated in any such analysis. When the inputs are principals for example, it is not just directed resourcing which matters but the extent to which school leaders are in control over how to expend those funds. It has been historically difficult to find and measure resources at detailed inputs level in any systematic way.

As the next section demonstrates, the Smarter Schools National Partnership in Victoria thus provide a rare opportunity in this research field to evaluate the impact of additional resources, granted under very broad conditions within a devolved policy environment.

III Background on the Smarter Schools National Partnerships

The Smarter Schools National Partnerships (SSNP) represent the largest joint Commonwealth-State intervention of its kind. This section briefly describes the institutional background to its onset and details the program's components.

3.1 NAPLAN

In 2008, the first ever standardised assessment was implemented across Australian states and territories. It accompanied other national and state-level education initiatives aimed at transparency and accountability including an intended national curriculum and mandated public reporting of school achievement results. Since then, every student in Australia has been required to participate in the National Assessment Program - Literacy and Numeracy (NAPLAN) in Years 3, 5, 7 and 9. The assessment is intended to measure essential skills under five domains: numeracy, reading, grammar, spelling and writing.

The commitment to displaying schools' NAPLAN results on MySchool since 2010, a publicly accessible website, demonstrated state and federal education authorities' intent to treat test scores as a major proxy of the quality of education offered by schools. The test has since dominated education discourse in Australia and while the impact of testing and accountability policies are beyond the scope of this paper, NAPLAN has become the first common measure of student achievement across Australia.

3.2 The SSNP

Shortly following the publication of results from the first round of NAPLAN test, The Smarter Schools National Partnerships (SSNP) were announced by the Council of Australian Governments in 2009. The Partnerships represented a new approach to policy-making in Australia, driven by Commonwealth funding for an area traditionally within state control, coupled with state input into the use of funds. Under the SSNP the Australian government pledged to provide significant additional funding to schools from all sectors and states to achieve two major objectives: the enhancement of educational outcomes and the development of an evidence base on effective practices in education. The latter objective is what makes the program data so important to analyse today - states

argued that best practice is best understood at the point of delivery leading to a departure in implementation from the founding documents of the SSNP COAG 2009. States were largely free to determine how the additional funding would be spent with the only requirement that expenditure would aim to "improve students' literacy and numeracy outcomes, strengthen the capacity of disadvantaged schools or drive continuous improvement in teaching." With some variation between states, this relatively unrestricted implementation turns the SSNP into a program that facilitated autonomous spending by schools using an additional set of funds. Another first for Australia was the introduction of reward funding for states that demonstrate progress against predetermined targets. As demonstrated below, reward funding was only a minor component of the program with most resources granted unconditionally to qualifying schools as facilitation funding to achieve the above goals³.

These Partnerships have been labelled Literacy and Numeracy NP, Low SES NP and Improving Teacher Quality NP. This paper evaluates the first two of these interventions due to insufficient data on the third as it was only implemented on an opt-in basis with minor takeup by schools⁴. The Commonwealth government distributed funds under the SSNP to states as a function of the ABS index of disadvantage while states were then asked to allocate funds for schools according to a state-determined threshold. The intervention can thus best be thought of as an external boost to school budgets, targeted at disadvantaged schools with the aim of lifting literacy and numeracy skills as measured by NAPLAN.

The level of SSNP funding by category is shown in Figure 1⁵. Of the \$2.5 billion in SSNP funding, 60% were assigned to the Low SES NP compared to 22% to the Literacy

³The reward component of the SSNP is not believed to have influenced school behaviour in any systematic way. The rewards were to be granted to state education authorities conditional on state-wide improvements. As the gain from meeting targets would not be directly felt by schools, we are not overly concerned with Hawthorne effects: unobserved changes in schools' behaviour driven by knowledge of their participation in the intervention.

⁴No clear distinction between Low SES NP and Literacy and Numeracy NP is supported by the data. Indeed the SES selection criteria was found to be binding for most schools. More importantly, selection under either category had no treatment variation apart from different funding levels. Our analysis using the level of SSNP funding as a treatment thus overcomes this labelling distinction

⁵Implementation costs comprise the remainder of the total NP funds.

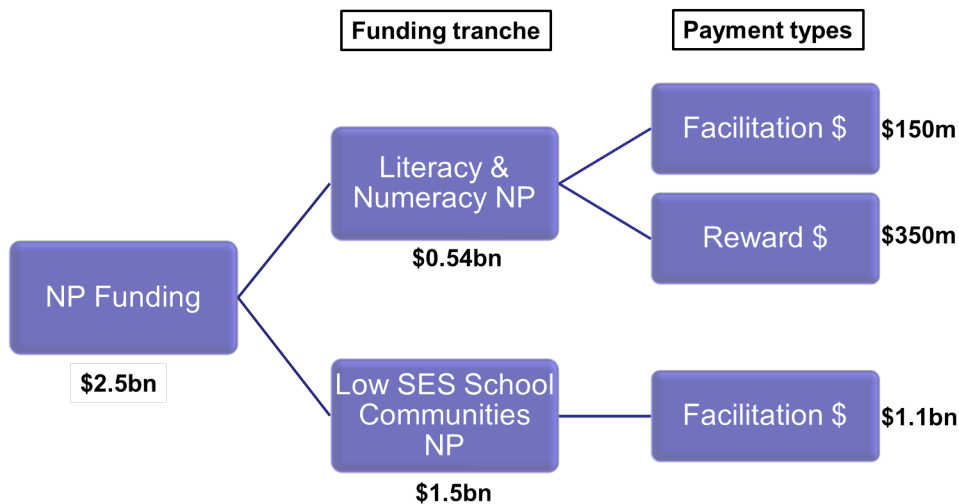


Figure 1: Funding under the Smarter Schools National Partnerships

and Numeracy NP with the remainder being spent on program administration. The funds received by schools directly are those categorised as facilitation funds, totalling \$1.25 billion throughout Australia with approximately 90% dedicated to the Low SES NP. Throughout all states and territories in Australia, 2656 schools in total, of which 75% were government schools, were selected into the SSNP.

Table 1 shows the national breakdown of SSNP as agreed by COAG (2009). Victoria is set to receive \$203 million by the conclusion of the Partnerships, representing 18% of all SSNP funding.

The assignment of schools into the SSNP in Victoria followed a state-wide criterion defined by the school's Student Family Occupation (SFO) index. Lower-skilled occupation groups receive higher weight thus a higher student family occupation index translates to a lower socioeconomic level.

There are two remarks to note in relation to the timing of resource distribution. The first is that facilitation funding was unconditionally agreed at the start of the Partnerships. States hence knew the level of funding they would receive from 2009 with certainty. Schools chosen for the program were additionally informed at the time of selection that funds would be distributed in 2010. Nevertheless, it is unlikely that schools chosen into the program manipulated their 2009 NAPLAN score as the funding decisions were reported in the second half of that year, after NAPLAN had been taken. For the purposes of evaluating the program impact as explained below, all schools in the SSNP had re-

ceived their allocation by 2010. The second point to note is that our analysis of the data in 2011 represents just over half the lifetime of the program as additional funds continue to be distributed in 2012 and 2013.

The following section details the data used in this paper.

IV Data

This paper analyses the effects of the SSNP on changes in students' skills as measured by NAPLAN. Since students undertake the tests biennially, the analysis examines changes between 2009, prior to the SSNP, and 2011, following a complete year of full implementation of SSNP. Treatment by the SSNP is defined as selection into the program and receipt of funding by the school⁶. Our data links administrative records from various divisions at the Department of Education and Early Childhood Development combining student level information on NAPLAN performance and family background as well as school level data on program participation, implemented initiatives, school finances and historical achievement. Table 2 demonstrates the implementation of the SSNP in Victoria.

334 schools attended by over 137,000 students in total were found to have received SSNP funding in Victoria. This represented one-quarter of all students in the state and more than one-third of Victoria's indigenous students. Indigenous students are overrepresented in the sample of treated students compared to all students in Victoria since the program targeted schools with high proportions of disadvantaged students.

Two samples were constructed from the population of government-schooled students in Victoria. The first, referred to as the primary school sample, includes students who move from Year 3 in 2009 to Year 5 in 2011. The second sample comprises secondary school students who move from Year 7 prior to the SSNP in 2009 to Year 9 in 2011.

⁶Though funding and selection into the program should be identical, we find two schools initially chosen have not received any SSNP funds by December 2011.

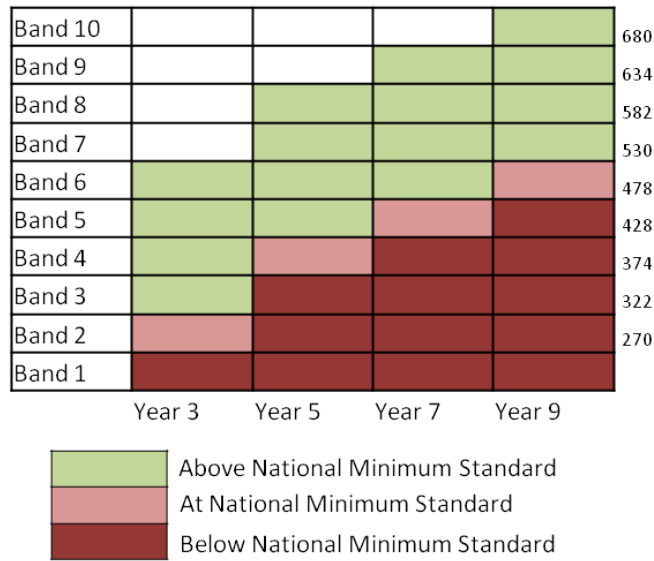


Figure 2: Vertical Scaling in the Measurement Instrument - NAPLAN

No.	Sample	2009	2011
1	Primary School Matched Sample	Year 3	Year 5
2	Secondary School Matched Sample	Year 7	Year 9

Keeping students with valid scores in NAPLAN Numeracy and Reading, estimation is conducted over approximately 35,000 and 30,000 students in the primary and secondary samples respectively. I use NAPLAN scores provided by the testing authority according to a vertical Rasch scale which covers all years of testing. NAPLAN questions are linked across grades in successive testing rounds. The unified scale then allows for the measurement of progress over grade levels as demonstrated by Figure 2 below.

Summary statistics comparing treated students to the general student population are presented in Table 3. SSNP students are more likely to come from a language background other than English (LBOTE) and are more than twice as likely to be indigenous (ATSI). As a result of the program's targeting of disadvantaged schools, 12% of treated students live in unemployed households compared to only 3% of the general student population. Approximately 45% of parents of students not in the SSNP are in a managerial or professional occupation while this is less than 19% of treated students. The differences in parental education between the two groups are further highlighted in this table. Com-

pared to the 29% of students with tertiary educated parents in non-SSNP schools, only 11% of SSNP students have a parent with a tertiary qualification.

V Empirical Strategy

The typical empirical challenge in causal evaluation is the inability to observe individuals in counterfactual states. In any intervention, subjects are either exposed or not exposed to a treatment T . Therefore, the econometrician can only observe an outcome Y under two states as presented below: The treated outcome for those who have been treated by T and the untreated outcome for those not chosen for the intervention.

$$Y_i = (1 - T_i) \cdot Y_i(0) + T_i \cdot Y_i(1) = \begin{cases} Y_i(0) & \text{if } T_i = 0 \\ Y_i(1) & \text{if } T_i = 1 \end{cases} \quad (1)$$

Under truly randomised allocation into treatment and control groups, the causal effect of the treatment is then simply given by

$$Y_i(1) - Y_i(0) \quad (2)$$

However, many interventions in education deliberately target groups of students in a non-random pattern. In particular, selection into treatment is often on the basis of low achievement or disadvantage. This may violate the assumption of orthogonality with the unobservables and leads to biased coefficient estimates. By tying program allocation to the 2008 school SFO, the SSNP policy allows for causal identification despite non-random assignment. In the vicinity of the specified SFO cutoff, unobservables are not assumed to vary discontinuously. This provides ‘local randomisation’ around the cutoff where program assignment can be considered exogenous for schools just above or below the threshold. The estimation strategy to be used in this instance is referred to as Regression Discontinuity (RD) design.

5.1 Identification

The RD design - due to Thistlewaite and Campbell (1960)⁷ - represents a way of estimating treatment effects in non-experimental settings by exploiting whether an observed allocation variable exceeds a designated cutoff. The idea of RD is to use a discontinuity in the level of treatment related to some observable predictor X_i to get a consistent estimate of average treatment effect at the cutoff. The canonical RD example presented in Figure 3 below estimates the treatment effect τ as the effects of the intervention on subjects' outcome variable at a value of X_i equal to the cutoff.

This predictor X_i may itself be associated with the potential outcomes, but association is assumed to be smooth. Under the assumption of continuity in all covariates at the threshold, any discontinuity in conditional distribution of the outcome at the cutoff value is interpreted as evidence of a causal effect of the treatment.

The identifying assumption might be violated if subjects can manipulate the forcing variable. This is empirically testable and as the next section shows, we do not see any evidence of discontinuities in school SFO around the cutoff point. Indeed, given that the assignment rule was a function of 2008 SFO, schools could not influence this index at the time SSNP allocation decisions were made in 2009.

Note that identification in RD designs does not necessitate assignment to the treatment be perfectly determined by the value of the predictor, X_i , commonly referred to as the forcing variable. The literature distinguishes between a sharp RD where treatment is a deterministic function of the forcing variable and fuzzy RD where a jump in the probability of assignment occurs at the cutoff value (Imbens and Wooldridge 2007). An example of a sharp RD would be an advancement rule in state high schools might require students to achieve a minimum score of 60% to move to a higher grade level. Students who score 60% or more are eligible to advance while those just below the cutoff are retained. A fuzzy RD on the other hand would be where students below 60% are referred to a panel which consequently makes the advancement decision in light of a number of other factors relating to the student.

⁷Since then this method has been increasingly used in economics, most notably by Hahn, Todd and van der Klaauw (2001) and Angrist and Lavy (1999). See Lee and Lemieux(2010) or Imbens and Lemieux (2007) for a detailed review.

To see how identification is obtained, consider treatment T_i such that:

$$T_i = 1\{X_i \geq c\} \quad (3)$$

where T_i is a deterministic function of X_i . Assuming continuity at x equal to a cutoff c , the conditional expectation of the observed outcome is given by:

$$\begin{aligned} E[Y_i(0)|X_i = c] &= \lim_{x \rightarrow c^-} E[Y_i(0)|X_i = x] \\ &= \lim_{x \rightarrow c^-} E[Y_i(0)|T_i = 0, X_i = x] \\ &= \lim_{x \rightarrow c^-} E[Y_i|X_i = x] \end{aligned} \quad (4)$$

Similarly,

$$E[Y_i(1)|X_i = c] = \lim_{x \rightarrow c^+} E[Y_i|X_i = x]$$

Thus the average treatment effect at the cutoff for a sharp RD can be estimated as

$$\tau_{SRD} = \lim_{x \rightarrow c^+} E[Y_i|X_i = x] - \lim_{x \rightarrow c^-} E[Y_i|X_i = x] \quad (5)$$

which is just the difference of two regression functions at the cutoff point (Lee and Lemieux 2010). In some situations, such as when there is imperfect take-up by program participants or imperfect adherence to the threshold in allocation rules, treatment is only partly determined by crossing the threshold. Using the same notation, fuzzy RD in this instance only requires a jump in the probability of treatment at the cutoff such that:

$$\lim_{x \rightarrow c^+} Pr[Y_i = 1|X_i = x] \neq \lim_{x \rightarrow c^-} Pr[Y_i = 1|X_i = x]$$

Here, the difference in outcomes must be scaled by the difference in the probability of treatment. As first noted by Hahn et al. (2001), the fuzzy RD estimator is analogous to the IV Wald estimator and under the assumption of monotonicity the treatment effect can be recovered by:

$$\tau_F = \frac{\lim_{x \rightarrow c^+} E[Y_i|X_i = x] - \lim_{x \rightarrow c^-} E[Y_i|X_i = x]}{\lim_{x \rightarrow c^+} E[T_i|X_i = x] - \lim_{x \rightarrow c^-} E[T_i|X_i = x]} \quad (6)$$

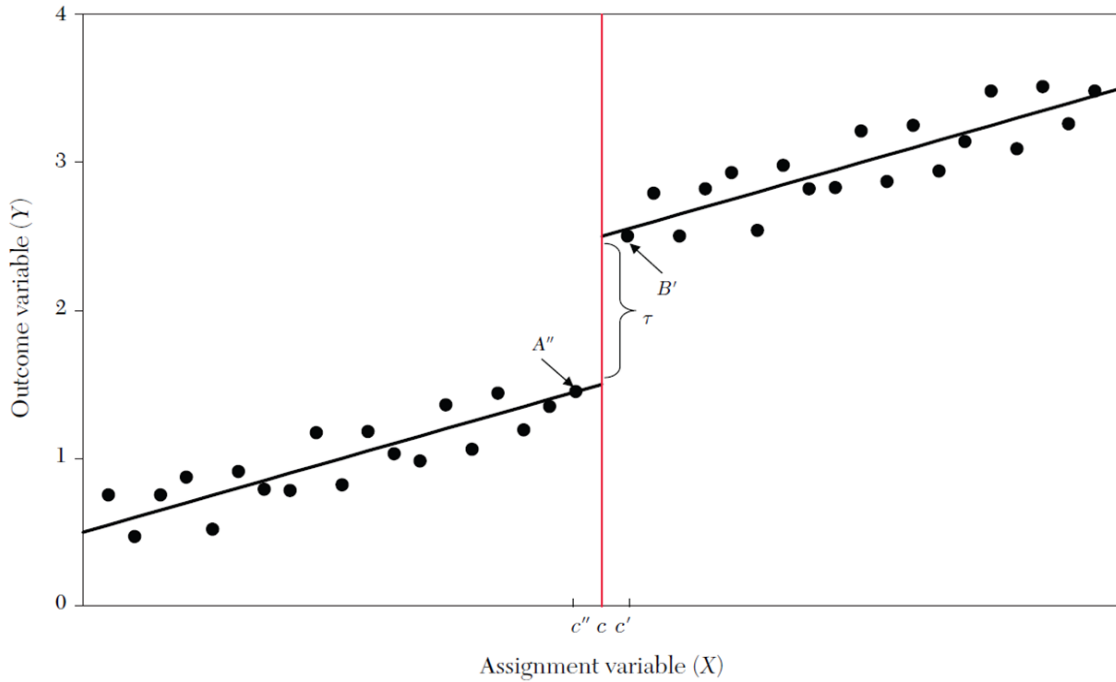


Figure 3

Where the numerator is the difference in conditional expectation of the outcome variable and the denominator is the jump in probability of treatment at the cutoff. It is important to recognise that the fuzzy RD estimator thus measures the average effect of the treatment on subjects at the threshold who have been allocated into the program as a function of their forcing variable.

Identification of program effects by RD designs are essentially dependent upon a discontinuity in the assignment variable at the threshold and a corresponding discontinuity in the outcome variable. Graphical analysis of schools' SFO by SSNP treatment status reveals imperfect adherence to the designated criteria. The probability of SSNP selection is plotted against rounded school SFO indexes in Figure 3. At the state set threshold of 0.7, the likelihood of treatment jumps from approximately 25% to over 77%. While crossing the threshold does not deterministically entitle schools to SSNP funding, the increased allocation likelihood along with the aforementioned continuity and monotonicity assumptions, allow for a fuzzy RD analysis of the program's effects on student achievement.

5.2 Estimation

In a fuzzy RD design, the probability of treatment discontinuously rises at the cutoff. The discontinuity can be used as an instrumental variable (IV) for treatment status. Indeed the Wald estimator presented above is numerically equivalent to the treatment coefficient in an IV setup (Jacob and Lefgren 2004).

Functional form is a concern in fuzzy RD as nonlinearities in the relationship between outcome variables and forcing variables may be misinterpreted as discontinuities due to the intervention. As a result flexible modelling should be employed to ensure results are not driven by functional form (Lavy 2011). This section demonstrates equivalent parametric strategies as well as a non-parametric technique to estimate treatment effects.

In a 2SLS framework let D_i be a dummy equal to one if the school crosses the assignment threshold. In a fuzzy RD design, this does not perfectly determine treatment T_i . We can set up a first-stage equation of treatment as a polynomial function of a vector of covariates \mathbf{X}_i and the indicator D_i for whether the subject has exceeded the threshold.

First stage:

$$T_i = \gamma_0 + \gamma_1 \mathbf{X}_i + \gamma_2 \mathbf{X}_i^2 + \dots + \gamma_n \mathbf{X}_i^n + \pi D_i + \zeta_i \quad (7)$$

Reduced form:

$$Y_i = \alpha_0 + \alpha_1 \mathbf{X}_i + \alpha_2 \mathbf{X}_i^2 + \dots + \alpha_n \mathbf{X}_i^n + \tau T_i + \eta_i \quad (8)$$

With one instrument this just identified model is characterised by robust finite sample properties for estimators of the treatment effect τ . However, some extensions can be applied to generalise the model. Varying treatment effects can be modelled for instance, by centring the forcing variable around the threshold and interacting treatment with this forcing variable. The local average treatment effect at the cutoff is then the coefficient on the treatment dummy.

A less restrictive non-parametric estimation technique has been additionally proposed by Lee and Lemieux (2010). The single equation for this method is presented in (??). Using one equation, local linear regression can be run on the forcing variable centred at the cutoff along with interaction terms with the treatment indicator within a specified

bandwidth around the threshold.

$$Y_i = \beta_0 + \rho T_i + \beta_1 \tilde{x}_i + \beta_2 \tilde{x}_i T_i + u_i \quad (9)$$

Interactions between the treatment and the centred running variable allow the smoothing function to differ on either side of the cutoff. The bandwidth h is chosen as the solution to an objective function minimising mean-squared error over local linear regressions⁸.

The results presented in the following sections employ both the IV technique and the nonparametric local linear regression with similar results. For the latter, I report results using a bandwidth which optimises the Lee and Lemieux (2010) cross-validation criteria although findings are robust to multiple binwidth choices. To evaluate the effects of SSNP on growth in student achievement, I set up an unrestricted value-added model where the outcome variable, represented by NAPLAN score in a particular domain, is regressed on prior NAPLAN score, student demographics and an interacted treatment dummy.

$$A_{idst} = \lambda A_{ids,t-2} + \alpha_1 \mathbf{X}_{ist} + \alpha_2 x_{ist} + \tau NP_{st} + \phi NP_{st} x_{ist} + \epsilon_i \quad (10)$$

The notation used is as before with achievement level A for student i in domain d at school s is the score she achieves in the particular test. This is determined by vectors of individual inputs \mathbf{X} , a school level forcing variable which influences program assignment \mathbf{x} and a dummy NP_{st} for treatment status. Finally any test is likely to contain a noisy measure of true student achievement, accounted for by ϵ . Various definitions of treatment are employed including binary participation/non-participation, length of participation and the level of funding received as a continuous treatment. The coefficient of interest is that on NP_{st} which shows the local average treatment effect i.e. the effect of participating in the SSNP on growth in student achievement for a student whose school was treated due to the cutoff rule.

⁸For more, see Lee and Lemieux 2010 for an extensive discussion of cross-validation procedures to choose optimal bandwidth.

VI Results

6.1 Non-parametric analysis

To facilitate graphical analysis a single variable is constructed to depict growth in achievement. Note that taking the difference between current and past achievement as a measure of growth imposes unrealistic assumptions on the rate of learning, namely that no knowledge attrition occurs. Instead, I construct conditional growth percentiles as in Helal, Justman and Houng (2012). The conditional growth percentile represents a student's percentile rank in a particular NAPLAN domain in 2011 compared to all those who had an identical score in 2009.

Modelling growth using this single variable, simple non-parametric results are first presented in Figure 4 below. The charts show local linear regressions of the conditional growth percentiles on the forcing variable fitted around the threshold SFO of 0.7. Two features are immediately noticeable in the graphs. The first is a significantly greater jump in Numeracy conditional percentiles at the cutoff compared to Reading. The second apparent result is a larger jump in the secondary school sample compared to students who went from Year 3 to Year 5. A negligible jump in Reading is found for the primary school sample with only a slightly larger jump for students in secondary school.

Conditional growth percentiles: 2011 rank relative to all students w/ same 2009 score

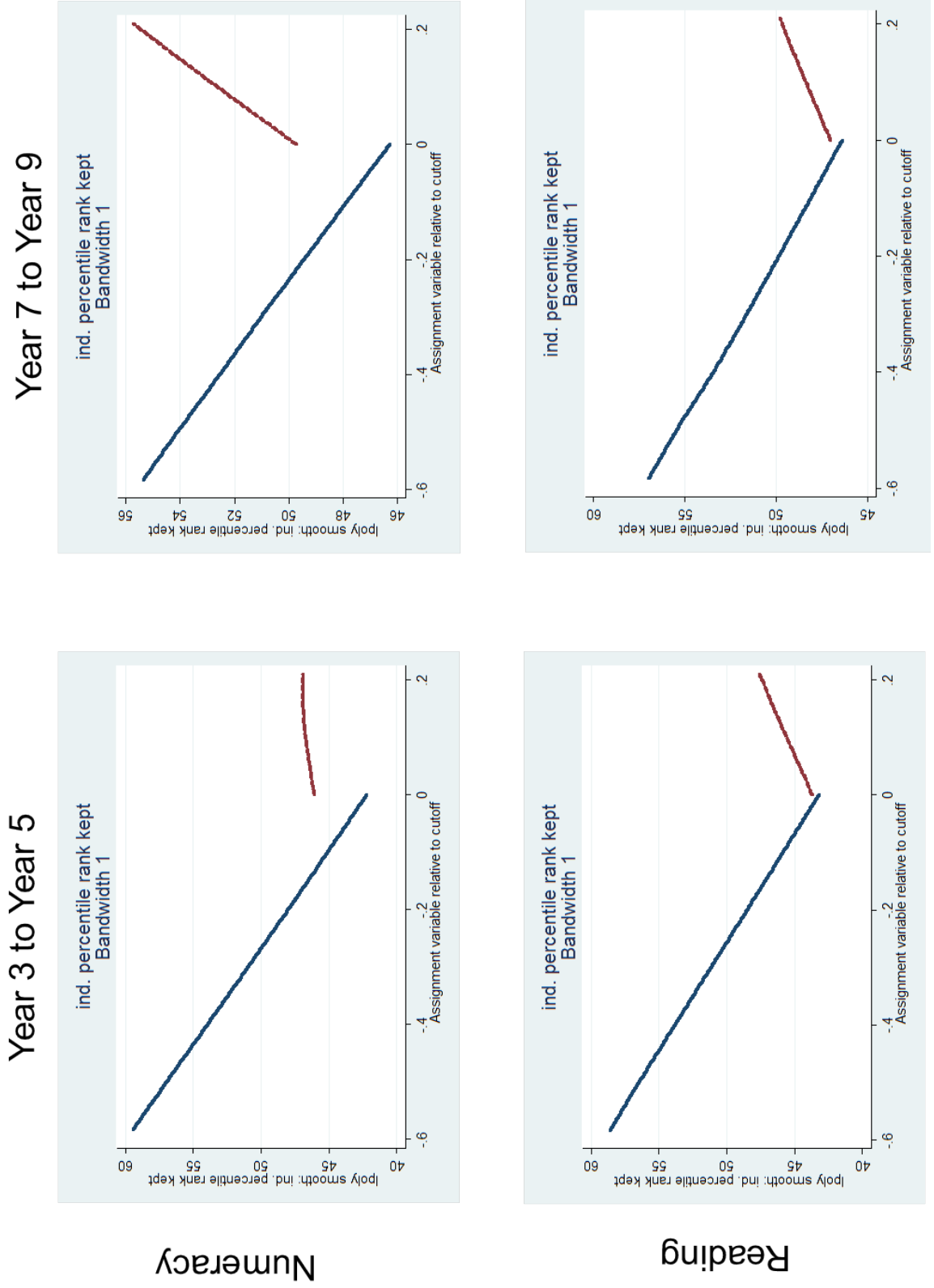


Figure 4

The graphs show Numeracy rankings rise by approximately 3 percentiles at the cutoff in primary school and 4 percentiles in secondary school. As demonstrated in equation (??), to estimate treatment effects these rises then need to be weighted by the increased likelihood of program assignment to estimate the previously outlined Wald estimator.

The ratio of the jump in the outcome variable to the jump in likelihood of treatment represents the nonparametric estimate of the SSNP local average treatment effect. In Reading, this was only 1 percentile point in the primary school sample and 6 points for secondary school students. This result implies no statistically significant improvement in Reading for primary school SSNP students compared to others at the margin who were not treated while 6 percentile points translate to approximately 16 points on the NAPLAN scale in secondary school. The SSNP effects on Numeracy appear to be substantially higher with a Wald estimate of 8 percentile points from Year 3 to Year 5 and 31 points from Year 7 to Year 9.

Compared to students who achieved the same score in 2009, students in SSNP schools thus appear to rank more than 1 standard deviation above their counterparts right below the program threshold. Given that the latter group is presumed to only differ in their treatment status, this suggests substantially positive effects on growth in student achievement due to the SSNP. The non-parametric estimates additionally suggest the SSNP may have had differential effects on students of different year levels as well as varying effects by domain.

OLS

Two parametric techniques are additionally employed to test the robustness of the above non-parametric estimates. The first is the value-added model presented in equation (??). In addition to prior achievement as a regressor, this model has the attractive feature of controlling for student background and family characteristics.

The model is presented along with a comparison specification in Table 4. Even-numbered columns directly reflect equation (??). They represent results of a value-added regression on observations within an optimised bandwidth around the cutoff of 0.7, with the running variable centred at the cutoff and interacted with treatment status. The local

average treatment effect, i.e. the effect of the treatment at the cutoff, is then given by the coefficient on the treatment status variable SSNP. A comparison model with similar controls is presented in odd-numbered columns. It estimates over the entire range of the selection criteria with the running variable and a dummy for SSNP participation as regressors.

Males are found to achieve higher growth than females in Numeracy while the opposite is true in Reading. The socioeconomic gradient is found to be larger in primary school than secondary school. Controlling for rurality, school size appears to be positively associated with Numeracy growth in primary school and both domains in secondary school. While larger schools may benefit from economies of scale or departmental teams, the relationship between size and growth need not be causal as successful schools may be attracting greater enrolments.

The results of Table 4 show that failing to account for the endogeneity of treatment by design underestimates the effects of the SSNP. The coefficient on treatment status in odd-numbered columns can be interpreted as an average treatment effect which is undoubtedly lower than the effect of the SSNP on its target population of disadvantaged students. From Year 3 to Year 5, column 1 shows SSNP participation is estimated to have contributed 4 additional NAPLAN points for a common starting level. However, column 2 highlights that the local average treatment effect of the SSNP is an additional 7 NAPLAN points. No statistically significant program effects were found in primary school Reading.

The effects of the SSNP on secondary school students are substantially higher. While average treatment effects appear negligible in both domains, columns (6) and (8) demonstrate the true effect of the SSNP on its targeted population of disadvantaged students at the program threshold. Treated students are shown to achieve an additional 27 point growth in Numeracy over their baseline score compared to similar non-treated students at the cutoff. The gain due to the SSNP is smaller yet still significant in Reading at 8 points. Considering that the National Minimum Standard on the NAPLAN scale rises by approximately 50 points over two years, treatment by the SSNP appears to add the equivalent of more than a full year of schooling in secondary school Numeracy.

6.2 IV

Finally, I present the results of the preferred model which treats the fuzzy RD design as an instrumental variables problem in Table 5. The reduced form of this model uses first stage fitted values from a regression of treatment on student covariates, the running variable and a dummy for crossing the threshold. Expanded student background covariates show students who have a language background other than English, with neither parent having completed high school, perform significantly worse than their peers. Their achievement growth is seven points lower in Numeracy and 6 points lower in Reading than otherwise identical students with the same baseline score. Indigenous students underperform by a similar magnitude at both primary and secondary school levels. Parental education is found to have the greatest impact on student achievement as primary students of tertiary educated parents experience a 12 point gain over students with the same score in 2009 whose parents do not have post-secondary education.

The coefficient on SSNP status shows the local average treatment effect at the threshold. It provides causal evidence of the effect of the program on growth in student achievement. The IV estimates are consistent with the figures found using both the non-parametric and OLS models. Table 5 show the SSNP has caused an additional improvement of 6 NAPLAN points in primary school Numeracy. The program does not appear to have had any effect on primary school Reading skills as measured by NAPLAN. The results at the secondary level show the SSNP has been more effective for students in higher year levels in both domains. An additional 12 points of growth in Reading skills were due to the SSNP. Numeracy scores were raised by 36 points on average as a result of participation in the SSNP. Validating the OLS results above, these figures show substantial program effects equivalent to approximately half a standard deviation in Reading and 1.5 standard deviations in Numeracy. The SSNP thus appears to have been successful in raising student achievement beyond typical growth between grade levels. The next section analyses these findings and shows variation in the program effects on different student subgroups.

6.3 Heterogeneity in Program Effects

In any treatment, it is likely that the intervention has differential effects on students of different types. The estimates reported above are averaged effects at the state-set cutoff. To determine the program's effects on various sub-populations of interest to policymakers, the preferred IV model was estimated separately for each group.

The first disaggregation estimates the program's effects on students of different socioeconomic background. 3 groups were created: the lowest quartile, the middle half and the top quartile of socioeconomic status. Though the program targeted schools with low SFO indexes, government schools in Victoria are mixed meaning students of all socioeconomic backgrounds were in the treated and untreated groups. I find students in the lowest SES quartile appear to have benefited more from the SSNP in primary school than students from the highest quartile though the difference is less than 3 NAPLAN points. The variation is much more pronounced in secondary school where the growth in disadvantaged students' Numeracy scores was 13 points higher than that of students in the highest SES quartile. A similar differential of 14 points was found in Reading. By providing additional funding to schools, the SSNP appear to have addressed resource issues for disadvantaged students who are likely not to have had access to such resources at home. It is likely that the program filled a gap in the limited academic support that could be offered by their comparatively less educated parents. The significance of the results at the secondary school level underscore this point. At this level of schooling, less educated parents' knowledge of the content of the curriculum hinders their ability to support students. Additional funding that schools spent on enhancing teaching strategies is likely to have addressed this shortfall in academic home support.

The IV model was additionally estimated for students of different academic ability. Quartiles of achievement in the base year i.e. Year 3 and Year 7 in 2009 were created for the primary and secondary samples. In Numeracy, the effect of participating in the NP was found to be most equitable in secondary schools with no significant differences in the program's value-added between students in all quartiles of base year score.

At the primary school level however, high achieving students in NP schools were found to have attained greater gains compared to students who were in the lower quartile of

achievement in Year 3 in 2009. No heterogeneity in NP effects on Reading growth was found at the primary or secondary school levels. Further analysis of the exact interventions implemented by schools offer insight into why this may have occurred. It is well-established in the education literature that the absence of meaningful challenge in learning leads to disengagement amongst students. This is particularly important for high achievers who are likely to disassociate from a class where undifferentiated teaching targets the lowest common denominator. An intervention like the SSNP equips teachers with strategies to differentiate their practice, offering various levels of difficulty within the classroom context. Higher treatment effects for high achievers imply that such a program that widens teachers' offerings in class finds most resonance amongst previously disinterested high achievers.

VII Discussion of Findings

Through access to detailed school financial data, I analyse the monetary impact of the SSNP on student achievement. Further investigation of the program implementation revealed some administrative regions in Victoria had a large proportion of eligible schools according to the SFO criteria. Rather than raise the threshold, DEECD regions split the funding over a greater number of schools, resulting in variation in dollars per student received within treated schools across Victoria.

The total funds received by each school as part of the SSNP were divided by the number of full-time enrolled students to determine per-student funding. The median level of funding received by treated schools was found to be \$719 per student, representing approximately 7% of all per-student funding received by schools annually. The above IV analysis was then repeated using the log of per-student dollars received by each school as the treatment variable. A dummy for whether the school exceeded the threshold was again used as an instrument for treatment. I find the program effects to be increasing in the level of funding received, though at a decreasing rate. The highest return on funds was found at approximately \$440 per student, \$280 below the median level of funds received by schools. Regional socioeconomic differences were found to have resulted in rates of SSNP funding that ranged from 2% of total annual funding to more than 23%. The

strong positive relationship between funding levels and growth in student achievement can partly explain the variation in outcomes between regions who participated in the program.

This substantial positive shock to school budgets was largely unconditional with schools free to choose how to spend the funds. The majority of schools were found to have implemented coaching programs for principals and/or teachers. These programs involved continuous monitoring and feedback to teaching staff and school leaders. 98% of schools were found to have used SSNP funding on literacy and numeracy coaches. These coaches observed lessons and provided advice on improving teaching methodology and assessment practices. 94% of schools additionally implemented a similar program for school leaders. This entailed coaches who observed principals' daily routines and advised on more effective administrative practices, staff management and communication strategies amongst others.

Finally I conduct an analysis of staff and student opinion surveys to assess potential pathways by which these changes are occurring. Staff and student opinion surveys regularly maintain high response rates and pertain to all aspects of life and learning at school and in the classroom. Using the annual teaching staff survey, I run levels and growth models with the factor of interest as the dependent variable and an IV setup as above. The results mirror similar models run on student samples, showing the greatest increase in both staff and student opinion due to the SSNP relates to teaching and learning inside the classroom. Students report an increase of approximately 40% of a standard deviation in their perception of the teaching practices they experience. Staff were found to exhibit rises of at least one-third a standard deviation in self-reported assessment of their teaching skills. Other channels by which the SSNP appear to have operated include higher reported support from the school leadership team, greater goal congruence amongst staff and a general improvement in team practice. All of the above were areas targeted directly or indirectly by principal coaches suggesting that a large part of the effect was related to principals' improved running of their schools.

VIII Conclusion

This paper has found substantial effects of the Smarter Schools National Partnership on growth in student achievement. The largest intervention of its kind, the results in this paper show the SSNP has succeeded on some fronts yet no evidence of improvements due to the program were found in other areas. Indeed the program varied in effectiveness by year level and domain. Secondary school students appeared to have gained more than their primary school counterparts while Numeracy was more positively affected than Reading. These results can be explained by the nature of those domains. Reading literacy skills for instance are more dependent on home resources than Numeracy. Moreover, there are key foundation skills necessary for Reading which may take longer to develop than some of the formulaic techniques in Numeracy problem solving.

Some important policy insights emerge from examining the drivers behind these results. Rather than fund new initiatives or highly specific programs, the SSNP was a new experience in Australia's education landscape. Funds provided to schools could be spent in the manner schools chose. It can therefore best be thought of as an intervention enhancing school autonomy. Despite full control over a school's budget in Victorian Government schools, principals often have limited discretionary spending remaining once salaries and mandatory support programs are fulfilled. As a result, it can be difficult for schools to develop tailored programs that serve its students best. With the SSNP, schools appear to have expanded already-run programs or embarked on new initiatives which principals anecdotally reported prior intent on implementing.

Interestingly, although schools could realistically apply any intervention that met SSNP goals, nearly all schools implemented coaching programs for principals and teachers. The details of these coaching programs merit further study to better understand why some succeed in raising growth in student achievement.

Potential implementation issues arise as part of our analysis. The discrepancy in per-student funding across regions is markedly high and is likely to explain differential growth rates due to the program. Therefore any simple analysis that fails to account for regional differences or funding levels is likely to report biased effects of the SSNP. This discrepancy may be explored by education authorities as the most disadvantaged

students received the least per-capita funding as a result of this program. Furthermore, the optimal funding rate was found to be less than the median provided to each school suggesting greater efficiencies in resource allocation can be attained.

Finally, it is worth reiterating that this paper evaluates the SSNP at an early stage of its lifetime. Similar interventions in other countries were found to have had dynamic effects on student achievement, particularly in Reading, that may not surface one year following the program. Moreover, given the substantial positive effects on staff and student opinion factors - themselves leading indicators of future success as evidenced by historical data - there are likely to be further gains from the Smarter Schools National Partnership.

References

- ANGRIST, J., AND V. LAVY (1999): "Using Maimonodes Rule to Estimate the Effect of Class Size on Scholastic Achievement," *Quarterly Journal of Economics*, 114(2), 533–575.
- BALLOU, D., W. SANDERS, AND P. WRIGHT (2004): "Controlling for Student Background in Value-Added Assessment of Teachers," *Journal of Educational and Behavioral Statistics*, 29(1), 37–65.
- BLACKBURN, M., AND D. NEUMARK (1993): "Omitted-Ability Bias and the Increase in the Return to Schooling," *Journal of Labor Economics*, 11(3), 521–544.
- (1995): "Are OLS Estimates of the Return to Schooling Biased Downward? Another Look," *The Review of Economics and Statistics*, 77(2), 217–230.
- BOYD, D., P. GROSSMAN, H. LANKFORD, S. LOEB, AND J. WYCKOFF (2006): "How changes in entry requirements alter the teacher workforce and affect student achievement," *Education Finance and Policy*, 1(2), 176–216.
- BOYD, D., H. LANKFORD, S. LOEB, J. ROCKOFF, AND J. WYCKOFF (2008): "The narrowing gap in New York City teacher qualifications and its implications for student achievement in high poverty schools," Working Paper 14021, National Bureau of Economic Research.
- BRANCH, G., E. HANUSHEK, AND S. RIVKIN (2012): "Estimating the Effect of Leaders on Public Sector Productivity: The Case of School Principals," Working Paper 17803, National Bureau of Economic Research.
- COELLI, M., AND D. GREEN (2012): "Leadership effects: school principals and student outcomes," *Economics of Education Review*, 31(1), 92–109.
- COLEMAN, J., E. CAMPBELL, C. HOBSON, J. MCPARTLAND, A. MOOD, AND F. WEINFELD (1966): "Equality of Educational Opportunity," Discussion paper, U.S. Government Printing Office, Washington, DC.
- COUNCIL OF AUSTRALIAN GOVERNMENTS (2008): "Council of Australian Governments Meeting Outcomes," COAG.
- CURRIE, J., AND D. THOMAS (2000): "School Quality and the Longer-Term Effects of Head Start," *Journal of Human Resources*, 35(4), 755–774.
- DHUEY, E., AND J. SMITH (2011): "How Important Are School Principals in the Production of Student Achievement?," Working Paper 2011-33, CLSRN.
- FULLER, W., C. MANSKI, AND D. WISE (1982): "New Evidence on the Economic Determinants of Postsecondary Schooling Choices," *Journal of Human Resources*, 17(4), 477–498.
- GREENWALD, L., V. HEDGES, AND R. LAINE (1996a): "The Effect of School Resources on Student Achievement," *Review of Educational Research*, 66(3), 361–396.
- (1996b): "Interpreting Research on School Resources and Student Achievement: A Rejoinder to Hanushek," *Review of Educational Research*, 66(3), 411–416.
- HAHN, J., P. TODD, AND W. VAN DER KLAUW (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69(1), 201–209.

- HANUSHEK, E., AND D. KIMKO (2000): "Schooling, Labor-Force Quality, and the Growth of Nations," *American Economic Review*, 90(5), 1184–1208.
- HANUSHEK, E., S. LINK, AND L. WOESSMANN (2011): "Does School Autonomy Make Sense Everywhere? Panel Estimates From PISA," Working Paper 17591, National Bureau of Economic Research.
- HANUSHEK, E., AND L. WOESSMANN (2008): "The Role of Cognitive Skills in Economic Development," *Journal of Economic Literature*, 46(3), 607–668.
- HANUSHEK, E. A. (1997): "Assessing the Effects of School Resources on Student Performance: An Update," *Educational Evaluation and Policy Analysis*, 19(2), 141–164.
- (2003): "The failure of input-based schooling policies," *Economic Journal*, 113(485), F64–F98.
- HELAL, M., B. HOUNG, AND M. JUSTMAN (2012): "Estimating value-added from student growth percentiles: Tracking student progress and evaluating school performance," Working paper, Melbourne Institute of Applied Economic and Social Research.
- HOXBY, C. (2000): "The Effects of Class Size on Student Achievement: New Evidence from Population Variation," *The Quarterly Journal of Economics*, 115(4), 1239–1285.
- IMBENS, G., AND T. LEMIEUX (2008): "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 142(2), 615–635.
- IMBENS, G., AND J. WOOLDRIDGE (2009): "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, 47(1), 5–86.
- JACOB, B., AND L. LEFGREN (2004): "Remedial Education and Student Achievement: A Regression-Discontinuity Analysis," *The Review of Economics and Statistics*, 86(1), 226–244.
- KANE, T. J., J. E. ROCKOFF, AND D. O. STAIGER (2008): "What Does Certification Tell us About Teacher Effectiveness? Evidence from New York City," *Economics of Education Review*, 27(6), 315–631.
- KANE, T. J., AND D. O. STAIGER (2008): "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation," Discussion paper, NBER.
- KRUEGER, A. (1999): "Experimental Estimates Of Education Production Functions," *The Quarterly Journal of Economics*, 114(2), 497–532.
- LAVY, V. (2011): "Regression Discontinuity Designs," CEMFI.
- LEE, D., AND T. LEMIEUX (2010): "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 48(48), 281–355.
- LEIGH, A. (2010): "Estimating teacher effectiveness from two-year changes in students' test scores," *Economics of Education Review*, 29(3), 480–488.
- LEITHWOOD, K., K. SEASHORE-LOUIS, S. ANDERSON, AND K. WAHLSTROM (2004): "How leadership influences student learning," Report, The Wallace Foundation.
- MURNANE, R., J. TYLER, AND J. WILLETT (2000): "Do the Cognitive Skills of School Dropouts Matter in the Labor Market?," *Journal of Human Resources*, 35(4), 748–754.

- MURNANE, R., J. WILLET, AND F. LEVY (1995): "The Growing Importance of Cognitive Skills in Wage Determination," *The Review of Economics and Statistics*, 77(2), 251–266.
- OECD (2010): "Education at a Glance 2010," Discussion paper, OECD.
- RIVKIN, S. (1995): "Black/white differences in schooling and employment," *Journal of Human Resources*, 30(4), 826–852.
- RIVKIN, S. G., E. A. HANUSHEK, AND J. F. KAIN (2005): "Teachers, schools, and academic achievement," *Econometrica*, 73(2), 417–458.
- THISTLETHWAITE, D., AND D. T. CAMPBELL (1960): "Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment," *Journal of Educational Psychology*, 51(6), 309–317.
- WORD, E., J. JOHNSTON, H. BAIN, B. FULTON, C. ACHILLES, M. LINTZ, J. FOLGER, AND C. BREDI (1990): "The State of Tennessee's Student/Teacher Achievement Ratio (STAR) Project: Technical Report 1985–1990," Discussion paper, Tennessee State Department of Education.

Table 1 Commonwealth Funding by State (\$m)

	2009	2010	2011	2012	2013	Total	Proportion
NSW	4	60	81	149	143	437	39%
VIC	2	28	38	69	66	203	18%
QLD	2	24	32	58	56	172	16%
WA	1	10	13	24	23	71	6%
SA	1	16	22	40	39	118	11%
TAS	1	7	10	18	17	53	5%
ACT	0	0	0	1	1	2	0.2%
NT	1	7	10	18	17	53	5%
Total	12	152	206	377	362	1109	

Table 2 - NP in Victorian Government Schools

	LIT/NUM NP	Low SES NP	Both	Total
Number of schools	96	220	18	334
Number of students	54,382	75,208	8,052	137,642
Proportion of all students	10%	14%	1.5%	25%
Number of indigenous students	286	900	60	1246
Proportion of all indigenous students	8%	24%	2%	34%

Source: Author's calculations using linked data from DEECD Schools, Strategy/Policy and Finance divisions.

Table 3 – Background Characteristics of Treated Students

	Students in non-NP schools	Students in NP schools
Male	51.4%	51.6%
LBOTE	18.1%	33.9%
ATSI	0.8%	1.9%
<i>Indicators</i>		
Mother and father both unemployed in last 12 months	3.4%	11.6%
Mother or father a manager or senior manager	44.9%	18.3%
Mother or father tertiary educated	29.1%	10.8%
Mother and father have Year 10 or less qualification	3.9%	11.5%
N	48,987	15,165

Table 4 - The Effect of the Smarter Schools National Partnership

	Year 3 to Year 5				Year 7 to Year 9			
	NUM_2011	NUM_2011	RDG_2011	RDG_2011	NUM_2011	NUM_2011	RDG_2011	RDG_2011
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NUM_2009	0.484*** (0.005)	0.438*** (0.022)	0.167*** (0.005)	0.179*** (0.023)	0.763*** (0.004)	0.756*** (0.015)	0.204*** (0.005)	0.230*** (0.016)
RDG_2009	0.180*** (0.004)	0.174*** (0.019)	0.515*** (0.004)	0.484*** (0.020)	0.119*** (0.004)	0.122*** (0.013)	0.595*** (0.005)	0.577*** (0.017)
SSNP	3.763*** (0.868)	6.757* (3.413)	1.700 (0.948)	4.212 (3.633)	1.518** (0.578)	27.33*** (3.123)	-0.320 (0.642)	7.819** (2.830)
Male	12.58*** (0.516)	10.39*** (2.404)	-6.591*** (0.563)	-0.883 (2.571)	5.935*** (0.440)	6.014*** (1.114)	-1.113* (0.489)	0.671 (1.161)
ICSEA_2011	16.64*** (0.763)	8.965* (3.852)	16.24*** (0.833)	16.17*** (4.126)	9.227*** (0.667)	7.547*** (1.796)	10.32*** (0.740)	6.608*** (1.838)
School_size2011	0.00380*** (0.001)	0.00365* (0.002)	-0.000883 (0.001)	0.00214 (0.002)	0.00326*** (0.001)	0.0125*** (0.002)	-0.000808 (0.0005)	0.00329* (0.001)
SFODensity2008	-20.87*** (2.050)		-27.90*** (2.238)	52.35 (85.89)	3.742 (2.040)		-0.532 (2.263)	
N	33083	1533	33175	1547	26370	4389	26369	4406
R-sq	0.575	0.484	0.580	0.534	0.749	0.725	0.646	0.620

Dependent variable is 2011 student Numeracy score in Year 5 (Year 7) for primary (secondary) schools. All models control for demographic and socioeconomic backgrounds as well as and prior scores in 2009. The local average treatment effect is given by the coefficient on the treatment status variable SSNP.

Even-numbered columns present value-added regressions within an optimised bandwidth around the cutoff, with the running variable centred at the cutoff and interacted with treatment status.

Odd-numbered columns present a model with similar controls estimated over the entire range of the selection criteria with the running variable and a dummy for SSNP participation as regressors.

Robust standard errors in parentheses * p<0.05, ** p<0.01, *** p<0.001

Table 5 - IV Local Average Treatment Estimates

	Year 3 to Year 5		Year 7 to Year 9	
	NUM_2011	RDG_2011	NUM_2011	RDG_2011
SSNP	5.686** (1.895)	-2.947 (2.065)	36.26*** (3.535)	12.10** (3.729)
NUM_2009	0.481*** (0.00452)	0.166*** (0.00492)	0.766*** (0.00483)	0.205*** (0.00499)
RDG_2009	0.181*** (0.00399)	0.513*** (0.00435)	0.130*** (0.00494)	0.727*** (0.00409)
male	12.48*** (0.510)	-6.786*** (0.556)	5.934*** (0.483)	3.265*** (0.500)
disLBOTE_2011	-7.393*** (1.207)	-6.411*** (1.315)	-5.218*** (0.966)	-5.436*** (1.023)
ATSI	-6.936** (2.371)	-5.192* (2.591)	-6.556** (2.404)	-6.419* (2.496)
mgr_2011	1.689** (0.634)	5.008*** (0.691)	-0.348 (0.581)	1.941** (0.617)
uni_2011	11.83*** (0.676)	9.729*** (0.736)	8.560*** (0.672)	10.26*** (0.712)
unemp_2011	0.392 (1.204)	0.163 (1.310)	1.194 (1.087)	-1.943 (1.148)
gr10_2011	1.674 (1.251)	-3.691** (1.359)	0.977 (1.010)	-2.466* (1.071)
SFODensity2011	-23.93*** (2.877)	-24.97*** (3.131)	-69.31*** (6.427)	-37.31*** (6.780)
school_size2011	-0.000993 (0.00115)	-0.00885*** (0.00125)	0.00143* (0.000672)	-0.000418 (0.000714)
N	33870	33959	24851	25413
R-sq	0.575	0.581	0.723	0.627

Results represent reduced form results of an IV-model with treatment estimated as a function of student covariates, the running variable and a dummy for crossing the threshold in the first stage. Dependent variable is 2011 student Numeracy score in Year 5 (Year 7) for primary (secondary) schools. The local average treatment effect is given by the coefficient on the treatment status variable SSNP.

Robust standard errors in parentheses

* p<0.05, ** p<0.01, *** p<0.001