

Connolly, Marie; Haeck, Catherine

Working Paper

Nonlinear class size effects on cognitive and noncognitive development of young children

Research Group on Human Capital - Working Paper Series, No. 18-01

Provided in Cooperation with:

University of Quebec in Montreal, School of Management Sciences (ESG UQAM), Research Group on Human Capital (GRCH)

Suggested Citation: Connolly, Marie; Haeck, Catherine (2020) : Nonlinear class size effects on cognitive and noncognitive development of young children, Research Group on Human Capital - Working Paper Series, No. 18-01, Université du Québec à Montréal, École des sciences de la gestion (ESG UQAM), Groupe de recherche sur le capital humain (GRCH), Montréal

This Version is available at:

<https://hdl.handle.net/10419/228763>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Research Group on Human Capital Working Paper Series

Nonlinear Class Size Effects on Cognitive and Noncognitive Development of Young Children

Working Paper No. 18-01

Marie Connolly and Catherine Haeck

December 2020 (revised version)



Groupe de recherche sur le
CAPITAL HUMAIN
ESG UQAM

<https://grch.esg.uqam.ca/en/working-papers-series/>

Nonlinear Class Size Effects on Cognitive and Noncognitive Development of Young Children

Marie Connolly and Catherine Haeck*

Research Group on Human Capital, Université du Québec à Montréal

December 2020

Abstract

We estimate the nonlinear impact of class size on student achievement by exploiting regulations that cap class size at 20 students per class in kindergarten. Based on student-level information from a previously unexploited and unique large-scale census survey of kindergarten students, this study provides clear evidence of the nonlinearity of class size effects on development measures. While the effects are largest on cognitive development, class size reductions also improve noncognitive skills for children living in disadvantaged areas. These findings suggest that sizeable class size reductions targeted at disadvantaged areas would achieve better results than a marginal reduction across the board.

JEL Codes: I21, I28, J24, C31

Keywords: class size, cognitive development, noncognitive development, kindergarten, nonlinear effects

* *Catherine Haeck*: haeck.catherine@uqam.ca. We gratefully acknowledge financial support from the Fonds de recherche du Québec - Société et Culture. We thank Fabian Lange, Daniel Parent, Dalibor Stevanovic, Philip Merrigan, anonymous referees, and also participants at the CEA, SCSE and SOLE meetings, the Boise State University seminar, the Université de Sherbrooke seminar, the GRCH workshop at the Université du Québec à Montréal and the IRLE visitors' workshop at UC Berkeley for their comments. We also want to thank the Institut de la statistique du Québec research data centre (CADRISQ) for great support throughout this project. David Lapierre provided skillful research assistance. The analysis is based on the Québec Survey of Child Development in Kindergarten restricted-access Micro Data Files available at the CADRISQ. All computations on these microdata were prepared by the authors who assume responsibility for the use and interpretation of these data.

1 Introduction

This article evaluates the nonlinear effects of class size on student achievement by exploiting regulations that set maximum class size at 20 pupils per class in kindergarten, adapting Angrist and Lavy’s (1999) causal identification strategy to account for nonlinearities. We carefully show that class size effects vary along the class size distribution: reducing class size from 15 to 12 does not produce the same effect on child development as reducing from 21 to 18. Mean effects may be hiding important variations, which has important implications for education policy and cost-benefit analyses. Our main contribution is to provide the first estimates of nonlinear effects of class size on cognitive and noncognitive measures of child development for kindergarten students using unique student-level quasi-administrative data. Focusing on kindergarten students is relevant because early-life interventions have long-lasting effects (Heckman, 2006), and because causal identification is facilitated by the fact that class composition manipulation is much less likely at the kindergarten level since school teachers and administrators know very little of the child before school starts.

Class size reduction is a popular educational reform. It is also costly, making it important to have good estimates of its benefits to be able to gauge its effectiveness vis-à-vis other education policies. The benefits of smaller classes on student achievement¹ have been documented in a number of studies using credible empirical strategies, but the debate remains ongoing (Hanushek, 1999, 2002). Identification is complicated by the fact that class size is not random in most settings: pupils with learning or behavioral disabilities may be placed in smaller classes, and children of high-income parents and of higher ability may also have access to schools with more resources and smaller classes. Theory also suggests that optimal class sizes are larger for better-behaved students. Models are often based on disruptions, with turbulent students in the classroom disrupting teaching and negatively affecting student learning (Lazear, 2001). The larger the class, the more likely disruptions will be. Lazear’s model is inherently nonlinear: the probability of disruption increases at a faster rate for

¹Some studies have also been interested in class size effects on noncognitive measures of development, such as disruptiveness, inattentiveness, motivation, effort, or self-confidence (Dee and West, 2011; Ding and Lehrer, 2011; Chetty et al., 2011).

small classes than it does for large classes, in which disturbances are already more common. In consequence, the effect of reducing class size depends on the size of the class, and class size reduction has a larger impact in smaller classes.

Yet the overwhelming majority of studies on class size in primary school presents estimates based on linear models, meaning that the estimated effect of reducing class size is constant over the class size distribution. Exceptions include Hoxby (2000), Urquiola (2006), and Hojo (2013).² Urquiola (2006) looks at the effects of class size reductions on third graders' achievement in Bolivia, where the maximum class size is 30. His comparison of two empirical strategies leads him to conclude that his evidence is consistent with nonlinearities in the class size effects, though without estimating the nonlinearities directly. Hojo (2013) uses a piecewise-linear function in a setting where maximum class size is 40. Using data on mathematics and science test scores of around 4,500 fourth graders in Japan, he finds that class size reduction generates a positive impact on test scores when the reduction occurs below 23. The relevance of these estimates for educational systems with overall much smaller classes on average, such as the United States where the average class size at the elementary school level was 21 in 2017 (OECD, 2017), is less clear. Hoxby (2000) uses school- and district-level data from Connecticut to estimate the impact of class size. She uses the natural logarithm of class size to account for nonlinearity. She finds that class size reduction does not improve September test scores in fourth and sixth grades. Class size is measured in prior school years and assumes that children do not change schools. In her setting, class size rules vary between districts which, as she mentions, implies that parents with unobserved characteristics positively related to test scores may choose to buy a residence in a district with smaller class sizes.

In this paper, we use student-level quasi-administrative data to estimate nonlinear effects of class size on child development in the Canadian context, which is comparable to the American context. In contrast with Hoxby (2000), class size rules are the same for everyone across the entire province, thereby eliminating this source of bias. Also, because we have

²There are also a few studies using postsecondary education data. Bandiera et al. (2010) estimate heterogeneous effects of class size on test scores, but for university students. Bedard and Kuhn (2008) also focus on higher education, this time looking at the nonlinear effects of class size on teacher evaluations.

student-level data on kindergarten children, we observe the actual class size of the child and relate it directly with their development several months after starting school. Our findings suggest that reducing class size from 20 to 19 in all classes would bring limited benefits at sizeable costs, but that reducing class size from 20 to under 15 in some targeted areas would highly benefit students' cognitive and behavioral skills. Such policy could even be cost-neutral, if paired with a small class size increase in the non-targeted areas. Our findings are different from Hoxby (2000) but our settings imply a few important differences. Although the range of class sizes we study is similar to Hoxby (2000), we focus on younger children who may benefit more from a reduction in that range. Indeed, Lazear's (2001) model suggests that the optimal class size is likely higher in better-behaved and older students than in younger kindergarten students. Also, since we have student-level data, we are able to specifically identify children living in finely-defined disadvantaged geographic areas, and find that only these children actually benefit significantly from smaller classes.

To conduct our analysis, we use previously-unexploited student-level data on more than 80 percent of kindergarten pupils in the province of Québec in 2012. We first present evidence on the nonlinearity of class size effects, then we modify Angrist and Lavy's (1999) and Fredriksson et al.'s (2016) empirical approaches to account for those nonlinear effects. The nonlinearity in our main model comes from the use of a logarithmic function of class size. We also estimate two alternative nonlinear specifications: the first one based on a quadratic function of class size, the second one using a step function to get the effect of a large class compared to a small one. In all models, we instrument class size using a function of school enrollment, akin to Angrist and Lavy's use of Maimonides' rule, as well as school enrollment thresholds where class size is expected to drop, following Fredriksson et al. (2016). The instruments are strong predictors of class size: actual class size tightly follows predicted class size. However, our setting is more in line with Project STAR (Krueger, 1999; Krueger and Whitmore, 2001; Chetty et al., 2001) in the sense that we study class size variations between nine and 24 students in kindergarten, while Angrist and Lavy (1999) look at variations between 20 and 40 students in third to fifth grades. In light of recent findings by Angrist et al. (2017) and Fredriksson et al. (2016), we also make sure our estimates are not biased by

teachers' strategic behavior or bunching around enrollment thresholds.

We find that reducing class size below 16 students per class positively impacts cognitive development, but that class size has to drop below 15 to have an effect on social competence and communication skills. We find no significant impact of class size reduction on emotional maturity. Average effects across the class size distribution, as typically estimated in the literature, are small for pupils' cognitive development but not statistically different from zero for noncognitive measures. Our results provide a deeper understanding of the effects of class size on student achievement by showing that mean impacts mask the fact that a marginal class size reduction of one or two students may be completely ineffective at improving cognitive and noncognitive outcomes.

Our data set allows us to disaggregate our results by gender, school poverty status, and child material deprivation quintile. We do not observe differences by gender: the gender-specific effects are fairly similar and in line with our overall results. Looking at the effects of class size for children in disadvantaged areas, we find large and significant improvements for both cognitive and noncognitive development when classes are small (15 students or fewer). Not only do students from disadvantaged backgrounds benefit the most, they are also those that are the farthest behind in terms of development.

From a policy standpoint, our findings suggest that while class size reduction policies always imply increased expenditures, they may not always trigger cognitive and noncognitive gains for students. Large and targeted class size reduction policies in disadvantaged areas are far more efficient than modest universal reduction policies. Such targeted policies could help level the playing field and promote more equal opportunities, as argued by Piketty and Valdenaire (2006) in their study of class size in France with a focus on disadvantaged areas. From a budgetary standpoint, our findings further suggest that large reductions in a limited number of classes could be financed by marginal increases in the vast majority of schools not experiencing high poverty rates.

The rest of the paper is organized as follows. We start with a presentation of the institutional context and of the data set in sections 2 and 3. Section 4 presents the identification strategy while section 5 discusses our findings. The final section concludes.

2 Institutional background

We present here some relevant institutional background on kindergarten and schools in Québec. The general context is quite similar to most of the United States. Children start kindergarten at age five: a child must be five years old by September 30 of the entering school year. Schools are typically divided between elementary schools, covering kindergarten through grade six, and secondary schools, for grades 7 and up. Elementary schools are on average smaller, and thus located closer to home. Children attend school 182 days a year, Monday to Friday, roughly from September to June, inclusively. Schedules vary from school to school, but total instruction time is usually just under five hours each day in kindergarten, not including lunch. Before- and after-school care, as well as lunch supervision, is subsidized by the Québec government, with a rate currently (and in 2012) set at C\$8 a day. Most children attend public schools, for which enrollment is free: 93% of kindergarten-level students in 2011-2012 were in public schools. Education is under provincial jurisdiction, not federal. Public spending on elementary and secondary education was 3.9% of GDP in 2009-2010, compared to 4% in the United States (MEESR, 2015). Until 2020, public schools in Québec were organized by school boards. Boards, or districts, were divided in parallel boards according to language: there were 60 French-language boards, nine English-language boards, and one board serving both French and English schools, located in a remote area. Two additional boards regrouped Indigenous communities from northern Québec. Most children attend school in French: 90% of students in 2011-2012 for kindergarten.³ Children are assigned to a public school based on the location of their residence. If they live within a school's catchment area, they are guaranteed a spot. Parents also have the option to enroll their child to another school within the same school board as availability permits. Changing school board is more complicated and rarely done.

Québec generously subsidizes early-childhood care, either directly funding public low-cost daycare centers, or through refundable tax credits available for parents sending their child to a private daycare. Those subsidies stop when the child reaches kindergarten age,

³Most statistics in this section come from the Databank of Official Statistics on Québec, available at www.bdso.gouv.qc.ca.

at which point the vast majority of children switch to the public elementary school system. Pre-kindergarten programs, also called junior kindergarten or four-year-old kindergarten in other contexts, have been offered in targeted schools from high deprivation areas as early as 1978 (Government of Québec, 2012). These programs were offered part-time (morning or afternoon, 5 days per week) on a voluntary basis. Parents had the option to send their child to pre-kindergarten if they lived in the school’s catchment area and their child was four years of age by September 30 of the school year. One of the goals of these programs was to facilitate the transition to school for low socioeconomic status children. In school year 2012-2013, our study period, 81,541 students were enrolled in kindergarten, but only 16,910 of these students had attended part-time pre-kindergarten programs in a school setting the previous year (roughly 21%). Since 2013, pre-kindergarten is now offered full-time in a larger number of schools, and the network is expanding.

Maximum class size is regulated by the Québec government in its agreement with the school teachers. The rule is therefore the same in all schools across the province. In Hoxby (2000), the rule varied by school district with a modal maximum class size of 25. While parents in Connecticut could strategically choose their area of residence to send their child to a school with smaller classes, parents in Québec do not have that option. In fact, they are generally completely unaware of the actual class size until the start of the school year. For kindergarten, the maximum class size has been set at 20 since year 2000. Private schools are not subject to this class size cap; the average class size in private schools is higher than in public schools. This cap at 20 creates a situation that is similar to that in most of the United States. Hoxby (2000) reports that most school districts in the United States have maximum class size rules ranging between 20 to 30 students. A majority of states have caps, or no hard cap but tie funding to class size. For example, California provides additional funding for classes with no more than 20 students in kindergarten (Gilraine et al., 2018). New York sets maximum class size at 20 as well, Texas has a cap of 22, and Florida uses 18.⁴

⁴To see a map of the United States along with class size information, look at Education Week’s infographics here: https://www.edweek.org/ew/section/infographics/13class_size_map.html.

3 Data

Our data come from the 2012 Québec Survey of Child Development in Kindergarten (QSCDK), a large-scale census survey conducted by the Institut de la Statistique du Québec (ISQ), the provincial equivalent to Statistics Canada. The QSCDK 2012 collected information on more than 80% of students enrolled in kindergarten in academic year 2011-2012 (Simard et al., 2013). The goal of this survey was to provide a detailed picture of the development of Québec children. The QSCDK has not been previously exploited to look at the effects of class size. In fact, only a couple of studies in the field of public health have used the QSCDK data (Laurin et al., 2015, 2018), making this paper the first study in economics to avail itself of these rich data. Overall, 98% of Québec’s kindergarten students are targeted by the survey. Identified at-risk students (students with learning difficulties or special needs) registered in classes with 50% or more at-risk students are excluded from the QSCDK target population, as well as schools catering specifically to special needs students. Also excluded from the target population are students in schools under federal jurisdiction (mostly schools on First Nations reserves) or in schools in Québec’s two Indigenous school boards, located in the northern areas of the province. Participation rate is extremely high: 69 out of 70 school boards and 88% of Québec’s schools participated in the survey, for a combined response rate at the student level of more than 80%.

Five domains of development are measured in the QSCDK: (1) cognitive and language development, (2) social competence, (3) emotional maturity, (4) communication skills and general knowledge, and (5) physical health. These five domains are assessed using the Early Development Instrument (EDI) developed at McMaster University in Ontario, Canada (Janus and Offord, 2007). The questionnaire filled out by teachers, online in 95% of the cases, is designed to assess the strengths and difficulties of children regarding the five domains.

In the EDI, the child’s development is measured using a total of 104 questions (Yes/No/Don’t know) split into three sections. Additional questions are included to gather information on child characteristics (15 questions), disabilities (17 questions), early childhood experiences (13 questions), and specific abilities (7 questions). In the EDI, the cognitive and language

development domain refers to the child’s competence in mathematics and reading. It is measured, among other things, by the child’s ability to recognize and use numbers from 1 to 10, read and write simple words or sentences, compare numbers, and identify concepts of time. Communication skills are measured by the ability of the child to listen, tell a story, pronounce clearly, and communicate their needs clearly. Social competence refers to the ability to play with others, follow the rules, act responsibly, adapt to change, and be autonomous. Emotional maturity captures whether the child has a tendency to hurt others physically or mentally, help others when in need, wait, and be patient, calm, and attentive. Additional information about the EDI can be found in Janus and Offord (2007). The EDI has been used not only in Canada, but also in Australia, the United States, and England (e.g., Janus et al., 2011; Brinkman et al., 2013), and has been shown to accurately measure the state of development of children. Janus et al. (2011) specifically test the validity of the measure, and show results against the well-known PPVT in four countries. Furthermore, recent research on the EDI shows that it strongly predicts children’s literacy and numeracy assessments at ages 8, 10, and 12 (Brinkman et al., 2013). The questionnaire is provided in the Appendix. While the EDI is not a high-stakes exam, it is a detailed measure with proven external validity. The advantage of not being a test, and definitely not a high-stakes test, alleviates concerns that can come from high-stakes testing (Lazear, 2006; Rockoff and Turner, 2010) and from score manipulation (Angrist et al., 2017).

One may be concerned by the fact that the EDI relies solely on teachers’ answers. Forget-Dubois et al. (2007) specifically test the validity of the EDI against more traditional test-based measures administered by trained professionals in a longitudinal survey of Québec children. They find that the EDI, especially its cognitive component, predicts later school achievement in grade 1 over and above the cognitive assessments and direct school readiness tests included in their survey, namely the PPVT-III (Dunn and Dunn, 1997), the Block Design subtest of the Wechsler Preschool and Primary Scale of Intelligence–Revised (Wechsler, 1989), the Number Knowledge Test (Okamoto and Case, 1996) and the Visually Cued Recall task (Zelazo et al., 2002). Their assessment concludes that the cognitive and language development score of the EDI is one of the best predictors of school achievement among the

measures available in their study. They further conclude that teachers’ assessment using the EDI questionnaire in kindergarten can be almost as effective as a formal battery of tests that rely upon trained professionals. However, the EDI has some limits which must be acknowledged. While all teachers in Québec have completed a 4-year-long university degree, combined with several months of training on school premises and are considered highly qualified, assessments may vary across teachers. To ensure the EDI is completed carefully by teachers, the ISQ ensures that a substitute teacher is hired to allow each teacher to answer the questionnaires in a separate room during the teacher’s working hours. Teachers take on average 15 to 20 minutes to answer each student’s questionnaire and time is monitored by the Québec Statistical Institute who runs the survey. Finally, teachers answer the questionnaires after having observed the child for several months.

We use as outcomes the cognitive, social, emotional, and communication measures computed from the various Yes/No questions and provided in the data set. Physical health measures are used as controls. In addition to the developmental measures, the QSCDK provides information at the school and student levels, which we also mostly use as control variables. At the student level we have information on gender, age in months, mother tongue, place of birth, number of years since child immigrated to Canada, quintile of social and material deprivation⁵ of residential area, whether the child attended childcare or attended pre-kindergarten in a school setting (also called 4-year-old kindergarten or junior kindergarten in other contexts), has social or learning disabilities (physical limitations, visual deficiencies, hearing deficiencies, problems at home, chronic health conditions, dental issues), and has required help from a non-teaching professional at school (nurse, speech therapist, psychoeducator, social worker, psychologist). School-level variables include total kindergarten enrollment, class size for each group, administrative region, school board, and teaching language (French or English).

⁵Social and material deprivation indices are provided in the database. The Institut de la statistique du Québec computes the indices using a principal component analysis based on five measures of the 2006 Census data for each dissemination area (about 400 to 700 persons in Canada) (Section 2.3, Institut de la statistique du Québec, 2013). Social deprivation mainly reflects a high proportion of individuals that are separated, divorced or widowed, and a high percentage of single families in the area. Material deprivation mainly reflects a high proportion of high school dropouts, a low proportion of individuals employed, and low average income in the area.

The unit of observation in the QSCDK data is the student. From the overall sample, we exclude students in private schools, which are not subject to the class size regulation (4.6 percent), and students in mixed⁶ classes (1.4 percent). We are left with 58,949 students in more than 3,600 classes spread across 1,484 schools. The average class size is 17.5 and the average enrollment is 59.0. Figure 1 shows the class size distribution in our sample. Class size varies between nine and 24,⁷ with less than two percent of students being in classes of more than 21 students. Student characteristics are reported in Table 1. Teachers filled out the questionnaire near the end of the school year, when students were six years old on average. Most students speak French at home (77%), attended childcare prior to kindergarten (62%), but did not attend pre-kindergarten in a school setting (79%). Around 6% of the children have learning disabilities or behavioral issues. Speech difficulties are present for 5% of the children. Indicators of health issues suggest that most children are in good health.

For ease of interpretation, we standardize each of the four development measures that we use as outcomes to have a mean of zero and standard deviation of one,⁸ with larger values meaning a more positive outcome. Figure 2 shows the average developmental scores by class size. It appears that averages of each of the four scores are decreasing with class size between nine and 14 students, but beyond that point, average scores seem to be relatively stable, and even slightly increasing with class size. This pattern holds for all measures of child development but is strongest for cognitive development and communication skills. Figure 2 motivates us to estimate nonlinear models of the effect of class size on child development. Figure 2 also shows that the standard errors do not increase with class size, thus minimizing concerns about measurement error in the survey instrument coming from teachers in large groups reporting less accurately.⁹ The next section will explain how we tackle identification issues and control for observables to obtain causal estimates.

⁶Mixed classes contain students from different grade levels, for example, pre-kindergarten students and kindergarten students.

⁷Disclosure rules do not allow us to release the number of classrooms at size 24 because it is too small, but students in those classes are in the estimation sample.

⁸Our results are also robust to using percentile ranks instead of standardized (or raw) scores.

⁹Larger standard errors in smaller classes mostly come from the fact that such classes are less common (smaller n).

4 Empirical strategy

Our goal is to estimate the nonlinear effect of class size (c_{ijs}) on four measures of child development (y_{ijs}), where i is a child, j their class, and s the school. While class size may positively impact student test scores through reduced student disturbance as modeled by Lazear (2001), student ability may also influence student assignment to different class sizes. To identify the causal relationship between class size and student scores, we cannot solely rely on correlation between scores and class size; ordinary least squares (OLS) estimations may yield biased estimates. A first approach to pin down the causal effect of class size is to look at experimental evidence, such as the one provided by the influential Project STAR.¹⁰ Given the absence of suitable experimental data in our context, we rely on an instrumental variable (IV) approach, following Angrist and Lavy (1999),¹¹ who explained that in Israel, enrollment and class size were positively related through Maimonides' rule, which dictates the maximum number of students per class. We use governmental regulation on class size that sets the maximum number of students per class per grade to identify the causal impact of class size on student achievement. The class size rule provides us with an instrumental variable for our potentially endogenous regressor, actual class size. The required identifying assumption is that residual child development, after controlling for observables, is not correlated with the class size rule function.

In Québec, class size cannot exceed 20 students in kindergarten. As a result, the class size function takes the following form:

¹⁰Project STAR, which took place between 1985 and 1989 in Tennessee, randomly assigned kindergarten to grade 3 students to regular or small classes. Krueger (1999) and Krueger and Whitmore (2001) find that students assigned to smaller classes had higher test scores compared to students assigned to regular sizes. Black students and those receiving a free lunch benefited more from being assigned to a small class, and most of the effects documented were found for medium to high performing students (Konstantopoulos and Li, 2012; Jackson and Page, 2013; Mueller, 2013). In a study on the long-term effects of Project STAR, Chetty et al. (2001) find that students in smaller classes were more likely to attend college, but that teacher experience and peer quality also matter for long-term outcomes such as college attendance and earnings.

¹¹Angrist and Lavy (1999), using data on third to fifth graders in Israel, exploit exogenous variations in class size due to school enrollment to instrument for class size. This approach has been applied to a number of different contexts: in Bolivia (Urquiola, 2006), Denmark (Browning and Heinesen, 2007; Nandrup, 2016), Japan (Akabayashi and Nakamura, 2014), the United States (Cho et al., 2012; Chingos, 2012), and Sweden (Fredriksson et al., 2013), to name a few. The estimated effects of class size reduction are often positive, but modest, and sometimes not statistically different from zero. Other approaches include using enrollment variations across time (Hoxby, 2000), or a general equilibrium framework (Gilraine et al., 2018).

$$n_s = \frac{e_s}{\left[\text{int}\left(\frac{e_s-1}{20}\right) + 1 \right]}, \quad (1)$$

where n_s measures the average class size in school s given enrollment e_s . For enrollment cohorts of one to 20 students, there is only one class in the school. For cohorts of 21 to 40 students, there should be two classes in the school of 10.5 to 20 students per class on average. For cohorts of 41 to 60 students, there should be three classes in the school of 13.7 to 20 students, and so on as enrollment increases. The rule applies to all cohort sizes. In our data set the maximum enrollment is 125, which should translate to seven classes of 17.9 students on average. In practice, nothing restrains schools from choosing not to have classes of more or less equal sizes. However, in our data we observe that class size differences within a school do not exceed three students for 95% of schools.

Figure 3 illustrates the class size function along with the actual average class size given enrollment. This figure clearly shows that the average class size follows tightly the government rule, albeit more so when the enrollment does not exceed 80 students, thus ensuring the relevance of the instrument. Angrist and Lavy's (1999) strategy implicitly assumes that the impact of class size is linear, or that the impact of class size is uniform across the class size distribution per unit of treatment. If the impact of class size reduction is not uniform, but instead varies across the class size distribution, then this strategy captures the average effect of class size reduction in the studied sample.

We modify the standard IV model to account for nonlinearity using mainly the natural log of class size ($\ln(c)$), but also a quadratic function of class size, which we instrument using n_s and n_s^2 , and a step function, also instrumented with n_s and n_s^2 . These models allow us to capture nonlinearities (Wooldridge, 2002, section 9.5), and form the basis for our main estimates of class size effects on child development.

Our first model is estimated using these equations:

$$y_{ijs} = \alpha + \beta \ln(c_{ijs}) + \theta X_i + \gamma W_s + \varepsilon_{ijs} \quad (2a)$$

$$\ln(c_{ijs}) = \pi_1 + \zeta_1 n_s + \theta_1 X_i + \gamma_1 W_s + \epsilon_{1,ijs} \quad (2b)$$

In the first stage of the model using logs (equation 2b), $\ln(c_{ijs})$ is explained by the class size rule n_s (equation 1) and other student- (X_i) and school-level (W_s) covariates, including functions of enrollment. The second stage (equation 2a) has a linear-log form: developmental score y_{ijs} is explained by predicted log of class size $\widehat{\ln(c)}_{ijs}$ from the first stage and covariates. Student controls (X_i) include dummy variables indicating gender, student's age in months, place of birth, whether the child attended childcare or pre-kindergarten, ten markers of health and behavioral problems (physical disability, visual deficiency, auditive deficiency, speech disorder, learning difficulties, emotional problems, behavioral problems, disadvantaged family environment, chronic health conditions, and dental problems), and dummies to indicate whether the child received help from various school professionals (nurse, speech therapist, psychoeducator, social worker, and psychologist). School controls (W_s) include enrollment, enrollment square, poverty index (high or low), social and material deprivation indices (highly advantaged, average, highly disadvantaged), teaching language (French or English), and school board dummies.

To validate the robustness of our results using the linear-log model, we also estimate a model including a quadratic function of class size and different models using various step functions. The point of using a step function is also to ease the interpretation of the nonlinear effect, albeit while losing some of the refinement coming from the use of the linear-log model.

The quadratic model is estimated using these equations:

$$y_{ijs} = \alpha + \beta_1 c_{ijs} + \beta_2 c_{ijs}^2 + \theta X_i + \gamma W_s + \varepsilon_{ijs} \quad (3a)$$

$$c_{ijs} = \pi_1 + \zeta_{1,1} n_s + \zeta_{1,2} n_s^2 + \theta_1 X_i + \gamma_1 W_s + \epsilon_{1,ijs} \quad (3b)$$

$$c_{ijs}^2 = \pi_2 + \zeta_{2,1} n_s + \zeta_{2,2} n_s^2 + \theta_2 X_i + \gamma_2 W_s + \epsilon_{2,ijs} \quad (3c)$$

The step function model is similarly defined as follows:

$$y_{ijs} = \alpha + \beta I[c_{ijs} > c] + \theta X_{is} + \gamma W_s + \varepsilon_{ijs} \quad (4a)$$

$$I[c_{ijs} > c] = \pi + \zeta_1 n_s + \zeta_2 n_s^2 + \theta_1 X_{is} + \gamma_1 W_s + \epsilon_{1,ijs} \quad (4b)$$

where the endogenous variable $I[c_{ijs} > c]$ is equal to one if class size c_{ijs} is greater than c and zero otherwise. We explore different c cutoffs ranging from 12 to 18. The estimated coefficients can be interpreted as the effect of being in a larger class ($c_s > c$) in kindergarten compared to a smaller one.

In some specifications, we also borrow from Fredriksson et al. (2016), who use a regression discontinuity design to estimate class size effects in Sweden, where class size is capped at 30. Instead of using a class-size function to instrument class size, they observe that the class size rule generates multiple discontinuities, at each multiple of 30 in kindergarten enrollment. They thus use a set of enrollment threshold dummies, and add segment dummies as control variables, in essence defining a separate instrument at each threshold. Following Fredriksson et al. (2016), we supplement the Angrist and Lavy (1999) framework by adding enrollment thresholds as instruments, exploiting the discontinuities at the different jumps in class size as enrollment at the school level increases. In our application, going from a school enrollment of 20 to 21 induces a different jump in class size than going from an enrollment of 40 to 41, and so on. Our thresholds are: school enrollment of 21 students and above (tr_1), 41 students and above (tr_2), 61 students and above (tr_3), 81 students and above (tr_4), and 101 students and above (tr_5). Enrollment segments are used as controls. The segments define bands of school enrollment: 1 to 10 students in the school; 11–30; 31–50; 51–70; 71–90; 91–110.

The literature has highlighted a number of potential threats to the validity of the empirical approach based on school enrollment: score manipulation, bunching around the thresholds, and class composition. We address each of these concerns as they come up in the next section. On class composition, while this is a legitimate concern in most applications, in our setting student assignment is done before most children ever set foot in the school. Since there are no formal modes of evaluation in childcare or pre-kindergarten in Québec, teachers have very limited knowledge of each student’s ability when they assign students to different classes. In this sense, studying the impact of class size in kindergarten is particularly attractive since it is less likely that class size within a school depends on student ability. If students with more difficulties are assigned to smaller classes within the same school, then empirical evidence using OLS would detect a negative impact of class size on student achievement.

However, in the methodology described above, predicted class size is used as opposed to actual class size, such that it is mainly variation in class size between schools that drives the identification, as opposed to variation within schools. Lastly, we cluster standard errors at the school level in all of our estimations. Clustering at the school district level leads to similar standard errors while using only robust standard errors leads to smaller standard errors.

5 Findings

We now present the results from our empirical analysis. We start with a graphical description of our data and follow with estimates from linear models, both using ordinary least squares and IV. We then present estimates coming from nonlinear models. We also investigate the heterogeneity of our nonlinear estimates. We conclude this section with some robustness checks.

5.1 Graphical description

Before we present findings from our model estimations, we start with a graphical description of the data. Figure 4 reproduces Figure 2, but this time showing on the Y -axis the average residual of class size after controlling for school-board fixed effects, in effect providing a visual version of the first stage when we exploit each enrollment threshold as an instrument. The lines are fitted to individual-level data, by band of school enrollment. The drops at each vertical line (i.e. when school enrollment passes a threshold) are substantial and statistically significant. Figure 5 shows the corresponding jumps in cognitive development, by plotting the average residual of cognitive development by school enrollment, after controlling for student- and school-level covariates. At each threshold we observe an increase in the cognitive development residual. The fitted lines by band in Figure 5 also all slope downward, except for the 40-to-60 enrollment band, for which the line is rather flat. Those two figures, which in essence replicate Figures 5 and 6 in Fredriksson et al. (2016), show how the instrumental variable strategy operates: the class size rule induces drops in class size, which are matched

to jumps in developmental scores.

5.2 Linear models

For the sake of comparison with the literature estimating linear effects, we first present estimates of linear effects of class size on development scores. Those can be found in Table 2. In specification 1, we estimate a linear model using ordinary least squares without controls. We add school-level controls in specification 2, and student-level controls in specification 3. Using OLS without controls, we find that both cognitive and noncognitive scores tend to increase with class size. Estimated coefficients are generally positive and statistically different from zero, but their magnitude is very small, ranging between 0.5 to 1.5 percent of a standard deviation (SD). These results differ from Hoxby (2000), who found negative and significant coefficients using the naive model. This difference provides some support to the idea that in contrast with Hoxby (2000), whose setting allowed parents with unobserved good characteristics to self-select their child into schools with smaller classes, here parents do not have that option. Once we include school-level controls, the relationship becomes economically and statistically equal to zero for cognitive development, social competence, and emotional maturity, but remains positive and statistically significant for communication skills (0.008 SD). Adding student-level characteristics as controls makes the last remaining statistically-significant estimate become non different from zero. Overall, our naive OLS results accounting for student and school characteristics suggest that class size has no effect on student development.

Once we turn to the IV results, also presented in Table 2, the story changes slightly, but only for cognitive development. In columns 4 and 5, we present the results first using the class size rule, then adding the enrollment dummies (thresholds and segments), to instrument class size as described in the previous section, along with our student- and school-level controls. The impact of class size on students' cognitive development becomes about three times larger at -0.012 SD, but is only significant under specification 4. This suggests that reducing class size by one student increases the cognitive score by 0.012 SD. This effect

may appear small, but it needs to be compared to other estimates of class size effects in kindergarten. In the STAR experiment, kindergarten students who benefited from a small class size assignment scored higher on the Stanford Achievement Test¹² by an average of 0.20 SD (Krueger, 1999). Small classes had on average 15.1 students in kindergarten compared to regular groups of 22.4 students. Krueger’s STAR estimate thus measures the impact of a variation of 7.3 students on average. If we multiply our estimates by 7.3, we get an impact of 0.088 SD for cognitive development, about half the size of Krueger’s STAR. When we turn to noncognitive measures, we do not find any significant impacts when using IV, but the signs of the estimated coefficients suggest that class size reduction may have a small positive impact on child development. We do not report the first-stage estimates corresponding to the IV estimates of Table 2, but F -statistics are 172 and 513, showing that the instruments are not weak.

5.3 Nonlinear effects of class size

So far, we have measured the average impact of class size on student outcomes, yet as seen in Figure 2 the effect of class size may vary along the class size distribution. We may reasonably expect class size variations to have a larger impact in smaller classes (Lazear, 2001). One explanation may be that a student’s marginal impact on the overall learning environment of the class is larger the smaller the number of students in the class. For example, talking with one’s neighbor is more likely to disturb a class of 12 students than a class of 30 students since the probability that other students are also talking at the same time increases with class size.

If class size could be considered exogenous, one could simply regress developmental scores on class size dummies to lift the linear restriction imposed by a model estimated by OLS in which class size explains development. Doing so (and including student- and school-level controls) yields estimates that present a nonlinear pattern shown in Figure 6. This pattern is similar to that observed in Figure 2. Endogeneity may still be an issue, but traditional

¹²The Stanford Achievement Test measures achievement in reading, word recognition, and math in kindergarten.

statistics to test for endogeneity in an IV context are uninformative if the impact per unit of treatment is not uniform (Lochner and Moretti, 2015). This means that we cannot rely on usual tests to verify the validity of our instruments in this case. Lochner and Moretti propose an exogeneity test in a context similar to ours, and show that if the test fails to reject exogeneity, then the OLS estimator is consistent and IV is not required. We thus perform the test proposed by Lochner and Moretti, and unfortunately find that class size remains endogenous for most outcomes, even in the context of varying treatment effects.¹³ We therefore have to move to our nonlinear IV approach to address endogeneity and allow for nonlinearity.

If the effect of class size was linear, then reweighting our data such that our class size distribution becomes uniform would lead to a comparable estimated coefficient. Table 3 provides OLS and IV estimates (columns 1 and 2, respectively) of models in which student weights have been adjusted to ensure that the class size distribution becomes uniform. When students from smaller classes are given more weight, we find much larger coefficients on class size (in absolute value): the OLS point estimate goes from -0.005 to -0.012, and the IV one from -0.012 to -0.055. This finding supports the idea that the impact of class size is not linear. Using $\ln(c_{ijs})$ as an explanatory variable appears like a natural extension, since this allows changes in class size to have a smaller impact as the class size increases. Columns 3 to 7 of Table 3 show the estimated coefficients of $\ln(c_{ijs})$, corresponding to the linear-log model presented in equations 2a and 2b. In columns 3 to 6, we use different sets of instruments and controls for enrollment. In column 7 we use LIML instead of 2SLS as an estimation method. All results are directionally identical. Raising class size has a statistically-significant negative impact only on cognitive development, although all coefficients are negative for the other outcomes. More specifically, raising class size by one percent decreases cognitive development by 0.0019 to 0.0024 of a standard deviation. In the range we are studying, an increase of one student generates an increase in class size of 10 to 5 percent (from 10 to 20 students per class). Therefore, increasing class size from 10 to 11 students would likely imply a predicted decrease in test score of 0.019 to 0.024 SD, while an increase from 19 to 20 students would

¹³Test results are available from authors upon request.

only result in a decrease of 0.008 to 0.012 SD. To facilitate interpretation of the results, Figure 7 presents the marginal effect of adding one additional student per class by size of class.¹⁴ We see that the marginal impact of class size is negative but decreasing in magnitude for all four measures. In other words, increasing class size has a negative impact on the different scores in small classrooms, but less so in larger classrooms. The marginal effects are however only statistically different from zero for the effect on cognitive development.

Nonlinear effects can also be captured by a quadratic or a step function. Table 4 present all three methods simultaneously. Panel A (at the top) presents our benchmark linear-log model. Panel B presents results from a quadratic function of class size instrumented following equation 3a to 3c. Finally, panel C presents results from various step functions specified in equations 4a and 4b. Each cell (line and column) is a separate regression. In Table 4, the outcome is specified at the top of each column, and the instrumented variable on each row. All regressions include school- and student-level covariates, including a quadratic function of enrollment. Results using a quadratic function of class size are in line with those using the log, but the impact of class size on social competence here appears statistically significant. Coefficients of the step function can be interpreted as the effect of being in a larger class ($c_s > c$) in kindergarten compared to a smaller one, where the average class size above and below c are shown in the last two columns of the table. When the cut-off c is equal to 12 or 13, the impact of class size is negative and significant for three of our four measures, namely cognitive development, social competence and communication skills. However, once the cut-off reaches 15, only the impact on cognitive development holds. At $c = 15$ we are comparing students in classes of 13.5 students with students in classes of 18.5 students. The estimated impact of increasing class size beyond the 15 cut-off therefore measures the impact of raising class size by 5 students, or by 37 percent. The step-function leads to an estimated impact of 0.11 SD while the linear-log model would suggest a comparable effect of 0.10 SD (0.00259×37). Taken together, our findings strongly suggest that class size does not have a linear impact on cognitive development. Results on behavioral outcomes are not as clear,

¹⁴The blue diamonds show the overall marginal effects corresponding to columns 5 of Table 3. The other two series correspond to the heterogeneous effect by quintile of material deprivation, which will be discussed in the following subsection.

but suggest nonlinearities in some cases. Measuring the average impact using a linear model thus possibly hides important variations along the class size distribution.

5.4 Heterogeneity of effects

To assess heterogeneity, we investigate whether the estimated nonlinear impacts of class size differ between certain subgroups of students. The top panel of Table 5 shows the estimated effects by subgroup. The bottom panel of the table shows the average scores for the different subgroups, to give context to the estimated effects.

First we look at gender differences. Columns 1 and 2 of Table 5 show the estimated impacts for boys and for girls separately (top panel). The results suggest that for both genders, the effect of being in a larger class is negative and statistically significant for cognitive development. The estimates for behavioral measures are also similar across gender, but not statistically significant. The bottom panel shows that boys score lower on all four measures relative to girls, but unfortunately do not seem to benefit more from class size reduction.

We now turn to the poverty level of the school. To estimate the level of poverty of a school we use the percentage of students having attended pre-kindergarten. As mentioned in the section on institutional background, pre-kindergarten in Québec was only available in neighborhoods where the poverty level was considered high. The QSCDK does not explicitly state whether pre-kindergarten is available at the school in question, but the teacher reports whether a child has attended pre-kindergarten or not. We therefore separate schools in two groups: (1) schools with less than 50% of children having attended pre-kindergarten, and (2) schools with 50% or more of their students reported as having attended pre-kindergarten. Columns 3 and 4 of Table 5 present estimates for these two subgroups. Around 14% of our students are in schools with a preponderance of children having attended pre-kindergarten. Results suggest that class size effects are generally stronger in disadvantaged areas (pre-kindergarten schools) for three of our four measures of development: cognitive development, social competence, and communication skills. Children in these schools also have lower scores on all four measures, and the difference with children in non-pre-kindergarten schools

is always statistically different from zero.

The QSCDK also provides the child’s quintile of material deprivation based on the location of the child’s residence. The area of residence is defined at a very fine geographical level: we have the 6-digit postal code of the child’s residence, and only 19 households on average live in one postal code in Canada. We can therefore look at students living in high-poverty areas separately (top quintile of material deprivation¹⁵) to further validate our previous findings using the preponderance of pre-kindergarten. Estimates for children living in the poorest areas (q5 of material deprivation) are presented in column 6 of Table 5, while estimates for all other children are presented in column 5 (Table 5, q1-q4). Figure 7 also shows the marginal effects corresponding to these results graphically. Again, we find important differences between children from poorer neighborhoods compared to all others. Children in high-poverty areas benefit more from class size reduction (top panel) and they also have significantly lower test scores (bottom panel) compared to all other children. This result is encouraging. It implies that class size reduction policies targeting disadvantaged areas in kindergarten could have a chance of leveling the playing field, by helping those who need it most. We also estimated the impact separately for boys and girls by material deprivation quintile.¹⁶ We find that boys in high-poverty areas benefit about two times more than girls when we look at the impact on cognitive development.

5.5 Robustness analyses

In this subsection, we assess the robustness of our previous results by performing various analyses. First, we re-estimate the IV linear-log model while restricting the number of students enrolled in the school to assess the sensitivity of our findings to the presence of very large schools or very small schools (columns 2 to 4, Table 6). Second, the literature on class size using class size rules or threshold dummies as instruments highlights two other threats to instrument validity: bunching and score manipulation. We first validate the robustness

¹⁵The first quintile (q1) refers to children living in Census dissemination areas at the bottom 20% of material deprivation (low-poverty areas), while the fifth quintile (q5) refers to children living in dissemination areas at the top 20% of material deprivation (high-poverty areas).

¹⁶Estimates are available from authors upon request.

of our results to the exclusion of regions where bunching could not be excluded (column 5, Table 6), and then to the exclusion of regions that exhibit patterns most closely related to score manipulation (column 6, Table 6). Our previous results are robust to both.

Restricting total school enrollment to be under 105 or over 15 does not appear to alter our main results (columns 2 and 4, Table 6). The number of very large schools (above 105) or very small schools (below 15) is rather small. Restricting our sample to schools with a total enrollment below 85 is more restrictive: we lose about 17 percent of our observations. The magnitude and sign of the coefficient remain the same, but it is no longer statistically significant.

Regarding bunching, McCrary (2008) first introduced the idea that in the absence of manipulation, the density of the observed units should be continuous around the threshold or that alternatively, bunching on either side of the threshold should not be observed. In the present application, one might worry that the number of schools is not evenly distributed around the different enrollment thresholds, and that parental and student characteristics differ around the thresholds (Urquiola and Verhoogen, 2009; Fredriksson et al., 2013, 2016). Fredriksson et al. (2013, 2016) show evidence of bunching. They find that the fitted value of class size predicts parental education, such that parental education is not distributed uniformly around the thresholds. In 1962, Sweden implemented a compulsory school law, which induced schools to change their catchment area in favor of those most in need. To address this issue, the authors use a one-school district approach to study the impact of class size on long-term outcomes.

During our observation period, no reform impacting class size or primary school catchment areas was enacted. Class size does however exhibit some form of bunching since local school districts were not delimited randomly in the past, and were by design constructed to reach a certain fraction of the population. School catchment areas are rarely modified, at least as long as new schools are not added. To the best of our knowledge, no new school opened in the province in 2012. Furthermore, the public education law prevents school principals from refusing school access to a child living in the catchment area of the school board, so long as the school has not reached full capacity. In this sense, while some form of bunch-

ing should be expected if school catchment areas are well defined to start with, strategic bunching due to principals' behavior should not be observed.

Another form of bunching could however emerge because of parental choices. In Québec, each child is assigned to a school through the school's catchment area. However, parents may choose a different school within the catchment area of the entire school board they live in. Parents are however encouraged to choose their local school because their child could be displaced at the start of any school year if the chosen school reaches its capacity. In this case, priority is given to students within the school catchment area. In practice, it is therefore possible that schools of higher perceived quality are always closer to full capacity, and therefore have larger class sizes. Parents who opt out of their local school are likely to be parents who are more involved in the schooling of their child. This implies that we could underestimate the true impact of class size reduction.

We formally investigate potential bunching in our data. We start by testing for bunching around the enrollment thresholds using various bandwidths. Table 7 presents the test results, which were conducted on the full sample, and on restricted samples. When we follow Fredriksson et al. (2016) and use a bandwidth of five on the full sample, we find no evidence of bunching (p -value of 0.183). However, when we use alternative bandwidths (optimal bandwidth and values of six or seven), the result no longer holds, and the p -values are under 0.05. We then tested for bunching at the region and school board levels, and our investigation revealed that we could not rule out bunching in two regions, and more specifically five school boards ¹⁷. When these two regions (or five school boards within these regions) are excluded, the tests (reported in the second and third panels of Table 7) no longer support the presence of bunching around the enrollment thresholds, except when a bandwidth of seven is used. Despite the evidence we just presented that shows that bunching is unlikely to be a substantive issue, we estimate our IV model on the subsample of our data that excludes students registered in the five school boards where bunching may have occurred (column 5, Table 6). We find that our main results continue to hold even when these school boards are

¹⁷Confidentiality regulations regarding data usage prevent us from naming these regions or investigating the reasons that might have led to stronger evidence of bunching.

excluded.

Another concern may be that teachers and school resources below and above the enrollment thresholds are different, so that we are in fact measuring not only the impact of class size, but the impact of a bundle of changes related to class size. Our results should be interpreted with this caveat in mind, but we do provide some checks that help alleviate this concern. For example, there is the possibility that teacher quality may be different around the thresholds, with more inexperienced teachers being sent to schools forced to open a new classroom. In our data, we have information on professionals in the school, but we do not know teacher quality or experience. However, we can observe how each teacher answers the 104 questions on child development, and the pattern of answers may reveal inexperience or laziness. We check for that in our next robustness test. Also, from an administrative standpoint, the order in which teachers choose their school is based on their experience. Many factors enter a teacher’s choice of school, but the number of students in a class in a given year is not likely to be at the top of their list, since they know they will likely teach at that school for a number of years. Finally, other resources, such as the number of psychoeducators, are unlikely to change around the thresholds since they depend on total enrollment, but we check for this possibility below.

We first start by looking at score manipulation, which can occur because teachers are deliberately cheating, but also because they are inexperienced or are simply shirking. In order to account for the possibility of score manipulation by teachers, we implement a procedure based on Quintano et al. (2009), which reweights the observations based on a fuzzy K -means clustering approach.¹⁸ This procedure identifies clusters most likely to have faced manipulation (outliers) and provides new weights for each observation. More weight is given to scores that are less likely to have been manipulated. If teachers in smaller classes (or larger classes) are more likely to manipulate scores, this procedure would provide a reasonable strategy to correct for that. In our application, we exploit the pattern of answers by teacher for each of the 104 sub-questions. For example, a teacher answering “yes” to every single question for all students in their class would be given much less weight in this

¹⁸ Angrist et al. (2017) use this procedure and provide evidence of score manipulation in Italy.

procedure. This lack of variance in a teacher’s answers could be caused by, for example, laziness, inexperience or unprofessionalism. More details on the procedure are provided in the Appendix. The last column of Table 6 shows our estimated coefficients when observations are reweighted to account for the possibility of score manipulation. The impact of class size on all four measures of development is very similar to that of the baseline specification. Our results are therefore not driven by observations of the outlier cluster. This suggests that, to the extent in which teachers’ behavior is reflected in the pattern of answers, our results are not driven by different teachers being assigned to smaller or larger classes.

Furthermore, when we look at school and student characteristics around the enrollment thresholds for these students, we do not find major differences. In Table 8, column 1 shows the estimated coefficients of a regression of cognitive development on the main covariates in our data. Results show that these covariates are highly correlated with cognitive development (this is also true for the other measures, but is not presented here for brevity). Columns 2 to 6 present the results of individual estimations of each of the covariates on tr_1 to tr_5 , meaning each cell (line and column) refers to a separate regression. The coefficients on threshold dummies are all virtually equal to zero. This suggests that the covariates are not related to the instruments (tr_1 to tr_5). We also test that all coefficients on threshold dummies are jointly equal to zero. The q -values associated with these F -tests are reported in column 7, and are above 0.1 for all covariates except one (whether the child receives help from a social worker, q -value = 0.08). It appears that more children above the second threshold (tr_2 equals one if enrollment is above 40) receive help from social workers. Schools with larger enrollment are more likely to have access to different professionals. Receiving help from a social worker is negatively correlated with cognitive development (column 1). Since children are more likely to be in larger schools with larger classes, our class size estimate would be overestimated if we did not control for these differences, which we do. Overall, we conclude that our observed covariates are not correlated with the thresholds, which suggests that student and school characteristics are distributed evenly across the thresholds.

Finally, following Hoxby (2000) and Angrist and Lavy (1999), we restrict our sample to schools with an enrollment that is within five students or two students of the enrollment

thresholds. Students above the thresholds are in smaller classes, while students below the thresholds are in larger classes. We estimate a simple discontinuity model where the variable *Treat* identifies students above the threshold. Again we include all student- and school-level controls. We find in Table 9 that class size is indeed smaller above the threshold: the coefficient on *Treat* is around -4.0 to -5.6 depending on the specification. When we focus on the first two columns, we find that the impact of class size on cognitive development is robust to the $+/- 5$ restriction, but not to the $+/- 2$ restriction. The coefficients are similar, but become not statistically different from zero when only students from schools with enrollment within two students of the thresholds are kept. However, when we focus on our main group of interest—students living in high material deprivation areas—we find that class size reduction continues to benefit them even when we restrict our attention to schools within two students of the enrollment threshold. These results suggest that reducing class size by around 5.6 students is associated with an increase of 0.41 SD in cognitive development. This effect is two times larger than the one estimated using the linear-log model. Taken together, our robustness analyses show that our main findings hold under several conditions. The impact of class size reduction on the cognitive development of children living in disadvantaged areas appears very robust.

6 Conclusion

Our results bring novel insights to the impact of class size on child development. Using previously unexploited census-based survey data, we confirm the presence of nonlinearities in the effect of class size on measures of cognitive and noncognitive development. Estimates from linear models produce average effects, and may mask substantial nonlinearities. Class size reductions to 15 students or below appear to benefit kindergarten children, especially in terms of cognitive development. Above 15 students, class size variation does not appear to influence child development, whether cognitive or noncognitive. The benefits are mainly concentrated among students from disadvantaged backgrounds (especially disadvantaged boys), who represent less than 15% of our population. They benefit the most, and are also

those that stand to benefit the most, given that their developmental scores are all lower than those of pupils that are not disadvantaged. The impact on the EDI cognitive component is not trivial since it has been showed to be a strong determinant of school achievement over and above the cognitive assessments and direct school readiness test such as the PPVT and the Number Knowledge Test (Forget-Dubois et al., 2007).

Together, our results suggest that increasing class size within the 16-to-22 range in advantaged areas would free up resources to reduce class size drastically in disadvantaged areas, consistent with the policy advice in Piketty and Valdenaire (2006). This cost-neutral promising policy could benefit disadvantaged students without harming other students. Since the number of teachers would be kept constant, the quality of teaching would also remain constant. Hence, the potentially adverse general-equilibrium effects of a universal class-size reduction policy could be avoided. If redistribution across schools is contentious, then perhaps the same result could be achieved by reducing class size more drastically in the earlier grades, and crowding the later grades within a school, given that early interventions tend to have larger impacts (Heckman, 2006).

In Québec, class size in kindergarten was established at 22 until 1999. It was reduced to 20 in 2000 and then to 19 in 2016. In our data, we simulated the impact of increasing the maximum class size rule from 20 to 22 while maintaining enrollment per school constant. We find that average class size increases from 17.1 to 17.9 and that 203 fewer teachers and classes are needed. If applied to all seven grades of primary school, a fairly small increase could free up a large number of resources to reinvest in children with greater needs.

In terms of magnitude, the linear effect we estimate on cognitive development using IV is roughly half what Krueger (1999) had found for the STAR experiment, when we scale up our estimate to reflect a similar change. Our nonlinear estimates coming from the step function also yield a number that is roughly half of Krueger’s (1999), and somewhat in line with the literature using class size regulations as instruments. Limits to our analysis come from the contemporaneous nature of the data: we look at the effect of kindergarten class size on kindergarten outcomes. While early outcomes matter and are crucial to the long-term development of a child, our study does not have the breadth of long-term studies such as

Chetty et al. (2011) or Fredriksson et al. (2013). A follow-up survey, or linkages using administrative files, would clearly help shed light on long-term effects. Another limit is that we only observe teacher-reported outcomes. While all teachers in Québec are professionals with a 4-year bachelor degree in education and teacher reports, after observing the child for at least six months, are likely to offer good evaluations of children relative to their classmates, the degree of standardization across teachers may not be as good as it would be if we were to use standardized testing. On the positive side however, concerns about teaching-to-the-test or score manipulation are probably more limited in our context. Moreover, as stated earlier, the relevance and reliability of the Early Development Instrument, which is the basis of the development measures used in our analysis, has been proven in the literature on child development. Finally, teachers were provided time outside of the classroom during their working hours to ensure they answer the questionnaire carefully and they were also guaranteed that their report would never be used to compute statistics at the school level (or even school board level), which certainly helped reduce the likelihood that teachers manipulate scores to improve the ranking of their school.

In conclusion, we highlight the nonlinear nature of the relationship between class size and child development, both cognitive and noncognitive. We study the impact within the class size bounds that we observe, that is mainly between 10 to 22 students per class, and show that allowing for nonlinearities of the class size effect is important, and could lead to different policy implications than estimates coming from a typical linear model.

7 References

Angrist, J. D., Battistin, E., & Vuri, D. (2017). In a Small Moment: Class Size and Moral Hazard in the Mezzogiorno. *American Economic Journal: Applied Economics*, 9(4), 216–249.

Angrist, J. D., & Lavy, V. (1999). Using Maimonides’ Rule to Estimate the Effect of Class Size on Scholastic Achievement. *The Quarterly Journal of Economics*, 114(2), 533–575

Bandiera, O., Larcinese, V., & Rasul, I. (2010). Heterogeneous Class Size Effects: New

Evidence from a Panel of University Students. *The Economic Journal*, 120(549), 1365–1398.

Bedard, K., & Kuhn, P. (2008). Where Class Size Really Matters: Class Size and Student Ratings of Instructor Effectiveness. *Economics of Education Review*, 27(3), 253–265.

Bitler, M. P., Gelbach, J. B., & Hoynes, H. W. (2006). What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments. *The American Economic Review*, 96(4), 988–1012.

Boozer, M., & Rouse, C. (2001). Intrascchool Variation in Class Size: Patterns and Implications. *Journal of Urban Economics*, 50(1), 163–189.

Brinkman, S., Gregory, T., Harris, J., Hart, B., Blackmore, S., & Janus, M. (2013). Associations between the Early Development Instrument at Age 5, and Reading and Numeracy Skills at Ages 8, 10 and 12: A Prospective Linked Data Study. *Child Indicators Research*, 6(4), 695–708.

Cattaneo, M. D., Jansson, M., & Ma, X. (2020). Simple Local Polynomial Density Estimators. *Journal of the American Statistical Association*, 115(531), 1449–1455.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR. *Quarterly Journal of Economics*, 126(4), 1593–1660.

Chingos, M. M. (2012). The Impact of a Universal Class-Size Reduction Policy: Evidence from Florida’s Statewide Mandate. *Economics of Education Review*, 31(2012), 543–562.

Cho, H., Glewwe, P., & Whitler, M. (2012). Do Reductions in Class Size Raise Students’ Test Scores? Evidence from Population Variation in Minnesota’s Elementary Schools. *Economics of Education Review*, 31(3), 77–95.

Dee, T. S., & West, M. R. (2011). The Non-Cognitive Returns to Class Size. *Educational Evaluation and Policy Analysis*, 33(1), 23–46.

Ding, W., & Lehrer, S. F. (2011). Experimental Estimates of the Impacts of Class Size on Test Scores: Robustness and Heterogeneity. *Education Economics*, 19(3), 229–252.

Dunn, L. M., & Dunn, L. M. (1997). Peabody Picture Vocabulary Test (3rd ed.). Circle Pines, MN: American Guidance Service.

Feir, D., Lemieux, T., & Marmer, V. (2016). Weak Identification in Fuzzy Regression

Discontinuity Designs, *Journal of Business & Economic Statistics*, 34(2), 185–196

Finn, J. D., & Achilles, C. M. (1999). Tennessee’s Class Size Study: Findings, Implications, Misconceptions. *Educational Evaluation and Policy Analysis*, 21(2), 97–109.

Forget-Dubois, N., Lemelin, J.-P., Boivin, M., Dionne, G., Séguin, J. R., Vitaro, F., & Tremblay, R. E. (2007). Predicting early school achievement with the EDI: A longitudinal population-based study. *Early Education and Development*, 18(3), 405–426. Available at <https://doi.org/10.1080/10409280701610796>

Fredriksson, P., Öckert, B., & Oosterbeek, H. (2013). Long-Term Effects of Class Size. *The Quarterly Journal of Economics*, 128(1), 249–285.

Fredriksson, P., Öckert, B., & Oosterbeek, H. (2016). Parental Responses to Public Investments in Children: Evidence from a Maximum Class Size Rule, *Journal of Human Resources*, 51(4), 832–868.

Gilraine, M., Macartney, H., & McMillan, R. (2018). Education Reform in General Equilibrium: Evidence from California’s Class Size Reduction (No. w24191). National Bureau of Economic Research.

Glass, G. V., & Smith, M. L. (1979). Meta-Analysis of Research on Class Size and Achievement. *Educational Evaluation and Policy Analysis*, 1(1), 2–16.

Government of Québec. (2012). Mieux accueillir et éduquer les enfants d’âge préscolaire, une triple question d’accès, de qualité et de continuité des services. Conseil supérieur de l’éducation. ISBN : 978-2-550-65006-5

Hanushek, E. A. (1999). The Evidence on Class Size. In Mayer, S. E., & Peterson, P. E. (Eds.), *Earning & Learning : How Schools Matter* (pp. 131–168). Washington, D.C., Washington, D.C.: Brookings Institution Press.

Hanushek, E. A. (2002). Evidence, Politics, and the Class Size Debate. In Mishel, L. & Rothstein, R. (Eds.), *The Class Size Debate* (pp. 37–66). Washington, D.C., Washington, D.C.: Economic Policy Institute.

Heckman, J. J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science*, 312(5782), 1900–1902.

Hojo, M. (2013). Class-Size Effects in Japanese Schools: A Spline Regression Approach.

Economics Letters, 120(3), 583–587.

Hoxby, C. M. (2000). The Effects of Class Size on Student Achievement: New Evidence from Population Variation. *Quarterly Journal of Economics*, 115(4), 1239–1285.

Institut de la statistique du Québec. (2013). Enquête québécoise sur le développement des enfants à la maternelle 2012, Portrait statistique pour le Québec et ses régions administratives, ISBN 978-2-550-68877-8

Jackson, E., & Page, M. E. (2013). Estimating the Distributional Effects of Education Reforms: A Look at Project STAR. *Economics of Education Review*, 32(2013), 92–103.

Janus, M., Brinkman, S. A., & Duku, E. K. (2011). Validity and Psychometric Properties of the Early Development Instrument in Canada, Australia, United States, and Jamaica. *Social Indicators Research*, 103(2), 283–297.

Janus, M., & Offord, D. R. (2007). Development and Psychometric Properties of the Early Development Instrument (EDI): A Measure of Children’s School Readiness. *Canadian Journal of Behavioural Science*, 39(1), 1–22.

Krueger, A. B. (1999). Experimental Estimates of Education Production Functions. *The Quarterly Journal of Economics*, 114(2), 497–532.

Krueger, A. B. (2003). Economic Considerations and Class Size. *The Economic Journal*, 113(485), F34–F63.

Krueger, A. B., & Whitmore, D. M. (2001). The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR. *The Economic Journal*, 111(468), 1–28.

Laurin, I., Fournier, M., Bigras, N., & Solis, A. (2015). La fréquentation d’un service éducatif préscolaire: un facteur de protection pour le développement des enfants de familles à faible revenu? *Canadian Journal of Public Health*, 106(7), eS14–eS20.

Laurin, I., Guay, D., Fournier, M., Blanchard, D., & Bigras, N. (2018). Quelle est l’association entre les caractéristiques résidentielles et du quartier et le développement de l’enfant à la maternelle? *Canadian Journal of Public Health*, 109(1), 35–42.

Lazear, E. P. (2001). Educational Production. *Quarterly Journal of Economics*, 116(3), 777–803.

- Lazear, E. P. (2006). Speeding, Terrorism, and Teaching to the Test. *Quarterly Journal of Economics*, 121(3), 1029–1061.
- Lochner, L., & Moretti, E. (2015). Estimating and Testing Models with Many Treatment Levels and Limited Instruments. *Review of Economics and Statistics*, 97(2), 387–397.
- McCrary, J. (2008). Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics*, 142(2), 698–714.
- Ministère de l'Éducation, de l'Enseignement supérieur et de la Recherche (MEESR) (2015), *Indicateurs de l'Éducation, Éducation préscolaire, enseignement primaire et secondaire, Édition 2014*. Gouvernement du Québec. Available at http://www.education.gouv.qc.ca/fileadmin/site_web/documents/PSG/statistiques_info_decisionnelle/indicateurs.2014_fr.pdf
- Mueller, S. (2013). Teacher Experience and the Class Size Effect – Experimental Evidence. *Journal of Public Economics*, 98(2013), 44–52.
- OECD (2017). Education at a Glance 2017: OECD Indicators, Éditions OCDE, Paris. Available at <http://dx.doi.org/10.1787/eag-2017-en>
- Okamoto, Y., & Case, R. (1996). Exploring the microstructure of children's central conceptual structures in the domain of number. *Monographs of the Society for Research in Child Development*, 61(1/2), 27–58.
- Piketty, T., & Valdenaire, M. (2006). L'impact de la taille des classes sur la réussite scolaire dans les écoles, collèges et lycées français. Estimations à partir du panel primaire 1997 et du panel secondaire 1995, Paris: Ministère de l'éducation nationale, 2006, 153 p.
- Quintano, C., Castellano, R., & Longobardi, S. (2009). A Fuzzy Clustering Approach to Improve the Accuracy of Italian Student Data. An Experimental Procedure to Correct the Impact of the Outliers on Assessment Test Scores. *Statistica & Applicazioni*, VII(2), 149–171.
- Rockoff, J., & Turner, L. J. (2010). Short-run Impacts of Accountability on School Quality. *American Economic Journal: Economic Policy*, 2(4), 119–147.
- Simard, M., Tremblay, M.-E., Lavoie, A., & Audet, N. (2013). *Enquête québécoise sur le développement des enfants à la maternelle 2012*, Québec, Institut de la Statistique du Québec, 99 p.

Urquiola, M., & Verhoogen, E. (2009). Class-Size Caps, Sorting, and the Regression-Discontinuity Design. *The American Economic Review*, 99(1), 179–215.

Urquiola, M. (2006). Identifying Class Size Effects in Developing Countries: Evidence from Rural Bolivia. *Review of Economics and Statistics*, 88(1), 171–177

Wechsler, D. (1989). *Manual for the Wechsler Preschool and Primary Scale of Intelligence-Revised*. San Antonio, TX: Psychological Corporation.

Zelazo, P. D., Jacques, S., Burack, J. A., & Frye, D. (2002). The relation between theory of mind and rule use: Evidence from persons with autism-spectrum disorders. *Infant & Child Development*, 11, 171–195.

8 Figures

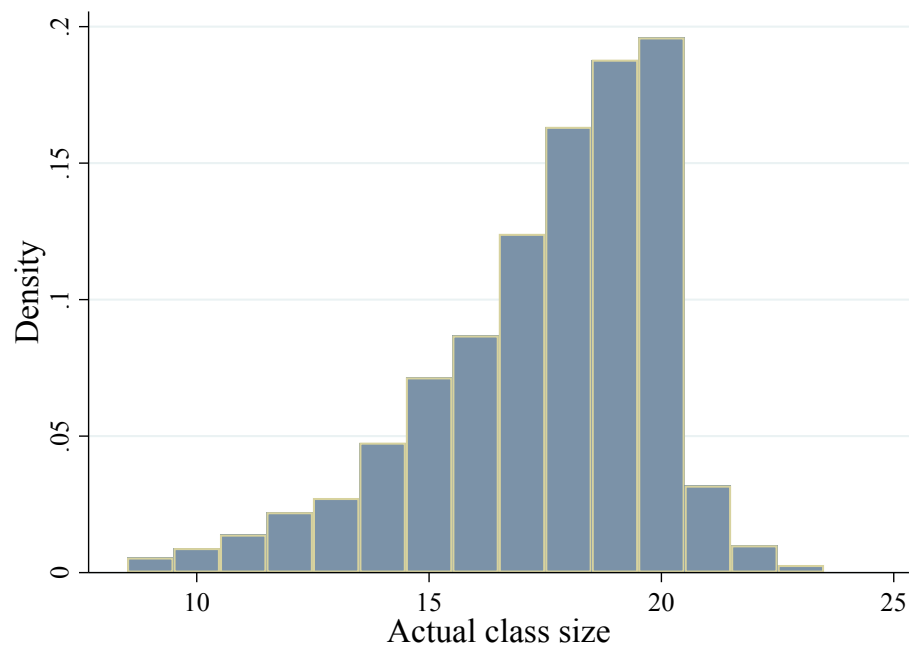


Figure 1: Histogram of Actual Class Size

Note: This figure shows the class size distribution in our data set, with class as a unit of observation.
Source: Authors' calculations using QSCDK 2012 data.

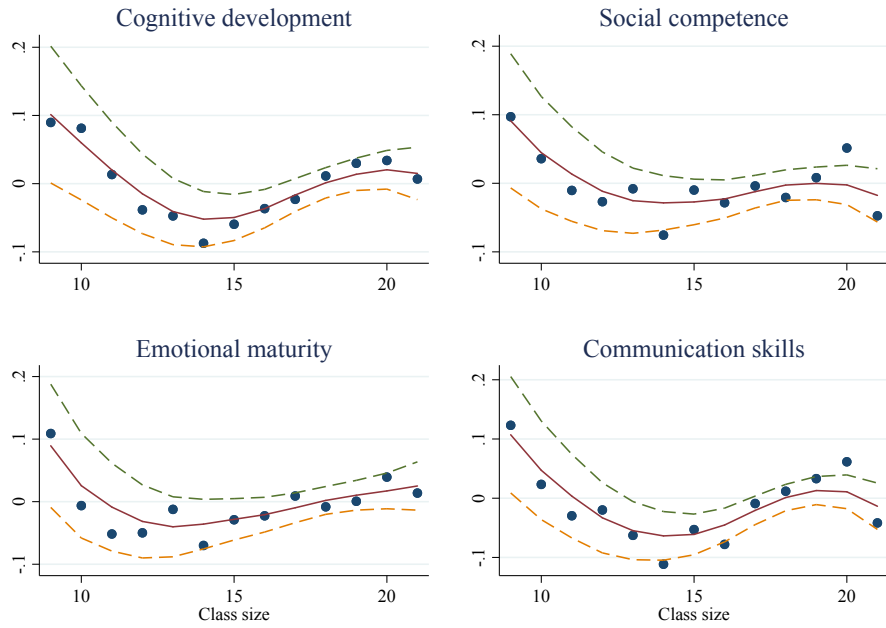


Figure 2: Average Developmental Score by Class Size

Note: In this figure, each dot corresponds to a class size (X -axis), and the Y -axis shows the average standardized developmental score for that class size. The solid line is a smoothed fit through the dots, while the dashed lines show 95% confidence intervals.

Source: Authors' calculations using QSCDK 2012 data.

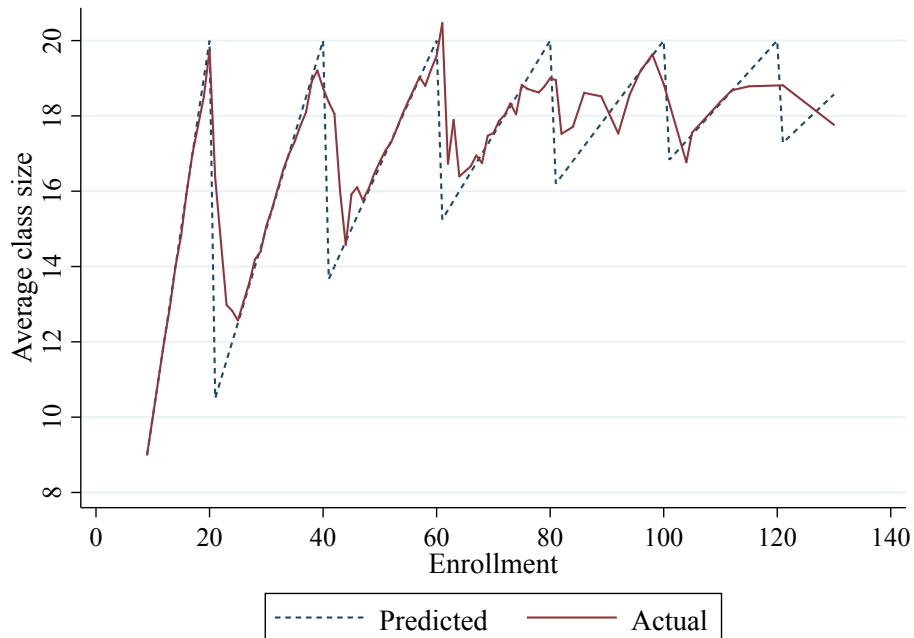


Figure 3: Average Predicted and Actual Class Size by School Enrollment

Note: This figure shows the actual (solid line) and predicted (dashed line) class size given enrollment.

Source: Authors' calculations using QSCDK 2012 data.

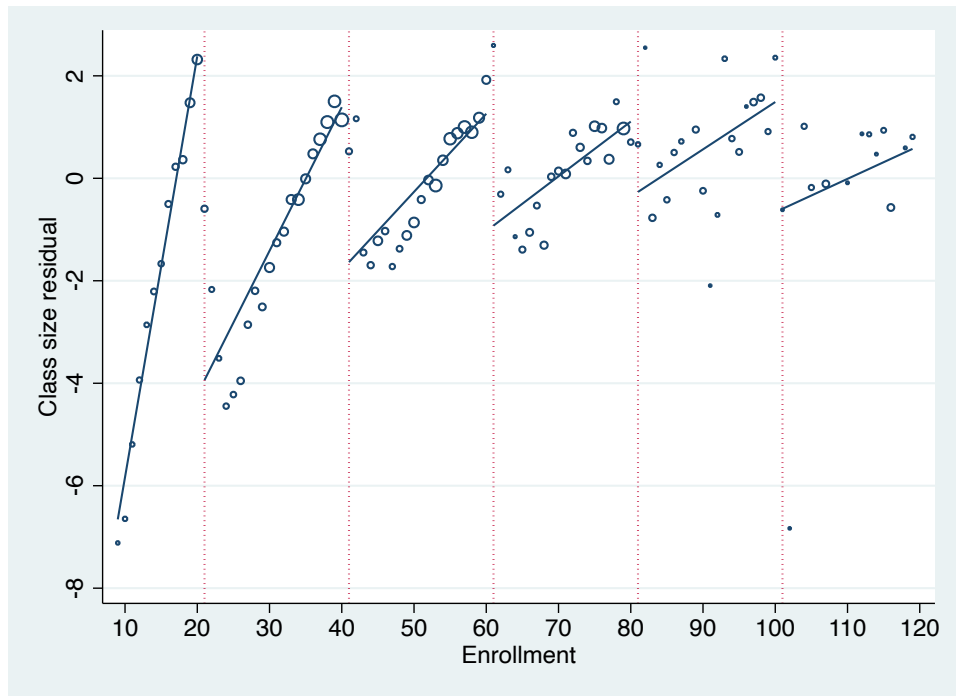


Figure 4: Class Size Residuals by Enrollment

Note: This figure shows residual class size, after controlling for school board fixed effects. The regression lines were fitted to individual data. The size of the dots is proportional to the number of observations.
Source: Authors' calculations using QSCDK 2012 data.

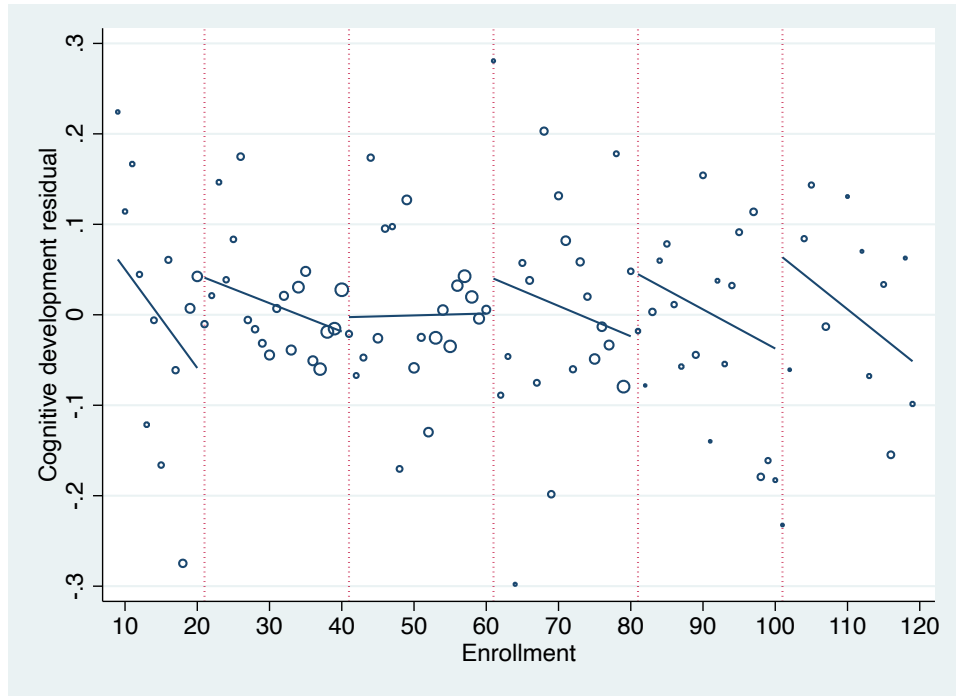


Figure 5: Cognitive Development Residuals by Enrollment

Note: This figure shows regression lines fitted to individual data of residual class size on enrollment, after controlling for student- and school-level covariates (see list in Table 2's note). The size of the dots is proportional to the number of observations.

Source: Authors' calculations using QSCDK 2012 data.

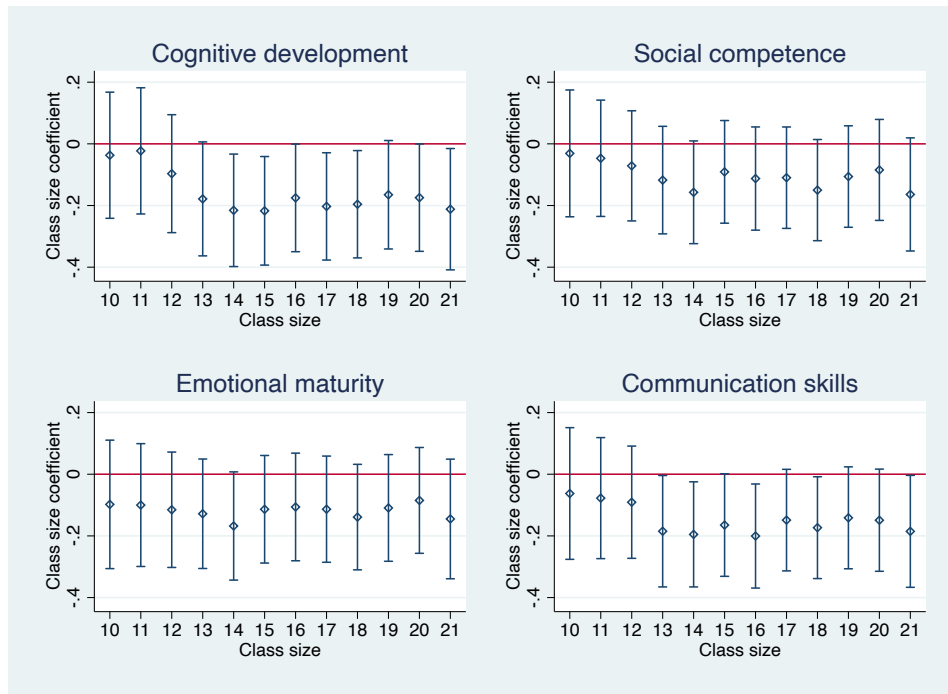


Figure 6: Nonlinearity: OLS Coefficients on Class Size Dummies

Note: This figure shows the estimated coefficients of class size dummies on developmental scores, including student- and school-level controls. Class size of nine is the omitted category. Bars show 95% confidence intervals, computed using standard errors clustered at the school level.

Source: Authors' calculations using QSCDK 2012 data.

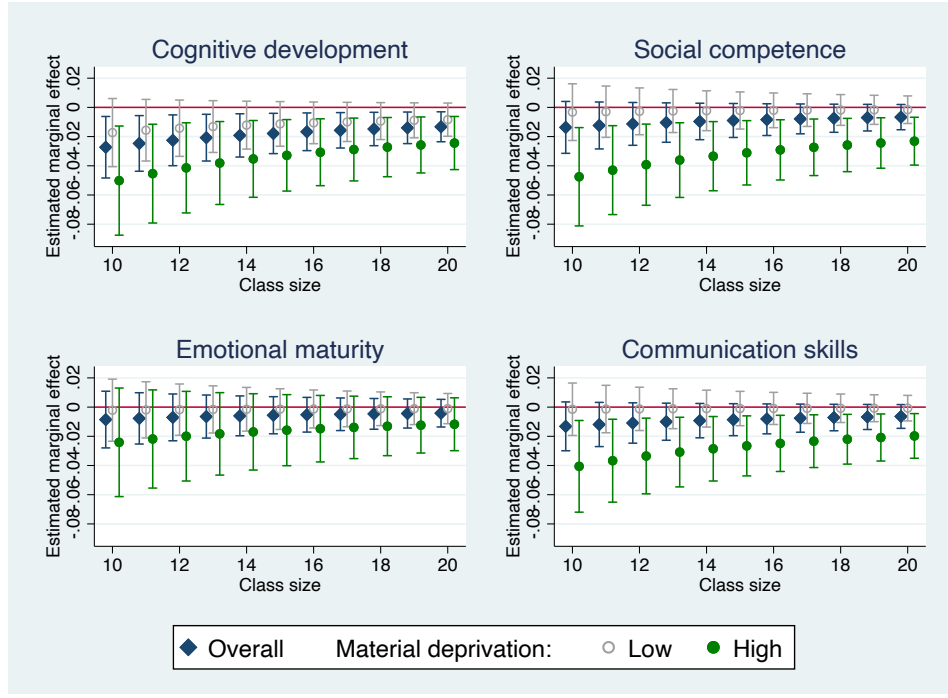


Figure 7: Estimated Marginal Effects by Class Size

Note: This figure shows estimated marginal effects by class size based on the logarithmic model estimated using IV with the size rule and its square (Table 3, column 5 for the overall series, Table 5, columns 5 and 6 for the series by quintile of material deprivation). Bars show 95% confidence intervals, computed using standard errors clustered at the school level.

Source: Authors' calculations using QSCDK 2012 data.

9 Tables

Table 1: DESCRIPTIVE STATISTICS - STUDENT

	Mean
Age (in months)	72.07
Std. dev.	(3.62)
Female	0.50
Mother tongue	
French	0.77
English	0.07
Other	0.16
Dysfunctional family	0.030
Disabilities	
Physical	0.005
Dental	0.004
Visual	0.003
Hearing	0.004
Chronic disease	0.004
Speech	0.057
Learning	0.065
Emotional	0.031
Behavioral	0.063
Receiving help from a	
Nurse	0.03
Speech therapist	0.08
Psychoeducator	0.04
Social worker	0.02
Psychologist	0.03
Attended childcare	
Yes	0.62
No	0.15
Missing	0.23
Attended pre-kindergarten	
Yes	0.17
No	0.80
Missing	0.03

Note: $N = 58,949$

Source: Authors' calculations using QSCDK 2012 data.

Table 2: LINEAR CLASS SIZE EFFECTS

	OLS estimates			IV estimates		N
	(1)	(2)	(3)	(4)	(5)	
Cognitive development						
<i>Class size</i>	0.008*	-0.001	-0.005	-0.012*	-0.012	58,777
	(0.004)	(0.004)	(0.004)	(0.007)	(0.009)	
<i>R</i> ²	0.00	0.14	0.20	0.20	0.20	
Social competence						
<i>Class size</i>	0.005	-0.001	-0.002	-0.004	-0.004	58,949
	(0.003)	(0.003)	(0.003)	(0.006)	(0.007)	
<i>R</i> ²	0.00	0.18	0.26	0.26	0.26	
Emotional maturity						
<i>Class size</i>	0.007**	0.002	0.001	-0.003	0.002	58,558
	(0.004)	(0.004)	(0.004)	(0.007)	(0.008)	
<i>R</i> ²	0.00	0.14	0.23	0.23	0.23	
Communication skills						
<i>Class size</i>	0.015***	0.008**	0.001	-0.006	-0.002	58,937
	(0.004)	(0.004)	(0.004)	(0.006)	(0.007)	
<i>R</i> ²	0.00	0.23	0.31	0.31	0.31	
First-stage <i>F</i> -statistic				512.5	172.4	
School controls	O	X	X	X	X	
Student controls	O	O	X	X	X	
<i>Z</i> : n_s	O	O	O	X	X	
<i>Z</i> : tr_1 to tr_5	O	O	O	O	X	

Note: Each cell contains the estimate from a separate regression. School controls include enrollment, enrollment square, poverty index (high or low), social and material deprivation indices (highly advantaged, average, highly disadvantaged), teaching language (French or English), and school board dummies. Student controls include dummy variables indicating gender, student's age in months, place of birth, whether the child attended childcare or pre-kindergarten, ten markers of health and behavioral problems (physical disability, visual deficiency, auditive deficiency, speech disorder, learning difficulties, emotional problems, behavioral problems, disadvantaged family environment, chronic health conditions, and dental problems), and dummies to indicate whether the child received help from various school professionals (nurse, speech therapist, psychoeducator, social worker, and psychologist). In specification (5), enrollment is controlled for using segments and allowing the slope for enrollment to vary within each threshold. Cluster robust standard errors are reported in parentheses. Significance is denoted using asterisks: *** is $p < 0.01$, ** is $p < 0.05$, and * is $p < 0.1$.

Source: Authors' calculations using QSCDK 2012 data.

Table 3: NONLINEAR CLASS SIZE EFFECTS

	Reweighted for uniform class size distribution			Nonlinear effect using $\ln(\text{class size})$				
	OLS	2SLS		2SLS			LIML	
	(1)	(2)		(3)	(4)	(5)	(6)	(7)
Cognitive development ($N = 58,777$)								
<i>Class size</i>	-0.012*** (0.004)	-0.055*** (0.017)	$\ln(c)$	-0.192* (0.109)	-0.217** (0.106)	-0.259** (0.102)	-0.240* (0.132)	-0.240* (0.132)
R^2	0.197	0.179	R^2	0.203	0.203	0.203	0.204	0.204
Social competence ($N = 58,949$)								
<i>Class size</i>	-0.001 (0.005)	-0.016 (0.015)	$\ln(c)$	-0.067 (0.091)	-0.094 (0.089)	-0.130 (0.086)	-0.120 (0.109)	-0.120 (0.109)
R^2	0.267	0.265	R^2	0.257	0.257	0.257	0.257	0.257
Emotional maturity ($N = 58,558$)								
<i>Class size</i>	0.000 (0.005)	-0.009 (0.016)	$\ln(c)$	-0.041 (0.101)	-0.055 (0.098)	-0.081 (0.094)	-0.015 (0.124)	-0.015 (0.124)
R^2	0.230	0.229	R^2	0.227	0.227	0.226	0.227	0.227
Communication skills ($N = 58,937$)								
<i>Class size</i>	-0.002 (0.006)	-0.018 (0.016)	$\ln(c)$	-0.086 (0.084)	-0.100 (0.083)	-0.125 (0.081)	-0.055 (0.102)	-0.055 (0.103)
R^2	0.300	0.298	R^2	0.313	0.313	0.313	0.314	0.314
First-stage F -statistic		21.9		549.5	532.1	117.6	129.2	129.2
Instruments (Z) and alternative controls for enrollment								
$Z: n_s$	O	X		X	O	O	O	O
$Z: \ln(n_s)$	O	O		O	X	X	X	X
$Z: tr_1 \text{ to } tr_5$	O	O		O	O	X	X	X
Enrollment thresholds and segments	O	O		O	O	O	X	X

Note: Each cell contains estimates from a separate regression. In the first two columns, observations are reweighted so that the class size distribution is uniform. Column 1 is estimated by OLS, all other columns rely on IV estimation (equations 2a and 2b), where the size rule and its log are used as instruments, along with enrollment thresholds as indicated. All specifications include school- and student-level controls. Enrollment is controlled for using enrollment and its square, unless specified otherwise. Cluster robust standard errors are reported in parentheses. Significance is denoted using asterisks: *** is $p < 0.01$, ** is $p < 0.05$, and * is $p < 0.1$.

Source: Authors' calculations using QSCDK 2012 data.

Table 4: NONLINEAR CLASS SIZE EFFECTS - QUADRATIC AND STEP FUNCTIONS

	Cognitive development (1)	Social competence (2)	Emotional maturity (3)	Communication skills (4)		
Panel A: Linear-log benchmark						
$\ln(c)$	-0.259** (0.102)	-0.130 (0.086)	-0.081 (0.094)	-0.125 (0.081)		
Panel B: Quadratic function						
Class size	-0.117** (0.053)	-0.115** (0.046)	-0.059 (0.051)	-0.066 (0.044)		
Class size ²	0.003* (0.002)	0.003** (0.001)	0.002 (0.002)	0.002 (0.001)		
Panel C: Step functions					Mean class size	
Each line and column is a separate regression					$c_s \leq c$	$c_s > c$
$I[c_s > 12]$	-0.242*** (0.08)	-0.132* (0.069)	-0.076 (0.076)	-0.117* (0.065)	11.0	17.9
$I[c_s > 13]$	-0.188*** (0.064)	-0.100* (0.055)	-0.057 (0.060)	-0.090* (0.052)	11.7	18.0
$I[c_s > 14]$	-0.147*** (0.053)	-0.075* (0.045)	-0.043 (0.049)	-0.069 (0.042)	12.6	18.3
$I[c_s > 15]$	-0.112** (0.047)	-0.053 (0.039)	-0.033 (0.043)	-0.052 (0.037)	13.5	18.5
$I[c_s > 16]$	-0.096** (0.044)	-0.044 (0.037)	-0.029 (0.04)	-0.046 (0.034)	14.2	18.9
$I[c_s > 17]$	-0.080* (0.043)	-0.033 (0.036)	-0.023 (0.039)	-0.038 (0.033)	15.1	19.2
$I[c_s > 18]$	-0.071 (0.05)	-0.026 (0.041)	-0.022 (0.045)	-0.036 (0.039)	15.9	19.7
N	58,777	58,949	58,558	58,937		

Note: Panel A corresponds to the linear-log baseline model in column 5 of Table 3. In panel B, each column is a separate estimation using IV where both class size and its square are instrumented (equations 3a to 3c). Instruments include rule size, rule size square and the thresholds (tr_1 to tr_5). In panel C, each cell stems from a separate estimation using IV of equations 4a and 4b, and shows the estimated coefficient on $\ln(\text{class size})$ or the class size indicator $I[c_s > c]$, for values of c going from 12 to 18. Instruments include $\ln(\text{rule size})$ and the thresholds (tr_1 to tr_5). All specifications include school- and student-level controls. School enrollment is controlled for using enrollment and enrollment square. Cluster robust standard errors are reported in parentheses. Significance is denoted using asterisks: *** is $p < 0.01$, ** is $p < 0.05$, and * is $p < 0.1$. Source: Authors' calculations using QSCDK 2012 data.

Table 5: HETEROGENEITY OF EFFECTS

Panel A	Gender		School with pre-kindergarten		Material deprivation	
	Boys (1)	Girls (2)	No (3)	Yes (4)	Low: q1-q4 (5)	High: q5 (6)
Cognitive development						
$\ln(c)$	-0.293** (0.119)	-0.222** (0.106)	-0.154 (0.121)	-0.642*** (0.233)	-0.164 (0.113)	-0.476*** (0.181)
N	29,635	29,142	51,048	7,729	47,082	10,356
R^2	0.191	0.197	0.205	0.227	0.197	0.220
Social competence						
$\ln(c)$	-0.146 (0.109)	-0.110 (0.083)	-0.042 (0.101)	-0.491** (0.201)	-0.031 (0.094)	-0.451*** (0.163)
N	29,711	29,238	51,207	7,742	47,223	10,386
R^2	0.237	0.194	0.260	0.259	0.250	0.283
Emotional maturity						
$\ln(c)$	-0.075 (0.116)	-0.085 (0.097)	-0.031 (0.109)	-0.345 (0.218)	-0.020 (0.103)	-0.229 (0.18)
N	29,484	29,074	50,867	7,691	46,905	10,324
R^2	0.182	0.135	0.229	0.235	0.221	0.25
Communication skills						
$\ln(c)$	-0.143 (0.096)	-0.101 (0.085)	-0.061 (0.093)	-0.568*** (0.205)	-0.014 (0.087)	-0.385** (0.152)
N	29,707	29,230	51,196	7,741	47,214	10,383
R^2	0.304	0.300	0.317	0.309	0.309	0.328
Average scores						
Panel B	Boys	Girls	No pre-K	Pre-K	q1-q4	q5
Cognitive development	-.11	.12	.01	-.06	.03	-.14
Social competence	-.22	.22	.00	-.03	.02	-.09
Emotional maturity	-.28	.28	.01	-.04	.02	-.08
Communication skills	-.13	.13	.01	-.05	.03	-.14

Note: Panel A shows the estimated coefficient on $\ln(\text{class size})$ when $\ln(\text{class size})$ is instrumented using $\ln(\text{size rule})$ and tr_1 to tr_5 (equations 2a and 2b). Student- and school-level controls included. School enrollment controlled for using enrollment and enrollment square. Panel B shows the average scores by subgroup. All score differences across subgroups are statistically significant with p -values smaller than 0.01. Cluster robust standard errors are reported in parentheses. Significance is denoted using asterisks: *** is $p < 0.01$, ** is $p < 0.05$, and * is $p < 0.1$.

Source: Authors' calculations using QSCDK 2012 data.

Table 6: ROBUSTNESS CHECKS

	Linear-log baseline (1)	Enrollment < 105 (2)	Enrollment < 85 (3)	Enrollment > 15 (4)	Bunching boards excl. (5)	Rewighted score manipulation (6)
Cognitive development						
$\ln(c)$	-0.259** (0.102)	-0.230** (0.111)	-0.189 (0.116)	-0.211* (0.128)	-0.209** (0.106)	-0.227** (0.099)
N	58,777	54,109	49,441	57,075	53,760	58,777
R^2	0.203	0.203	0.202	0.202	0.202	0.199
Social competence						
$\ln(c)$	-0.130 (0.086)	-0.095 (0.092)	-0.093 (0.096)	-0.078 (0.106)	-0.114 (0.091)	-0.140 (0.096)
N	58,949	54,257	49,573	57,246	53,923	58,949
R^2	0.257	0.257	0.257	0.256	0.252	0.254
Emotional maturity						
$\ln(c)$	-0.081 (0.094)	-0.028 (0.103)	-0.03 (0.107)	-0.059 (0.117)	-0.109 (0.098)	-0.085 (0.088)
N	58,558	53,905	49,255	56,868	53,553	58,558
R^2	0.226	0.227	0.228	0.226	0.221	0.234
Communication skills						
$\ln(c)$	-0.125 (0.081)	-0.074 (0.088)	-0.088 (0.091)	-0.096 (0.098)	-0.119 (0.086)	-0.136* (0.072)
N	58,937	54,245	49,561	57,234	53,911	58,937
R^2	0.313	0.314	0.315	0.313	0.308	0.306

Note: Each coefficient results from a separate estimation of the IV model from equations 2a and 2b, where $\ln(\text{size rule})$ and tr_1 to tr_5 are used as instruments. Column 1 corresponds to the baseline linear-log model in column 5 of Table 3. Columns 2 to 4 restrict observations to schools whose total enrollment is under 105 and 85 or above 15, respectively. Column 5 excludes observations in the five school boards where bunching could not be ruled out. Column 6 use weights that account for the possibility of score manipulation (see text and Appendix for details). All specifications include school- and student-level controls. Cluster robust standard errors are reported in parentheses. Significance is denoted using asterisks: *** is $p < 0.01$, ** is $p < 0.05$, and * is $p < 0.1$.

Source: Authors' calculations using QSCDK 2012 data.

Table 7: DENSITY TESTS AROUND ENROLLMENT THRESHOLDS

	Bandwidth		Effective N		p -value/ t -stat	Excluding 2 regions	Excluding 5 school boards
	Left	Right	Left	Right			
RD density, p -value (Cattaneo et al., 2020)	5.20	4.33	567	188	0.044	no	no
	5	5	567	233	0.183	no	no
	6	6	664	274	0.031	no	no
	7	7	757	308	0.002	no	no
	5.75	4.25	397	153	0.382	yes	no
	5	5	397	186	0.721	yes	no
	6	6	465	220	0.314	yes	no
	7	7	542	246	0.091	yes	no
	5.01	4.24	494	177	0.215	no	yes
	5	5	494	219	0.277	no	yes
	6	6	582	256	0.096	no	yes
	7	7	673	289	0.017	no	yes
DC Density, t -stat (McCrary, 2008)					-0.138	no	no
					0.842	yes	no
					0.993	no	yes

Note: This table shows various test statistics testing for bunching around enrollment thresholds. The top panel shows p -values associated with Cattaneo et al.'s (2020) test, using the optimal bandwidth (first row), and bandwidths of five, six, and seven. Effective N refers to the number of schools within the bandwidths around the thresholds. The bottom panel shows the t -statistic from McCrary's (2008) test. Two regions or five school boards showing evidence of bunching are excluded sequentially from the sample, as indicated in the last two columns.

Source: Authors' calculations using QSCDK 2012 data.

Table 8: CHILDREN AND SCHOOL CHARACTERISTICS AROUND THRESHOLDS

	<i>This column is one regression</i>	<i>Each row contains the estimates from a separate regression</i>					
	Cognitive development (1)	tr_1 (2)	tr_2 (3)	tr_3 (4)	tr_4 (5)	tr_5 (6)	q -value of F -test (7)
Female student	0.150*** (0.009)	-0.014 0.970	-0.032 1.024	0.232 0.811	-0.262 0.958	0.046 1.042	0.683
Born in Canada	0.047** (0.023)	0.018 1.000	0.017 1.000	-0.364 0.965	0.084 1.000	-0.009 0.993	0.523
Dysfunctional family	-0.866*** (0.041)	0.008 0.985	0.007 1.000	-0.092 0.865	-0.005 1.000	0.098 0.954	0.635
Disabilities							
<i>Physical</i>	-0.755*** (0.104)	-0.003 1.000	-0.024 0.706	0.012 0.969	0.023 0.949	-0.009 1.000	0.456
<i>Visual</i>	-0.381*** (0.115)	0.002 1.045	-0.014 0.973	-0.011 0.969	-0.031 1.020	0.063 0.536	0.380
<i>Hearing</i>	-0.681*** (0.092)	0.004 0.425	-0.005 1.000	-0.005 1.000	-0.004 1.000	0.007 1.000	0.725
<i>Chronic disease</i>	-0.371*** (0.110)	0.000 1.000	-0.008 0.976	-0.001 1.000	-0.023 0.997	0.038 0.874	0.817
<i>Dental issues</i>	-0.636*** (0.126)	0.002 0.991	0.004 1.000	0.032 0.905	-0.043 0.962	0.010 1.000	0.707
Help from a school professional							
<i>Nurse</i>	-0.180*** (0.042)	-0.079 0.058	0.015 1.000	0.246 0.836	0.014 1.000	-0.016 1.000	0.147
<i>Speech therapist</i>	0.564*** (0.028)	-0.059 0.052	0.124 0.806	-0.065 1.000	0.178 0.988	-0.049 1.000	0.198
<i>Psychoeducator</i>	0.067* (0.035)	-0.022 0.698	0.092 0.938	0.049 1.000	-0.112 0.975	-0.012 1.000	0.668
<i>Social worker</i>	-0.146*** (0.048)	-0.031 0.120	0.169 0.063	-0.074 0.959	-0.116 0.988	0.138 0.979	0.082
<i>Psychologist</i>	0.290*** (0.041)	-0.026 0.091	0.061 0.828	0.005 1.000	0.085 0.950	-0.101 0.952	0.404
<i>Other professional</i>	0.355*** (0.021)	-0.06 0.653	0.084 0.999	-0.292 0.956	0.216 1.000	0.203 1.000	0.699
Attended pre-kindergarten	0.054*** (0.019)	0.138 0.062	-0.306 0.958	-0.013 0.989	0.277 0.994	-0.608 0.996	0.137
School poverty index (High)	-0.086*** (0.029)	-0.050 1.000	-0.176 1.000	-0.777 0.954	1.555 0.950	0.183 1.000	0.640
School language (French)	-0.171*** (0.043)	-0.059 0.842	-0.658 0.661	0.681 0.820	-0.656 1.000	1.025 0.962	0.454
N	53,760						

Note: Excludes school boards in which bunching cannot be rejected ($N = 53,760$). Column 1 shows the coefficients estimated using OLS for a model where the cognitive development score is explained by all the covariates (dummy variables) listed in the rows. Columns 2 to 6 report results from separate regressions (one per row) estimated by OLS, where the variable listed on the left is explained by the five thresholds (tr_1 to tr_5). Column 7 shows the q -value associated with the F -statistic testing for the joint significance of the coefficients on the four threshold dummies. Cluster robust standard errors are reported in parentheses for specification (1) and q -values are reported under each tr_1 to tr_5 coefficient. Significance is denoted using asterisks: *** is $p < 0.01$, ** is $p < 0.05$, and * is $p < 0.1$.

Source: Authors' calculations using QSCDK 2012 data.

Table 9: ROBUSTNESS CHECKS AROUND THRESHOLDS

	Number of students from discontinuity					
	All students		High deprivation (q5)		Low deprivation (q1-q4)	
	+/- 5 (1)	+/- 2 (2)	+/- 5 (3)	+/- 2 (4)	+/- 5 (5)	+/- 2 (6)
Class size						
<i>Treat</i>	-4.050*** (0.131)	-5.616*** (0.344)	-4.110*** (0.198)	-5.630*** (0.406)	-4.071*** (0.139)	-5.646*** (0.351)
<i>N</i>	26,208	9,275	3,894	1,141	21,722	7,922
<i>R</i> ²	0.674	0.837	0.681	0.898	0.671	0.825
Cognitive development						
<i>Treat</i>	0.065* (0.035)	0.075 (0.076)	0.130* (0.068)	0.411** (0.205)	0.047 (0.037)	0.037 (0.080)
<i>N</i>	26,144	9,258	3,885	1,141	21,667	7,905
<i>R</i> ²	0.206	0.217	0.243	0.263	0.198	0.208
Social competence						
<i>Treat</i>	0.017 (0.028)	0.076 (0.063)	0.071 (0.051)	0.118 (0.161)	0.008 (0.028)	0.056 (0.067)
<i>N</i>	26,208	9,275	3,894	1,141	21,722	7,922
<i>R</i> ²	0.258	0.267	0.304	0.317	0.251	0.258
Emotional maturity						
<i>Treat</i>	0.075** (0.031)	0.116* (0.07)	0.057 (0.048)	0.311* (0.168)	0.082** (0.032)	0.090 (0.069)
<i>N</i>	26,007	9,214	3,876	1,136	21,544	7,870
<i>R</i> ²	0.231	0.230	0.290	0.298	0.223	0.218
Communication skills						
<i>Treat</i>	0.044* (0.026)	0.105** (0.053)	0.163*** (0.060)	0.208 (0.143)	0.009 (0.024)	0.070 (0.050)
<i>N</i>	26,205	9,274	3,894	1,141	21,719	7,921
<i>R</i> ²	0.319	0.319	0.327	0.299	0.317	0.324

Note: Only includes students in schools with enrollment within x students of the discontinuity ($+/- x$). Each cell is a separate regression where the variable *Treat* equals one if enrollment is above the threshold (21 to 25, 41 to 45, 61 to 65, etc.) and equals zero otherwise. All specifications include school- and student-level controls. Standard errors clustered at the school level are reported in parentheses. Significance is denoted using asterisks: *** is $p < 0.01$, ** is $p < 0.05$, and * is $p < 0.1$.

Source: Authors' calculations using QSCDK 2012 data.

Appendix (not for publication) to Nonlinear Class Size Effects on Cognitive and Noncognitive Development of Young Children

Marie Connolly and Catherine Haeck

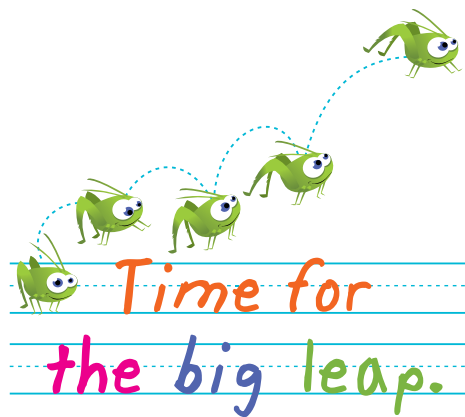
May 2019

1 Score manipulation

In order to account for the possibility of score manipulation (cheating) by the teachers, we implement a procedure based on Quintano et al. (2009), which re-weights the observations based on a fuzzy K -means clustering approach. This procedure reassigns more weight to observations that are less likely to have been manipulated. This method is implemented in four steps. First, for each class and each developmental score, we compute the mean and standard deviation, the total number of missing answers to each sub-question (non-response), and the average Gini (homogeneity of answers across all sub-questions). Second, we use a principal component analysis to reduce the number of dimensions to assess score manipulation. The principal component analysis results demonstrate that the first two components explain at least 79% of the observed variance. The first component is correlated with class mean score, standard deviation, and homogeneity of answers. This component is labeled the outlier identifier. The second component is correlated with the number of missing answers and is labeled the class participation identifier. Third, we use a fuzzy K -means cluster analysis to regroup classes into different clusters and ultimately identify the cluster in which score manipulation is more probable (thereafter referred to as the “outlier cluster”). Such a cluster would be defined as one having high negative values on the first component (indicating high average scores and low variability of scores within class) and values close to zero on the second component (indicating a small number of missing answers). Contrary to a regular K -means cluster analysis, the fuzzy K -means analysis does not assign an observation to a unique cluster. Instead, each observation is assigned a probability of

belonging to a specific cluster based on the Euclidean distance to its centroid. As a final step, we exploit this particularity by re-weighting each observation i by w_i , the complement to the probability μ of belonging to the outlier cluster a : $w_i = 1 - \mu_{ia}$.

2 QSCDK 2012 Questionnaire



Québec Survey of Child Development
in Kindergarten, 2012.

Early development instrument[©]

Québec 2011-2012

Instructions

For any additional information, contact the Direction des stratégies et opérations de collecte of Statistique Québec by calling 418-691-2404, or 1-800-561-0213 toll free.

Please complete and mail back the questionnaire in the prepaid envelope.

Fill in the circles like this ● or like this (X) not like this x

Québec



Avec la participation de :

- Ministère de la Santé et des Services sociaux
- Ministère de l'Éducation, du Loisir et du Sport
- Ministère de la Famille et des Aînés



AVENIR D'ENFANTS
DES COMMUNAUTÉS ENGAGÉES

Offord
CENTRE
FOR CHILD
STUDIES



EARLY DEVELOPMENT INSTRUMENT
a population-based measure for communities

If any of the information on the label is incorrect or missing, please make changes clearly below.

1. Class assignment:

- ☒ Kindergarten

2. Child's Date of Birth:

day		month		year			

3. Sex:

- ☐ F ☐ M

4. Postal Code:

--	--	--	--	--	--	--	--

5. Class Type: (see Guide)

- ☐ Kindergarten
☐ Preschool/Kindergarten
☐ Preschool/Kindergarten/Grade 1
☐ Kindergarten/Grade 1
☐ Other

6. Date of Completion:

day		month		2 0 1 2			

7. Identified Special Needs: (see Guide)

- ☐ Yes ☐ No (Go to question 9)

8. Student at risk or EHDA: (see Guide)

- ☐ At Risk ☐ EHDA

9. Child considered ESL:

- ☐ Yes ☐ No

- ☐ The parent refuses to let the teacher fill out the child's questionnaire.

Return the uncompleted questionnaire with the refusal coupon and the child's label.

10. French Immersion:

- ☐ Yes ☐ No

11. Other Immersion

- ☐ Yes ☐ No

12. Aboriginal (North American Indian, Metis or Inuit):

- ☐ Yes ☐ No ☐ Don't know

13. Child's first language(s):

- ☐ English only
☐ French only
☐ Other only (see Guide)

--	--	--	--

☐ English & French
☐ English & other

--	--	--	--

☐ French & other

--	--	--	--

☐

--	--	--	--

 other and

--	--	--	--

 other

(Refer to Guide for language codes in "other" categories. If you do not know the "other" language code, use "998").

14. Communicates adequately in his/her first language:

- ☐ Yes ☐ No ☐ Don't know

15. Student Status: (see Guide)

- ☐ in class more than 1 month
☐ in class less than 1 month
☐ moved out of this class
☐ moved out of the school
- } **Do not complete**

16. Student is repeating this grade:

- ☐ Yes ☐ No

SECTION A – PHYSICAL WELL-BEING

1. About how many regular days (see Guide) has this child been absent since the beginning of this school year?

,
Number of days absent

Since the beginning of this school year, has this child sometimes (more than once) arrived:

	Yes ↓	No ↓	Don't know ↓
2. over- or under-dressed for school-related activities	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. too tired/sick to do school work	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. late	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. hungry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Would you say that this child:

	Yes ↓	No ↓	Don't know ↓
6. is independent in washroom habits most of the time	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. shows an established hand preference (right vs. left or vice-versa)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. is well-coordinated (i.e., moves without running into or tripping over things)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How would you rate this child's:

	Very good/ Good ↓	Average ↓	Poor/ Very poor ↓	Don't know ↓
9. proficiency at holding a pen, crayons, or a brush	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. ability to manipulate objects	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. ability to climb stairs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. level of energy throughout the school day	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13. overall physical development	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

SECTION B – LANGUAGE AND COGNITIVE SKILLS

How would you rate this child's:	Very good/ Good ↓	Average ↓	Poor/ Very poor ↓	Don't know ↓
1. ability to use language effectively in English	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. ability to listen in English	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. ability to tell a story	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. ability to take part in imaginative play	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. ability to communicate own needs in a way understandable to adults and peers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. ability to understand on first try what is being said to him/her	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. ability to articulate clearly, without sound substitutions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Would you say that this child:	Yes ↓	No ↓	Don't know ↓
8. knows how to handle a book (e.g., turn pages)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. is generally interested in books (pictures and print)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. is interested in reading (inquisitive/curious about the meaning of printed material)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. is able to identify at least 10 letters of the alphabet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. is able to attach sounds to letters	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13. is showing awareness of rhyming words	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14. is able to participate in group reading activities	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15. is able to read simple words	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16. is able to read complex words	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17. is able to read simple sentences	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18. is experimenting with writing tools	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
19. is aware of writing directions in English (left to right, top to bottom)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
20. is interested in writing voluntarily (and not only under the teacher's direction)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
21. is able to write his/her own name in English	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
22. is able to write simple words	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

SECTION B – LANGUAGE AND COGNITIVE SKILLS (continued)

Would you say that this child:	Yes ↓	No ↓	Don't know ↓
23. is able to write simple sentences	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
24. is able to remember things easily	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
25. is interested in mathematics	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
26. is interested in games involving numbers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
27. is able to sort and classify objects by a common characteristic (e.g., shape, colour, size)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
28. is able to use one-to-one correspondence	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
29. is able to count to 20	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
30. is able to recognize numbers 1-10	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
31. is able to say which number is bigger of the two	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
32. is able to recognize geometric shapes (e.g., triangle, circle, square)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
33. understands simple time concepts (e.g., today, summer, bedtime)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
34. demonstrates special numeracy skills or talents	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
35. demonstrates special literacy skills or talents	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
36. demonstrates special skills or talents in arts	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
37. demonstrates special skills or talents in music	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
38. demonstrates special skills or talents in athletics/dance	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
39. demonstrates special skill or talents in problem-solving in a creative way	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
40. demonstrates special skills or talents in other areas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
If yes, please specify _____			

SECTION C – SOCIAL AND EMOTIONAL DEVELOPMENT

How would you rate this child's:

	Very good/ Good ↓	Average ↓	Poor/ Very poor ↓	Don't know ↓
1. overall social/emotional development	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. ability to get along with peers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Below is a list of statements that describe some of the feelings and behaviours of children. For each statement, please fill in the circle that best describes the child now as much as possible. It might be necessary to use the period of time since the beginning of the school year as your reference frame.

Would you say that this child:

	Often or Very true ↓	Sometimes or Somewhat true ↓	Never or Not true ↓	Don't know ↓
3. plays and works cooperatively with other children at the level appropriate for his/her age	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. is able to play with various children	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. follows rules and instructions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. respects the property of others	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. demonstrates self-control	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. shows self-confidence	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. demonstrates respect for adults	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. demonstrates respect for other children	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. accepts responsibility for actions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. listens attentively	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13. follows directions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14. completes work on time	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15. works independently	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16. takes care of school materials	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17. works neatly and carefully	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18. is curious about the world	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
19. is eager to play with a new toy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
20. is eager to play a new game	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
21. is eager to play with/read a new book	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

SECTION C – SOCIAL AND EMOTIONAL DEVELOPMENT (continued)

Would you say that this child:	Often or Very true ↓	Sometimes or Somewhat true ↓	Never or Not true ↓	Don't know ↓
22. is able to solve day-to-day problems by him/herself	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
23. is able to follow one-step instructions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
24. is able to follow class routines without reminders	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
25. is able to adjust to changes in routines	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
26. answers questions showing knowledge about the world (e.g., leaves fall in the autumn, apple is a fruit, dogs bark)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
27. shows tolerance to someone who made a mistake (e.g., when a child gives a wrong answer to a question posed by the teacher)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
28. will try to help someone who has been hurt	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
29. volunteers to help clean up a mess someone else has made	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
30. if there is a quarrel or dispute will try to stop it	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
31. offers to help other children who have difficulty with a task	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
32. comforts a child who is crying or upset	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
33. spontaneously helps to pick up objects which another child has dropped (e.g., pencils, books)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
34. will invite bystanders to join in a game	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
35. helps other children who are feeling sick	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
36. is upset when left by parent/guardian	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
37. gets into physical fights	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
38. bullies or is mean to others	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
39. kicks, bites, hits other children or adults	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
40. takes things that do not belong to him/her	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
41. laughs at other children's discomfort	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
42. can't sit still, is restless	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43. is distractible, has trouble sticking to any activity	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
44. fidgets	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
45. is disobedient	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

SECTION C – SOCIAL AND EMOTIONAL DEVELOPMENT (continued)

Would you say that this child:	Often or Very true ↓	Sometimes or Somewhat true ↓	Never or Not true ↓	Don't know ↓
46. has temper tantrums	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
47. is impulsive, acts without thinking	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
48. has difficulty awaiting turn in games or groups	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
49. cannot settle to anything for more than a few moments	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
50. is inattentive	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
51. seems to be unhappy, sad or depressed	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
52. appears fearful or anxious	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
53. appears worried	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
54. cries a lot	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
55. is nervous, high-strung or tense	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
56. is incapable of making decisions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
57. is shy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
58. sucks a thumb/finger	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

SECTION D – SPECIAL CONCERNS

1. Does the student have a problem that influences his/her ability to do school work in a regular classroom? (based on parent information, medical diagnosis, and/or teacher observation).
- | Yes
↓ | No
↓ | Don't know
↓ |
|-----------------------|-----------------------|-----------------------|
| <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
- If answer is No/Don't know go to question 5a

2. If you answered "Yes" to the previous question, please mark all that apply. Base your answers on your observation as a teacher or on a medical diagnosis and/or parent/guardian information.

	Yes Observed ↓	Yes Parent info/ Medical diagnosis ↓		Yes Observed ↓	Yes Parent info/ Medical diagnosis ↓
a. physical disability	<input type="radio"/>	<input type="radio"/>	g. behavioural problem	<input type="radio"/>	<input type="radio"/>
b. visual impairment	<input type="radio"/>	<input type="radio"/>	h. home environment/ problems at home	<input type="radio"/>	<input type="radio"/>
c. hearing impairment	<input type="radio"/>	<input type="radio"/>	i. chronic medical/health problems	<input type="radio"/>	<input type="radio"/>
d. speech and language impairment	<input type="radio"/>	<input type="radio"/>	j. unaddressed dental needs	<input type="radio"/>	<input type="radio"/>
e. learning disability	<input type="radio"/>	<input type="radio"/>	k. other:	<input type="radio"/>	<input type="radio"/>
f. emotional problem	<input type="radio"/>	<input type="radio"/>	(specify) _____		

3. If the child has received a diagnosis or identification by a doctor or psychological professional please indicate. (Please use the Guide for diagnostic codes.)

Code:

If « Other » (code 33), specify: _____

	Yes ↓	No ↓	Don't know ↓
4. Is the child receiving any school-based support (e.g. ressource person, special equipment or other)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5 a. Is the child currently receiving further assessment?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5 b. Is the child currently on a waitlist to receive further assessment?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5 c. Do you feel that this child needs further assessment?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

SECTION E – COMPLEMENTARY INFORMATION

		Yes ↓	No ↓	Don't know ↓
The child:				
1.	Had been in non-parental child care on a regular basis prior to kindergarten entry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	Had attended a kindergarten class for 4-year-olds in a public school?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3 a.	Child's place of birth	Canada ↓ <input type="radio"/>	Outside Canada ↓ <input type="radio"/>	Don't know ↓ <input type="radio"/>
3 b.	If the child was born outside Canada, how long has he or she been living in Canada?	Less than 2 years ↓ <input type="radio"/>	2 years or more ↓ <input type="radio"/>	Don't know ↓ <input type="radio"/>
4.	Since the beginning of the school year, has the child received the services of a non-teaching professional at school?	Yes ↓	No ↓	Don't know ↓
a.	Nurse	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b.	Speech therapist	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c.	Psychoeducator (see Guide)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d.	Social worker	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e.	Psychologist	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f.	Other (specify) _____	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5.	Since the beginning of the school year, has one of the child's parents or guardian participated in a parent meeting (either in person or by telephone)?	Yes ↓ <input type="radio"/>	No ↓ <input type="radio"/>	Don't know ↓ <input type="radio"/>
6.	In your opinion, is one of the child's parents (or guardian) able to understand when you communicate with him/her (verbally or in writing)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

If you have any comments about this child and her/his readiness for school, list them below:

Thank you for your cooperation.

Go on the next child

Please put all the questionnaires, the group list, and parental refusal coupons
in the envelope and return it to the Institut de la statistique du Québec.

