The effect of information shocks in primary schools. Evidence from a quasi-experimental study.

Fritz Schiltz[δ] and Kristof De Witte[‡,δ]

δ: Katholieke Universiteit Leuven (KU Leuven), Leuven Economics of Education Research, Naamsestraat 69, 3000 Leuven (Belgium), kristof.dewitte@kuleuven.be; fritz.schiltz@kuleuven.be.

‡: Maastricht University, Top Institute for Evidence Based Education Research, Kapoenstraat 2, MD 6200 Maastricht (the Netherlands), k.dewitte@maastrichtuniversity.nl.

February 14, 2017

Abstract

School inspectorate reports provide transparency in the market for schools. It allows consumers (parents and pupils) to (re)consider school choice more objectively. To obtain causal evidence, we exploit the fact that school inspectorates select schools randomly – enforced by law - for a detailed assessment of school quality. Unfavorable reports induce reputation effects, which might trigger a lower demand for the school. This paper contributes to the growing literature on school information and parental choice by exploiting information shocks in primary schools in an educational context without standardized central examinations and with school inspection reports as the sole source of objective school quality information parents can rely on. Our findings indicate that parental demand for schools is strongly affected by shocks in transparency, following the publication of, especially positive, school inspection reports. These responses vary with the school's location and socio-economic composition.

Keywords: School choice; School inspectorate; Information shock; Quasi-experiment; socio-economic composition.

JEL-classification: I20, L10

1 Introduction

This paper answers the following two research questions by quasi-experimental evidence: '*Do parents respond to shocks in school transparency resulting from school inspections?*' And '*How are schools' compositions affected by information shocks?*'. Answering these questions provides new evidence on the mediating role of school inspections to improve quality through enhanced competition.

School choice has been proposed as a means to increase competition and improve school quality (MacLeod & Urquiola, 2015). In the absence of catchment areas, schools are competing for students and the funding assigned to them. Proponents argue that competition in the education market improves student achievement. This claim is based on the idea that markets supply goods efficiently, which is ensured by the concern of sellers over their reputation. In the education market, schools compete over students and free school choice encourages enrollment in high quality schools (Friedman, 1962; Macleod, 2007). Opponents indicate that possible side-effects might occur if competition is not properly put in place. For example, school choice could lead to more school segregation if schools are allowed to select students (e.g. Hsieh & Urquiola, 2006).

Even if schools cannot select students, further decentralizing the education landscape - and hence fostering competition – will only affect school performance if parents are informed about the relative quality of schools (Mizala & Urquiola, 2013). In other words, transparency on the added value (or quality) of schools is a necessary condition to reap the benefits of school choice and competition – in addition to the non-selectivity of schools. School inspections have been proposed in most western education systems as a major instrument for controlling and promoting the quality of schools (Ehren, 2016). Given the positive relationship between school accountability and academic outcomes (Schutz, West, & Wossmann, 2007), school inspections can be seen as a 'low stakes' accountability tool for policy makers (Burgess, 2016).

In recent years, school inspectorates have taken a center stage in the toolbox of policy makers. Two mechanisms can be identified which motivate the increasing importance of school inspectorates. On the one hand, they allow parents to make well informed choices about the school they send their children to. Inspection reports allow parents to objectively compare schools, and, hence, foster quality improvement through increased competition between schools. On the other hand, central government agencies are able to supervise autonomous schools in a decentralized school system, enabling them to guarantee certain quality standards. Intuitively, both mechanisms are strongly intertwined: The ability of central governments to ensure a certain quality threshold will likely depend on the competitive forces present in the education market, which in turn depends on the degree parents are able to discern schools based on their quality.

A broad (qualitative) literature studies the relationship between school inspections and school performance. In their review, de Wolf & Janssens (2007) conclude that the (by then) existing studies cannot identify a clear causal relationship between school inspections and school quality. They state that "*It is, in any event, typical that the studies make almost no use of research designs that focus more strongly on exposing causal relationships, such as the use of control groups and more quasi-experimental approaches.*" (p.388). Based on qualitative studies, they find that "*principals and teachers believe performance indicators are important*[1]*, while parents and pupils take very little notice of these indicators when choosing schools*" (p.379). Furthermore, the authors suggest that the type of transparency (i.e. positive or negative performance scores) might have different effects on school quality. Also, de Wolf & Janssens (2007) argue that a distinction between short- and long-term effects is hardly ever made in the available literature: "*the studies examining the effects of inspection on school achievement focus primarily on the year in which the inspection visit took place or the year immediately following. The question is whether the intended effects of inspection visits on pupils' educational achievements would appear so rapidly.*"(p.388). For example, changes in school policy are typical short-term effects (e.g. replacing the school principal) while the impact (e.g. test scores) of this policy change does not appear directly in (administrative) data.

Klerks (2012) updates this literature review and identifies fourteen relevant articles. Her findings suggest a lack of evidence that school inspections lead to the improvement of the educational quality, at least not directly. Narrowing down the available studies to those applying the most advanced empirical methodologies (e.g. using fixed effects models) reveals rather small causal effects of school inspections in the Netherlands (Luginbuhl, Webbink, & de Wolf, 2009), and even negative effects in the UK (Rosenthal, 2004).

This paper contributes to the literature in several ways. First, given that student test scores do not immediately change after an information shock provided by the education inspectorate

[1] Also more recent qualitative research supports this finding (Gaertner, Wurster, & Pant, 2013)

(Klerks, 2012), we focus on other outcomes that might be influenced in the short run. In particular, we focus on parental responsiveness to the publication of school inspection reports. Parental responsiveness will be assessed by looking at changes in enrollment growth and schools' socio-economic composition, reflecting parental school choice.

Looking at schools' socio-economic composition is the second contribution of this paper and the first study to include school SES as an *outcome* variable. Detailed and centralized data on schools' socio-economic composition (see section 4) allows us to study the impact of information shocks. This contribution complements the work of Hastings, Van Weelden, & Weinstein, 2007 stating that "by lowering information costs policy makers may enable families to act on their true preferences and fully benefit from school choice" (Hastings et al., 2007, p.27). If capacity constraints are reached in schools, a change in demand following a school inspection cannot further increase the school size but it offers school administrators leeway to leverage this excess demand. Koning & Van der Wiel (2013) studied the demand side of school choice by looking at parental responses to information shocks (see below) while our dataset and the unique setting of Flanders enables us to analyze the supply-side 'pull factors' that might affect the composition of schools. In Flanders, where urban schools reached their capacity constraint, a change in demand following a school inspection cannot further increase the school size but it offers school administrators leeway to leverage this excess demand.

Third, to identify the causal effect of an information shock by school inspectorates our identification strategy requires that schools are randomly selected for inspection. In the absence of a random inspection (e.g. targeting low performing schools), the results would be endogenous due to selection bias. We show in section 3 and 4 that school inspections are completely random and exogenous in our setting, such that we can rely on quasi-experimental identification strategies. In particular, we will introduce a Difference-in-Differences (DiD) design with fixed effects, tailored to the educational setting in Flanders for which we have a rich and unique administrative panel dataset, combined with data obtained from the school inspectorate.

Fourth, thanks to our specific setting, we can single out the effect of information shocks provided by school inspection reports from other school quality information. We focus in particular on Flanders, the Dutch speaking region of Belgium, which makes an unique setting for our purposes since there are no standardized central examinations. As a result, school inspection reports are the sole source of objective school quality information parents can rely on.

The randomness in school selection by inspectorates combined with the absence of other accountability tools results in a truly exogenous shock in school transparency. In addition, by assessing the impact of these shocks on short-term outcomes (enrolment and socio-economic composition of schools) we are able to overcome the problems faced in earlier work (section 2).

The remainder of the paper is structured as follows. Section 2 reviews the literature, while section 3 introduces the setting of school inspections in Flanders. Section 4 presents the data. Our empirical strategy is developed in section 5. Section 6 presents the results of our analysis and extends these by various robustness checks and sensitivity analyses. Section 7 discusses our results and concludes.

2 Literature review

This section reviews the state of the art of the literature on school information and education market outcomes[2]. Following the selection criteria of Klerks (2012), we focus our attention on studies applying the most advanced empirical methodologies, with a particular focus on school inspection data.

Rosenthal (2004) reports for the English education inspectorate, the Office of Standards in Education (Ofsted), a significant negative effect of school inspections, although small in magnitude and disappearing over time. Apart from problems related to the lagged impact on quality measures, it should be noted that the reputation of schools, and hence their composition will also influence average performance. Rosenthal applies a fixed effects "within estimator" to account for this possible source of bias and to single out the impact of school inspections. However, when the quality of English schools moves together with Ofsted inspections and this data is readily available to parents, it could be that the estimated effect of a school policy change – following inspection – is not captured in this setting. Parents could consider other school choices, affecting the school's composition. This mechanism has not been studied yet appears very likely considering the importance and availability of accountability measures in education (apart from inspection data) to parents in England[3]. As a result, it is not clear from

[2] We do not consider studies looking at other market outcomes, e.g. property prices (Figlio & Lucas, 2004).
[3] Burgess, Wilson, & Worth (2013) exploit a policy change (following a referendum on school policy) in school accountability in Wales but not in England to obtain a Difference-in-Difference estimate on the

this study whether the negative impact follows parental choices or changing school policies, as targeted by 'hard governance' interventions by Ofsted (Grek, Lawn, Ozga, & Segerholm, 2013).

Luginbuhl et al. (2009) follow a similar empirical approach and obtain a positive and significant effect of school inspections on test scores, especially for mathematics. In contrast to the Ofsted selection process, school are selected non-randomly in the Netherlands and hence it cannot be stated that "*the estimates from this model are free from the problem of endogeneity bias*" (Luginbuhl et al., 2009). Repeating their analysis on a more restrictive, but essentially randomly drawn, sample of schools, the authors cannot find any significant effect due to the limited number of inspections at randomly selected schools.

Mizala & Urquiola (2013) attempted to study the direct effect of information shocks by looking at enrollment growth in Chilean schools, following the publication of award-winning schools. Applying a sharp regression discontinuity (RDD) they find no consistent evidence that enrolment rates are affected by obtaining an award. Experimental evidence by Hastings & Weinstein (2008) revealed that the extent to which parents are aware of school quality depends strongly on their socio-economic status (SES). This might explain the limited effect as observed by Mizala & Urquiola (2013) since award-winning schools already consist of children of high SES parents, who do not adjust their behavior following this shock in transparency. This result is also in line with the aforementioned findings of Burgess, Wilson, & Worth (2013) where high performing schools in Wales were unaffected by changes in accountability pressures.

On the contrary, a study conducted in the Netherlands by Koning & Van der Wiel (2013) finds that negative (positive) school quality scores decrease (increase) the number of students choosing a school. This effect is measured after one year and follows from the publication of school rankings in a national newspaper. The positive effects are particularly large for the academic school track[4]. In addition, and contrasting the findings of Hastings & Weinstein (2008), the authors cannot find differences in information responses between socio-economic groups. The ranking of schools constructed by the newspaper were based on school inspection reports. These reports contain information on the average grade students achieve at the centralized exam in their final year of education; the percentage of students who from third grade on leave the school

importance of accountability in education. They find that accountability strongly affects school effectiveness (in terms of test scores), but not in schools belonging to the top quartile of the league tables.

[4] Note that this mechanism follows the finding of de Wolf & Janssens (2007) that the type of transparency (i.e. positive or of negative performance scores) might have different effects.

with a diploma without any delay; and the net percentage of students who in third grade are within school tracks that are above or below their school advice. Although the empirical strategy of this study is solid (it considers short-term effects and overcomes selection bias), it fails to identify the impact of a shock in information. In other words, it measures the effect of increased coverage by a national newspaper. Since central examinations are available in the Netherlands, and published by school inspections, a change in ranking cannot be seen as a shock in transparency. A school's ranking will move together with test scores which complicates disentangling the impact of information and simultaneous changes in quality, as observed by parents. Koning & Van der Wiel (2013) try to overcome this issue by exploiting the lag between registration and publication of quality indicators.

More recently, Nunes, Reis, & Seabra (2015) studied the effect of ranking publications on school enrolments and closures. Unlike Koning & Van der Wiel (2013), the authors analyze a dataset comprising data before and after publication of rankings in the newspaper. This allows them to single out the effect of information on parental choice. Using a pooled regression approach, they find that fewer parents send their children to poorly rated schools and these are likely to face closure after publication. This effect appears to be stronger for private schools. As we have stressed in the studies by Rosenthal in the UK (2004) and Koning & Van der Wiel in the Netherlands(2013), also in Portugal parents have access to quality indicators. Based on publicly available national test scores, school rankings are constructed and published by a newspaper. In this educational context, also other indicators of quality - apart from its ranking - influence the reputation of the school since "*at least some parents already had some information on school quality*" (Nunes et al., 2015).

In summary, the effect that is obtained by the studies reviewed above can be defined as the impact of *additional information* on parental choice, for *parents who are uninformed* on the quality distribution of schools. Parents that are well aware of this distribution – possibly consulting other quality measures – do not seem to respond to changes in information (Burgess et al., 2013; Hastings & Weinstein, 2008). The randomness in school selection by inspectorates combined with the absence of other accountability tools results in a truly exogenous information shock. In addition, by assessing the impact of these shocks on short-term outcomes for a large and randomly drawn sample, we are able to overcome the problems faced in earlier work. Also, we contribute to the existing literature by including school SES composition as an *outcome* variable. Studying the impact of information shocks on schools' composition complements the work of Koning & Van der Wiel (2013) by considering supply-side ‘pull factors’ that might affect this composition. The Flemish educational system serves as an ideal setting to study the responsiveness of parents and the ability of schools to ‘cherry-pick’ students. This unique setting will be explained in the next section.

3 Setting

3.1 Flemish education system

Flanders makes an interesting application because of its freedom of educational choice. All schools are subsidized in centralized way. The Flemish government aims to improve the opportunities of disadvantaged students by complementing funding for low SES schools. In order to distribute these additional resources, data needs to be obtained about students’ parental (economic and ethnic) background (Nusche, Miron, Santiago, & Teese, 2015). This unique data allows us to study the effect of inspection reports on socio-economic composition (see 4.).
Pupils (and their parents) in Flanders are totally free to select their preference school (i.e. there are no catchment areas). Moreover, every person has the right to establish an educational institution, creating a broad heterogeneity in the Flemish educational landscape. Providers include local community governments, private providers (mostly catholic schools) and a centralized state school system, all operating within overlapping geographic regions.

Since there are no tuition fees in compulsory education – irrespective of the provider - and schools are not allowed to select students, the (perceived) quality of schools remains the sole means to attract students and compete with other schools (MacLeod & Urquiola, 2015). Our motivation to focus on elementary schools in subsequent analyses is threefold. First, we expect parents to be more sensitive to elementary inspections due to a higher degree of homogeneity. Secondary schools can be more heterogeneous by offering some sort of vocational education and can thus not always be considered to be competitors. They aim at a ‘different market niche’. In Flanders, elementary schools offer the same education, irrespective of the location and provider of education. Second, total yearly expenditures that schools can charge to parents is limited to a fixed ceiling in elementary schools (i.e. 85 euro per school year). This ceiling is set by the central government, to guarantee effective free school choice since it prevents schools from influencing the student (socio-economic) composition. This policy is not yet implemented in secondary schools and hence, socio-economic composition could be affected by a school's (financial) policy. Third, inspection reports of elementary schools are published within two months following an inspection in contrast to secondary schools - where this sometimes takes a full school year (Penninckx, 2015). This allows the specification of a clear cut-off year in our DiD framework, developed in section 5.

In the decentralized Flemish system, school principals and teachers have significant autonomy. There are no central examinations, but school inspectorates monitor whether the content, budget formulation, and teaching activities are sufficient to attain the minimal goals and competences set by the central government. These goals are centrally imposed and equal for all (general) primary schools in Flanders (Hindriks, Verschelde, Rayp, & Schoors, 2010).

This unique educational setting allows us to single out the impact of information shocks (through school inspections) on parental school choice since other objective tools to compare schools are inexistent in Flanders (Woessmann, Lüdemann, Schütz, & West, 2007).

3.2 School inspectorates

School inspections are essential parts of 'evidence-based governance' concepts and have been implemented by many countries as a major strategy to assure and improve the effectiveness and quality of education systems (Altrichter & Kemethofer, 2015). Altrichter & Kemethofer (2015) break down inspectorates into a continuum between 'soft' and 'hard' governance. Examples of soft inspections are Austria, and to a lesser extent, Sweden, Ireland and the Czech Republic. The approach of these countries inspectorates' can be characterized by regular visits to schools. During these visits, schools are offered feedback based on some SWOT analysis, without classifying schools as failing or well-performing. Moreover, 'failing' schools, face no or only limited consequences (Ehren et al., 2015).

Contrasting models of 'hard' inspection can be found in England or the Netherlands where the main focus of inspectorates lies on targeting inspection visits of 'failing' schools. This type of inspection does rely on a set of standards (primarily student achievement data) in contrast to self-assessments in soft governance models.

The high reliance of school inspections on student achievement data in countries following a hard governance models complicates disentangling the impact of school inspections on enrolment growth and socio-economic composition of schools, as stated in the introduction. In

Flanders, inspections are the only source of accountability of schools towards the government (and parents) because of the absence of central examinations. The aforementioned argument that schools can only compete on quality due to the absence of tuition fees emphasizes the importance of school inspection reports as some sort of objective quality indicator for parents. Schools in Flanders receive public funding conditional on a 'favorable' judgement by the school inspectorate. The judgements given to schools can also be 'restricted favorable' and 'unfavorable'. In the former case, schools will face a follow-up inspection three years later to determine proper adjustments are put in place by the school. The latter category is assigned to schools that fail to achieve attainment targets and development goals set by the central government. These schools will in theory be closed but are effectively offered a 'second chance' to develop a strategy in order to overcome the structural deficiencies identified by the inspection. Hence, schools are never closed in reality and a negative evaluation only affects enrollment rates through its effect on the perceived quality of the school (Penninckx, 2015).

The Flemish approach towards school inspections can be seen as in between a soft and hard approach. That is, schools are classified as favorable, restricted favorable or unfavorable but do not face consequences, apart from deteriorating reputation. Moreover, school inspectorates do not form their judgement based on student achievement data since such data is not available. Due to the absence of this alternative quality indicator, parents can (and likely do) consult inspection reports available online (www.onderwijsinspectie.be). The final reports are published between 5 and 8 weeks following an inspection[5]. Currently, schools are inspected approximately once every ten years, which is a significantly lower frequency compared to other countries (De Volder, 2012). The inspection adopts a 'differentiated approach' which allows inspecting some schools more often. However, until now the inspectorate has not made use of this possibility (Penninckx, 2015). Exceptions occur when follow-up inspection are conducted (see above). Therefore, in our analysis, we only consider elementary schools that were not subject to an earlier inspection in the period 2010-2015[6]. As a result, the sample of inspected schools can be considered random. We will illustrate this graphically and statistically in the following section.

---

[5] In addition, a summary of the 'current state of education' is compiled based on inspection reports of the past school year. Although this document is aimed at informing the Minister of Education, it is also available online.

[6] Note that we can observe which reports are due to a follow-up and those that are not. School inspectorates indicate in every report whether it is a standard inspection (SI), follow-up (FU), or other.

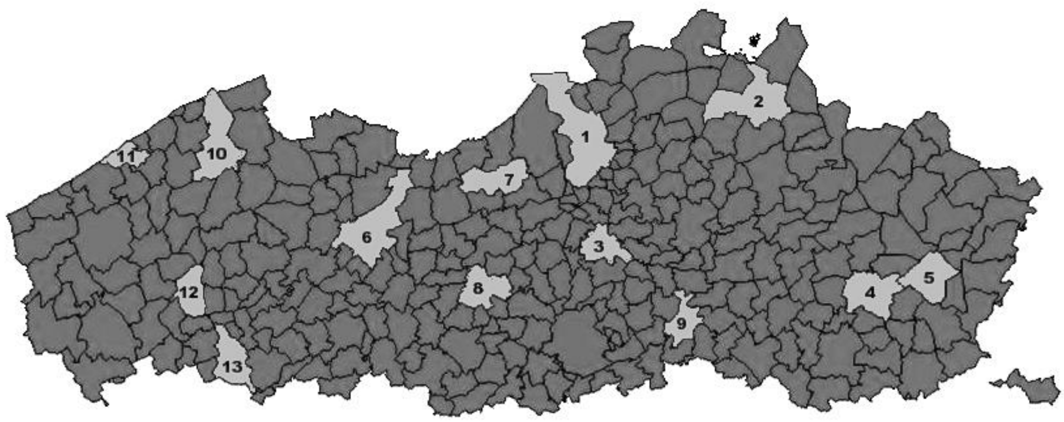
4 Data

4.1 Summary statistics

Our dataset is a result of merging school inspection data and administrative data covering the period of 2010-2015. We only look at elementary schools because we can isolate the effect of school inspections (see 3.1.). In Flanders, general elementary schools offer the same education, irrespective of the province the school is located in. However, some schools offer kindergarten and/or secondary schooling in the same administrative school. This might affect parents' choice who prefer not to switch between primary and secondary school. In Flanders, 92,1% of elementary schools jointly offers kindergarten, 40,1% secondary schooling, and 36,5% offers all three. In section 5 we will add school-level fixed effects such these (time-invariant) characteristics are taken into account when estimating the impact of school information shocks.

In the remainder of this paper we separate our analysis between urban and rural areas. This choice allows us to acknowledge the importance of distance for parental school choice. Proximity has been identified as one of the main drivers of school choice (Hastings, Kane, & Staiger, 2005; Koning & Van der Wiel, 2013). Although we do not have detailed geographical information, we can run the analyses for urban and rural areas separately to compare possible differences in mechanisms. In addition to our concern to account for proximity, we want to consider increasingly binding capacity constraints in urban areas in Flanders. Reputation effects will likely channel through differently in urban and rural areas because of this educational context. For example, if the capacity of a school is reached, it cannot grow further following a positive shock. Therefore, we hypothesize less (positive) responsiveness of urban schools.

Figure 1: City centers in Flanders



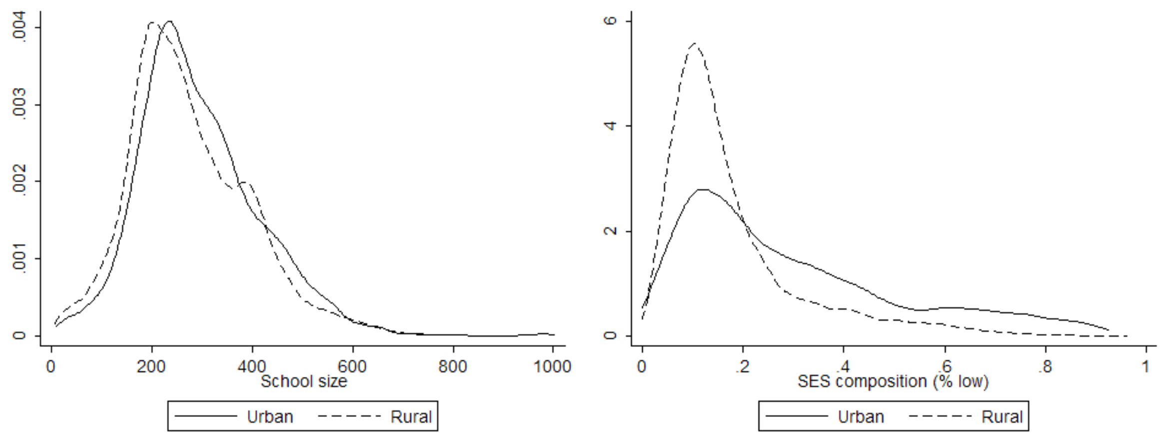
We define urban schools as schools located in one of the thirteen 'main cities'. These cities correspond to the latest revision of Flanders' structural plan (Flemish Government, 2011) and are indicated in Figure 1. All thirteen cities combined account for 24% (1.478.542) of Flanders' total population and 38% of its foreign population (Statbel, 1/1/2015). The number of students in compulsory education (in primary schools) equals 1,1 million (385.354) in total, and 266.570 (70.307) if we only consider the main cities in Flanders[7]. Although these thirteen cities are the largest in Flanders, their size is sufficiently small to assume city-wide competition without stratification (Hastings et al., 2005). It is therefore possible, and likely, that parents will not move to another city but will pick the school in their choice set that offers the best education.

Applying the above definition of urban schools, the number of schools in our dataset is lower in cities compared to rural areas: around 663 schools compared to 1582 rural schools. For each school we have 6 years (2010-2015) resulting in 3978 relative to 9489 observations.

Figure 2 displays the distribution (Kernel densities) of schools in rural and urban areas. It displays school size and socio-economic (SES) composition, which will be the outcome variables in our analyses. The left panel shows the school size and appears rather similar between areas. On average, urban schools are slightly larger (294 vs 275 students) and also the maximum is higher in urban areas (1004 vs 906 students). The standard deviation is equal, suggesting a relatively similar dispersion of school size in both areas.

The right panel shows the socio-composition of schools, measured as the percentage of students in the schools where the students' mother did not attain a secondary schooling degree. As can

Figure 2: School size and composition in rural and urban areas.



[7] More detailed statistics, and inspection reports (see below) are included in Table A 1.

be expected, this share is significantly higher in urban areas (29% vs 18%). Clearly, schools with a share of low-SES students that is beyond 25% are less common in rural areas. As a result, the socio-economic composition distribution of schools is much more dispersed in urban areas (SD 0.22 vs 0.14). However, we can see that also in rural areas, extremely segregated schools exist. The maximum values correspond to 96% and 93% for rural and urban areas, respectively.

The inspection reports cover school years 2012-2014. In this period, 767 elementary schools were subject to inspection (71,4% of the total of 1074 schools). Relative to the total of 2245 elementary schools (663 urban and 1582 rural schools), the inspectorate covered 34% of all schools in Flanders. The majority (59.5%) of all schools received a positive evaluation. The remaining 40.5% can be further disentangled into restricted positive (40,4%) and negative (0.1%). In the remainder of this paper, we group 'restricted positive' and 'negative' into the same category, as displayed in Table 1.[8]

Table 1: School inspections, coverage and classification.

Year	Inspected schools	Coverage (%)	Positive (%)	Negative (%)
2012	304	13.54	61.84	38.16
2013	268	11.94	58.21	41.79
2014	195	8.69	55.38	44.62
All	767	34.17	59.51	40.49

Note: The total coverage can be obtained by adding up yearly coverage rates since our panel is strongly balanced.

4.2 Random selection of schools for inspection

In section 3.2 we argued that the school inspectorate selects schools for inspection in a random manner. In order to assess the validity of this statement we compare the distribution of variables in our dataset. We compare two sets of elementary schools: those schools inspected ('treatment'), irrespective of the outcome, and all other schools ('control'). We selected four variables and compare before-2012 (first inspection) values between treatment and control groups: school size, school growth prior to 2012, teacher absenteeism, and class size.

[8] In our dataset, only 3 schools received a 'negative' evaluation. Considering this very low number, schools that are classified as 'restricted' positive are viewed as 'negative' evaluations in Flanders. As can be expected, our subsequent analyses hold irrespective of including these schools.

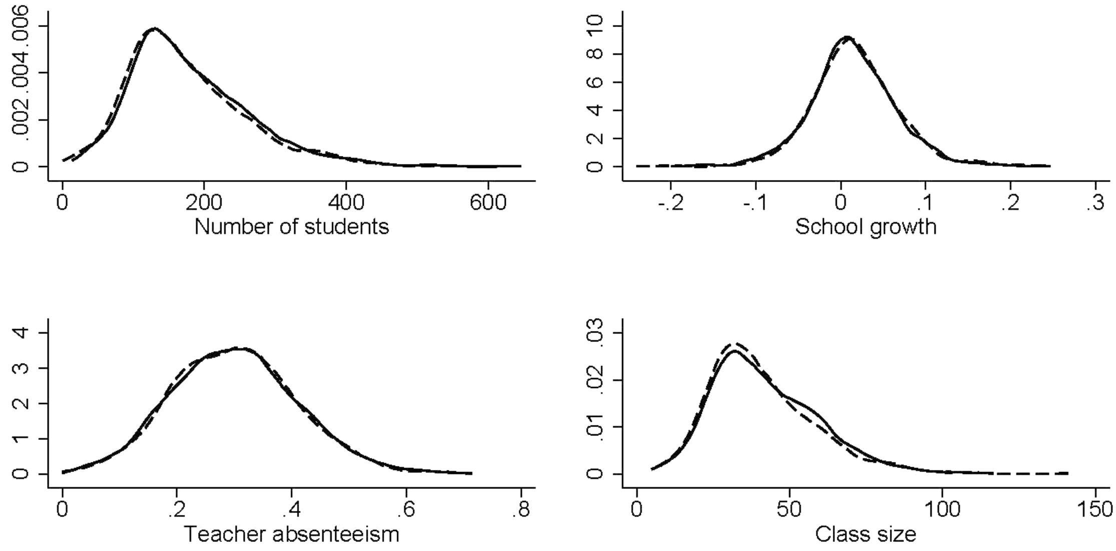
In Figure 3, densities of both groups of elementary schools are plotted for all four variables. From these graphs we can see that schools that were inspected and do not significantly differ from schools that did not receive a visit from the inspectorate[9]. We check this graphical analysis by calculating the Kolmogorov-Smirnov statistic and find that none of the compared variables differs significantly between inspected and non-inspected schools. The null hypothesis of a common underlying distribution cannot be rejected. Also, additional t-tests cannot reject the equality of means. This strengthens our statements that the selection process of schools for inspection is random. The low frequency of inspections and apparent exogeneity in the selection of schools for inspections leads to shocks in transparency. In the following section, we will assess the impact of these shocks on enrolment and the socio-economic composition of schools facing a (negative or positive) evaluation.

5 Empirical strategy

Schools (in Flanders) are not able to change their price to compete over the 'market of students'. In addition, due to the absence of standardized exams, there is few information available for parents to compare the quality of schools. We hypothesize that any increase in transparency will shift market share from low-quality schools to high-quality schools. Two important assumptions need to be met. First of all, the sample of schools should be random. That is, low-performing schools should not face inspection more often. Second, the release of the reports should increase transparency. If parents have full information over the quality distribution of schools, this additional information will not result in market share shifts. Now we have established the random nature of the school selection process (see 4.2), we will develop an empirical strategy to estimate the 'treatment' effect (inspection report publication) faced by schools.

[9] The graphs presented in Figure 3 are obtained using the full sample of schools in Flanders. If we repeat our analysis using only the subsample of urban schools, the resulting graphs are no longer smooth due to a smaller number of observations but treatment and control groups remain comparable. This graphical analysis is again confirmed by calculating the significance levels of the Kolmogorov-Smirnov statistics. Graphical and statistical tests for both samples are available upon request.

Figure 3: Comparison of treatment and control group.



School inspectorates report on the school quality. Other quality indicators are not available to parents and schools do not face closure when they receive a negative school evaluation (see 3.2). As a result, the identification strategy does not suffer from endogeneity in the choice set since the number of schools in an area is not affected by school inspections. All things equal, only the available information to parents is affected.

We estimate whether school growth (both enrolment and socio-economic composition) is affected by the release of inspection data. This estimation allows us to check whether information, as perceived by parents, has actually changed. If inspection reports reveal information that was known to all parents already, we expect no change in enrollment rates after this release. The opposite holds for new information, as parents might reconsider their school choice. By comparing school growth before and after the release of school inspection reports, we assess the responsiveness of parents to this information.

5.1 Theoretical framework

In particular, we apply a Difference-in-Difference framework to obtain the parameter estimate of parental responsiveness with respect to our outcome variables. Following the notation of Cameron & Trivedi (2005) we specify the binary independent variable:

$$D_{it} = \begin{cases} 1 \text{ if } i \text{ receives treatment in period } t \text{ or before,} \\ 0 \text{ otherwise.} \end{cases} \quad (1)$$

In the notation of Athey & Imbens (2006), D_{it} can also be written as $Treatment_i * Time_t$. Specifying the regression equation of a fixed effects model for Y_{it} to obtain the coefficient on D_{it}:

$$Y_{it} = \gamma_i + \delta_t + \beta_1 X_{it} + \beta_2 D_{it} + \varepsilon_{it} \quad (2)$$

Where δ_t = time-specific fixed effect ($Time_t$), and γ_i = individual-specific fixed effect, which includes the treatment dummy ($Treatment_i$). The latter can be eliminated by first differencing:

$$Y_{it} - Y_{it-1} = \delta_t - \delta_{t-1} + \beta_1(X_{it} - X_{it-1}) + \beta_2(D_{it} - \overline{D}_{it-1}) + (\varepsilon_{it} - \varepsilon_{it-1}) \quad (3)$$

The obtained estimator of the treatment effect (β_2) is called the first-differences estimator. If data are available pre- and post-treatment and if not every *i* received the treatment, β_2 can be measured consistently. In that case, the first-differences estimator of the fixed effects model reduces to a simple estimator called the Differences-in-Differences (DiD) estimator.

The interpretation of β_2 as a causal parameter relies on the assumption that after controlling for X_{it}, $Treatment_i$ (included in γ_i), and a time-specific fixed effect (δ_t), the post-treatment difference between treatment and control groups can be completely attributed to the treatment effect:

$$E(Y_{0it}|\gamma_i, X_{it}, \delta_t, D_{it}) = E(Y_{0it}|\gamma_i, X_{it}, \delta_t) \quad (4)$$

The DiD estimator relies strongly on the above model assumption, which is often not made explicit. The main advantage of this estimator follows from its applicability when repeated cross section data rather than panel data are available (Cameron & Trivedi, 2005, p.768). Because of this reason, economists have applied the DiD model for causal inference following policy shocks. However, for both types of data, it should be assumed that the time effects δ_t are common across treatment and control group. This corresponds to the 'parallel trends assumption' (PTA), which will hold, by definition, if the treatment is truly random. The PTA is captured by equation (4).

In essence, there is no deep difference between DiD and fixed effects. DiD is a version of fixed-effects estimation using aggregate data (Angrist & Pischke, 2008, p.170) [10]. In a balanced panel,

[10] In a DiD model, fixed effects are included at the group level (G) such that the expected value of the outcome variable for $i \in G$ equals the mean of *G*: $E(Y_{0iGt}|s,t) = \delta_t + \gamma_G)$ (Angrist & Pischke, 2008, p.170). In equation (4), γ_i will be substituted by γ_G. The assumptions imposed by including individual-specific fixed effects are essentially less restrictive. It is straightforward to see that group level characteristics are also captured by individual-specific fixed effects. Hence γ_i by definition includes γ_G, and hence, γ_G will be dropped when estimating the regression equation.

the same result will be obtained by estimating using either individual fixed effects, or with a dummy variable for treated individuals. If panel data are available, both first-differences (DiD) or deviation from means estimators ("within estimator") are appropriate. However, if $T > 2$ and ε_{it} is iid (what we need to assume), the first-differences estimator is less efficient (Cameron & Trivedi, 2005, p.705). Hence, because of this, we choose to apply the within estimator in a fixed effects regression[11]:

$$Y_{it} - \bar{Y}_i = \delta_t - \bar{\delta} + \beta_1(X_{it} - \bar{X}_i) + \beta_2(D_{it} - D_i) + (\varepsilon_{it} - \bar{\varepsilon}_i) \tag{5}$$

Estimation by fixed effects regression to remove the individual-specific fixed effect γ_i is consistent, conditional on the exogeneity of the treatment D_{it}. If D_{it} and ε_{it} are correlated for any i and t our estimate of β_2 will be inconsistent (Wooldridge, 2007). The exogeneity of the treatment again corresponds to (4) and has been explained in sections 3.2 and 4.2.

5.2 Application

Adapting (2) to our specific setting, we obtain the regression equation:

$$School_{i,t} = \gamma_i + \beta_1 Time_t + \beta_2 SI_{i,t} + \varepsilon_{i,t} \tag{6}$$

$School_{i,t}$ indicates the school size or school SES composition at school i in time t. This is the first study to include socio-economic composition as an outcome variable. We include fixed effects at the school level (γ_i). Note that a schools' size and composition can be affected in two ways: parents choosing a school when enrolling their children in first grade primary schools, and switching schools after this initial decision. Both mechanisms are likely related to the school structure (e.g. secondary education in the same schools reduces later switching costs). Including school-specific fixed effects enables us to account for time-invariant (un)observed characteristics. These include the aforementioned structure, but also the type of provider (e.g. public or private), and other school-particular factors that remain constant over time.

In a general two-period DiD setting, the dummy variable $Time_t$ separates the period before inspection (=0) and after the publication of the inspection report (=1). However, given that inspections have taken place between 2012 and 2014 (see 4.1), $Time_t$ corresponds to a set of

[11] Note that by including school-specific fixed effects, we also satisfy the Stable Unit Treatment Value Assumption (SUTVA) which is necessary for causal inference but does not necessarily hold when applying DiD (Angrist, Imbens, & Rubin, 1996).

dummies for every year which accounts for year-specific growth changes. Considering the exogenous timing of school inspections, simultaneous shocks in enrollment together with inspections are very unlikely.

Our variable of interest is $SI_{i,t}$ and indicates whether school i received a school inspection in time t (or before) and equals 0 otherwise. This indicator variable can be seen as D_{it} in (2) and can also be written as $Treatment_i * Time_t$. The dummy variable $Treatment_i$ is not included in (6) since estimation by fixed effects will eliminate this time-invariant variable. The associated parameter β_2 indicates the sensitivity of school growth and SES composition to inspection evaluations.

First, we estimate the impact of school inspections on school size and SES composition. This corresponds to estimating equation (6). Second, we include an additional interaction to separate the effect for negative and positive evaluations. Hence, this model can be estimated using the following regression function:

$$School_{i,t} = (6) + \beta_2 SI_{i,t} * Neg_i + \beta_2 SI_{i,t} * Pos_i \qquad (7)$$

Where Neg_i and Pos_i indicate the type of evaluation school i received. In (7) the term $\beta_2 SI_{i,t}$ drops, since the inspection is either positive or negative. As a result, we estimate four models: (6) and (7) for enrollment and (6) and (7) for SES composition. Every set of four models is estimated separately for rural and urban schools in order to properly account for different mechanisms described in section 4.

6 Results

Estimation results are displayed in Table 2 and Table 3 for rural and urban schools, respectively. As we explained in section 4, the number of schools in urban areas (663) is much lower than the number of rural schools (1582). This can be seen from the bottom row in Tables 2 and 3. Despite this difference in number, the share of inspected schools is equal in both areas (34%), further supporting the claim that schools are selected randomly for inspection. In all models, explained variance is relatively low (between 1 and 7%). This might be due to the lack of geographical data, which has been proposed as an important determinant of parental school choice (Simon Burgess, Greaves, Vignoles, & Wilson, 2014; Hastings et al., 2005). In both tables, we purposely display estimated coefficients of the year dummies to illustrate the trend in enrolment and socio-economic composition of schools. In rural schools (Table 2), we report rising coefficients on school size, indicating an overall growth trend, consistent with the educational setting in Flanders. In contrast, the share of low-educated mothers of children in schools appears to be decreasing every year between 2010 and 2015.

In Table 2, the results are presented of estimating regression equations (6) and (7) for enrolment and SES composition in rural elementary schools. We find that school inspections, on average, increase school growth while school SES composition does not significantly change. Further disentangling the impact of school inspections, we find that the positive enrolment growth effect can be explained by schools receiving a positive evaluation in rural areas. A negative evaluation does not appear to affect school growth. It does however, lower the share low-educated mothers in a school. This finding seems counterintuitive since we do not expect negatively evaluated school to attract high educated parents. Taking a closer look at the data, we find that the share of low SES children in rural areas is generally low (Figure 1), with some exceptions of schools consisting only of low-SES students. The coefficient suggesting this finding is no longer significant if outlying low SES schools are excluded from the analysis[12].

Contrasting the findings of Hastings, Van Weelden, & Weinstein (2007) that low SES families experience higher information costs, we cannot find a significant change of SES composition in schools facing a positive evaluation. This suggests that school growth following a positive inspection is driven in a relatively symmetric manner between low- and high-SES parents. If this were not true, we would expect the share of low SES children to change significantly in response to a positive evaluation.

Table 2: Regression results (1)

Rural	Enrolment (students)		SES composition (% low)	
Variables	(6)	(7)	(6)	(7)
School inspection in t	2.993**		-0.000811	
	(1.313)		(0.00160)	
Positive evaluation		3.654**		0.00176
		(1.595)		(0.00212)
Negative evaluation		2.076		-0.00437**
		(2.019)		(0.00220)
2011	1.757**	1.757**	-0.00449***	-0.00449***
	(0.781)	(0.781)	(0.000617)	(0.000617)
2012	4.473***	4.461***	-0.00757***	-0.00762***
	(0.939)	(0.941)	(0.000889)	(0.000889)

[12] If we leave out schools that contain mainly low SES students (low educated share > 50%), we keep 97% of our observations in rural areas. Also, all subsequent analyses are robust to this adjusted sample.

2013	6.537***	6.533***	-0.00972***	-0.00974***
	(1.126)	(1.126)	(0.00111)	(0.00111)
2014	9.815***	9.814***	-0.0109***	-0.0109***
	(1.277)	(1.277)	(0.00133)	(0.00132)
2015	12.56***	12.56***	-0.0128***	-0.0128***
	(1.441)	(1.441)	(0.00154)	(0.00154)
Constant	269.2***	269.2***	0.184***	0.184***
	(0.797)	(0.797)	(0.000780)	(0.000779)
School FE	YES	YES	YES	YES
Observations	9,492	9,492	9,489	9,489
R-squared	0.048	0.048	0.032	0.034
Number of schools	1,582	1,582	1,582	1,582

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

When we shift our attention towards urban schools, we find that the enrolment growth trend is similar to the one we found in rural schools. In contrast to rural schools, the decreasing share of low-educated mothers seems to come to a halt in 2014. Linking this finding to immigration statistics, we are able to explain the trend reversal due to an increasing inflow of immigrants in Flanders after 2014 (Frontex, 2015). This effect did not significantly influence rural school composition, indicating that the inflow of immigrants remained restricted to urban areas.

Compared to rural areas, school size does not respond strongly to school inspections. It is unlikely that parents in urban areas are indifferent to shocks in information. A more probable explanation of this limited response could be a structurally zero growth level in urban areas (see section 3). Parents possibly want to send their children to positively evaluated schools but are not able to do this due to capacity constraints. Interestingly, we when look at coefficients obtained for changes in school composition, we find that positive school evaluations significantly reduce the share of low SES children in urban schools. In contrast to rural schools, where parental responses appeared symmetric between SES groups of parents, this cannot hold in urban schools. If parents would respond equally to a positive school inspection report - as they do in rural areas -, the coefficient on school composition would be zero, especially considering the fact that schools are not allowed to select students.

This finding suggests that, due to the inability of schools to grow further, school administrators seem to leverage the increased demand to 'improve' their SES composition following the publication of a school inspection report[13]. Hence, parents actually respond equally sensitive to the release of quality information, as illustrated in rural schools, but are not granted equal access to their school of choice.

[13] This finding is robust to the restricted sample applied before.

Table 3: Regression results (2).

Urban	Enrolment (students)		SES composition (% low)	
Variable	(6)	(7)	(6)	(7)
School inspection in t	-0.681		-0.00805**	
	(2.129)		(0.00359)	
Positive evaluation		-0.610		-0.0127***
		(2.493)		(0.00466)
Negative evaluation		-0.789		-0.000897
		(3.088)		(0.00448)
2011	3.354***	3.354***	-0.00305***	-0.00305***
	(0.970)	(0.970)	(0.00114)	(0.00114)
2012	7.769***	7.770***	-0.00420**	-0.00426**
	(1.548)	(1.549)	(0.00184)	(0.00184)
2013	10.95***	10.95***	-0.00518**	-0.00513**
	(1.818)	(1.818)	(0.00252)	(0.00252)
2014	14.36***	14.36***	-0.00362	-0.00364
	(2.182)	(2.182)	(0.00317)	(0.00317)
2015	18.58***	18.58***	-0.00470	-0.00471
	(2.425)	(2.425)	(0.00350)	(0.00350)
Constant	285.5***	285.5***	0.299***	0.299***
	(1.179)	(1.179)	(0.00160)	(0.00160)
School FE	YES	YES	YES	YES
Observations	3,978	3,978	3,978	3,978
R-squared	0.067	0.067	0.010	0.013
Number of schools	663	663	663	663

Robust standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

6.1 Robustness checks

Despite the similarity between inspected and non-inspected schools (Figure 3), we cannot assume that ex-post positively and negatively evaluated schools were comparable ex-ante. Therefore, it might be the case that the composition of schools is unaffected by school inspections because positively evaluated schools only contain high SES parents, and vice versa for negative evaluations. If this were true, an asymmetric response of parents (e.g. only high SES parents respond to inspection data) would result in insignificant coefficients on size and composition following a negative evaluation and on composition following a positive evaluation. In this case, the significant coefficient on size for positively evaluated schools could be attributed to high SES parents sending their children to these schools, away from not (yet) inspected schools.

Figure 4: Comparing the distribution of school composition before inspections.

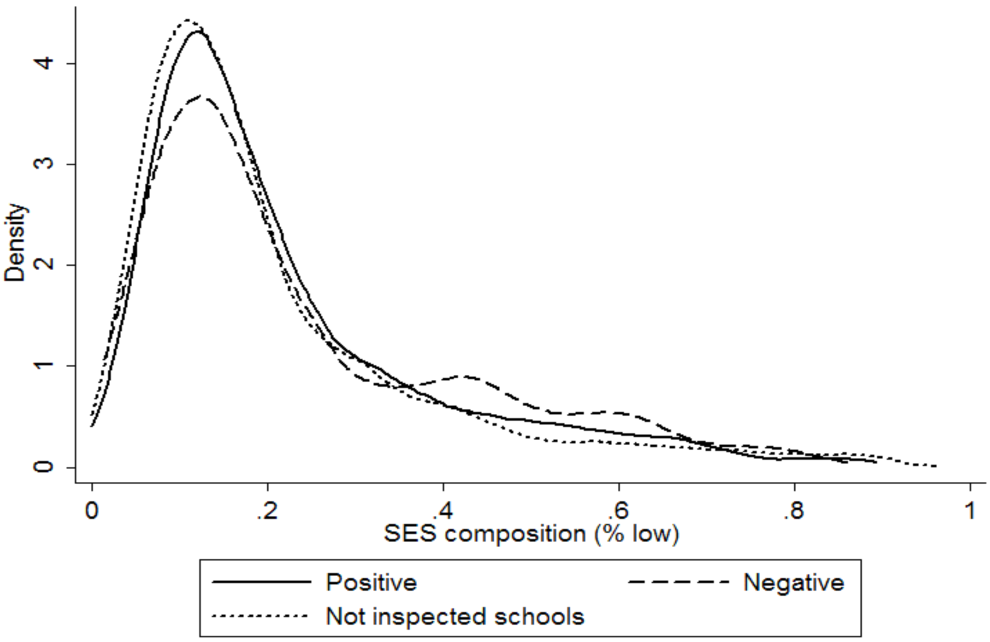


Figure 4 plots the density functions of positively, negatively, and not inspected schools using data before 2012. Although the density of low SES schools (>40%) is greater in negatively evaluated schools, the overall density largely coincides with positively evaluated schools and schools that did not face an inspection. Hence, we can conclude that the absence of (1) significant changes in school composition following an inspection and (2) significant changes in school size following a negative evaluation is not simply due to the ex-ante composition of schools. Parents in rural areas appear to respond more strongly to positive evaluations, and irrespective of their socio-economic background.[14]

In the economics literature, fixed effects models have received a lot of attention since weaker assumptions (although still strong, see above) are required compared to those needed to establish causation with random effects (RE) models (Cameron & Trivedi, 2005). However, when causation is clear, random effects may be more appropriate as an estimation method. RE models can be preferred because of several practical weaknesses of the FE model (Cameron & Trivedi, 2005, p.715). For example, coefficients on time-invariant variables cannot be estimated since they are 'absorbed' by the FE model. In our application, this includes the 'treatment' variable.

[14] This finding is consistent with evidence for the Netherlands in Koning & Van der Wiel (2013).

Using an Hausman specification test we are unable to reject the hypothesis that differences in coefficients are not systematic. In Tables A2 and A3 RE estimation results are presented. As is clear from these tables, the results from the RE approach are very similar to the FE results.

[To be extended: Quantile regressions; restricted sample (<2014); dual panel Fig 4]

7 Conclusion and discussion

In contrast to many other countries, school inspections are not targeted in Flanders. If this were the case, a negative correlation between inspections and school performance will automatically appear (Rosenthal, 2004). We have shown that the selection process of school inspections can be considered exogenous and the resulting information is the only source of 'objective' data available to parents due to the absence of central examinations. As a result, growth rates can be compared before and after the 'shock', within a Difference-in-Difference framework. In particular, as the education market becomes more transparent, parents see more clearly which schools offer high value. Flanders makes an interesting application because of its freedom of educational choice. All schools are subsidized and pupils (parents) are totally free to select their preference school (there are no catchment areas). Since there are no tuition fees in elementary school and schools are not allowed to select students, the (perceived) quality of schools remains the sole means to attract students and compete with other schools. As a result, this increase in transparency will shift market share from low-quality schools to high-quality schools. Many studies have investigated the relationship between parental choice (e.g. property prices) and information on student achievement (Clapp, Nanda, & Ross, 2008; Figlio & Lucas, 2004), while evidence is still lacking on the responsiveness of parents to information on the added value of schools.

Our findings indicate that parents do value the information revealed by inspection reports. We find that parents respond strongly to school inspection reports when capacity constraints are not reached. That is, in rural areas where school choice is effectively free, schools grow faster after receiving a positive inspection report. The opposite does not hold when schools are evaluated negatively. Interestingly, this effect appears to be equally driven by parents from different socio-economic background. No significant difference in school composition can be observed in rural schools, following positive school inspections (which increases schools size).

We run separate regressions for urban areas to account for possible differing mechanisms as a result of capacity constraints faced by many urban schools. In these schools, we find no significant change in school size following the publication of school inspection reports. This in contrast with findings for rural schools, though consistent with schools' capacity constraints. When we consider a school's socio-economic composition as the outcome variable, we find that a positive evaluation 'improves' the composition of the school while a negative evaluation has no significant impact.

This finding suggests that, due to the inability of schools to grow further, school administrators seem to leverage the increased demand to 'improve' their SES composition following the publication of a school inspection report. Hence, parents actually respond equally sensitive to the release of quality information, as illustrated in rural schools, but are not granted equal access to their school of choice.

Policy implications are straightforward. School inspections are clearly important for school choice since parents do care about this information. Considering the importance attached by parents to school inspection data, policymakers might shift their attention to evaluating the performance of inspectorates. Studies have shown a high variety between countries (M. Ehren et al., 2015) while improving school inspection practices has only received little attention (M. Ehren, Altrichter, McNamara, & O'Hara, 2013; M. C. Ehren, 2016). Also, non-selectivity of schools should be guaranteed to allow equal consideration of all children. Mechanisms should be implemented to monitor real freedom in school choice. If this goal is achieved, school inspections will reach their full potential as policy instruments to improve student achievement through enhanced competition, without leading to further segregation of schools.

References

Altrichter, H., & Kemethofer, D. (2015). Does accountability pressure through school inspections promote school improvement? *School Effectiveness and School Improvement*, *26*(1), 32–56.

Angrist, J. D., Imbens, G. W., & Rubin, D. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association.*

Angrist J. D. and J. S. Pischke. (2008). Mostly Harmless Econometrics : An Empiricist's Companion. *Massachusettts I Nstitute of Technology and The London School of Economics*, (March), 290.

Athey, S., & Imbens, G. W. (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica*, *74*(2), 431–497.

Burgess, S. (2016). *Human Capital and Education : The State of the Art in the Economics of Education* (No. 9885).

Burgess, S., Greaves, E., Vignoles, A., & Wilson, D. (2014). What Parents Want: School Preferences and School Choice. *Economic Journal*, *125*(August 2014), 1262–1289.

Burgess, S., Wilson, D., & Worth, J. (2013). A natural experiment in school accountability: The impact of school performance information on pupil progress. *Journal of Public Economics*, 1–45.

Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics - Methods and Applications.*

Clapp, J. M., Nanda, A., & Ross, S. L. (2008). Which school attributes matter? The influence of school district performance and demographic composition on property values. *Journal of Urban Economics*, *63*(2), 451–466.

De Volder, I. (2012). *External school evaluation in Europe. A comparative study.* Antwerp, Apeldoorn: Garant.

de Wolf, I. F., & Janssens, F. J. G. (2007). Effects and side effects of inspections and accountability in education: an overview of empirical studies. *Oxford Review of Education*, *33*(3), 379–396.

Ehren, M., Altrichter, H., McNamara, G., & O'Hara, J. (2013). Impact of school inspections on improvement of schools-describing assumptions on causal mechanisms in six European countries. *Educational Assessment, Evaluation and Accountability*, *25*(1), 3–43.

Ehren, M. C. (2016). School Inspections and School Improvement; the Current Evidence Base. In *Methods and Modalities of Effective School Inspections* (pp. 69–85).

Ehren, M., Gustafsson, J. E., Altrichter, H., Skedsmo, G., Kemethofer, D., & Huber, S. G. (2015). Comparative Education Comparing effects and side effects of different school inspection systems across Europe. *Comparative Education*, *51*(3), 375–400.

Figlio, D. N., & Lucas, M. E. (2004). What's in a Grade ? School Report Cards and the Housing Market. *American Economic Review*, *94*(3), 591–604.

Frontex. (2015). *Annual Risk Analysis.*

Gaertner, H., Wurster, S., & Pant, H. A. (2013). The effect of school inspections on school improvement. *School Effectiveness and School Improvement*, *3453*(April 2014), 1–20.

Grek, S., Lawn, M., Ozga, J., & Segerholm, C. (2013). Governing by inspection? European inspectorates and the creation of a European education policy space. *Comparative Education*, *49*(4), 486–502.

Hastings, J. S., Kane, T. J., & Staiger, D. O. (2005). Parental Preferences and School Competition: Evidence from a Public School Choice Program. *National Bureau of Economic Research*, (Working Paper 11805), 160.

Hastings, J. S., Van Weelden, R., & Weinstein, J. M. (2007). Preferences, Information, and

Parental Choice Behavior in Public School Choice. *National Bureau of Economic Research Working Paper Series, No. 12995.*

Hastings, J. S., & Weinstein, J. M. (2008). Information, School Choice, and Academic Achievement: Evidence from Two Experiments. *Quarterly Journal of Economics*, *123*(4), 1373–1414.

Hindriks, J., Verschelde, M., Rayp, G., & Schoors, K. (2010). *School autonomy and educational performance: within-country evidence* (No. 2010082).

Hsieh, C. T., & Urquiola, M. (2006). The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program. *Journal of Public Economics*, *90*(8–9), 1477–1503.

Imai, K., & Kim, I. S. (2016). *When Should We Use Linear Fixed Effects Regression Models for Causal Inference with Longitudinal Data ?* (Diss. Working paper).

Klerks, M. (2012). The effect of school inspections: a systematic review. *School Improvement*, 1–33.

Koning, P., & Van der Wiel, K. (2013). Ranking the Schools : How School Quality Information Affects School Choice in the Netherlands. *Journal of the European Economic Association*, *11*(2), 466–493.

Luginbuhl, R., Webbink, D., & de Wolf, I. (2009). Do Inspections Improve Primary School Performance? *Educational Evaluation and Policy Analysis*, *31*(3), 221–237.

Macleod, W. B. (2007). Reputations , Relationships , and Contract Enforcement. *Journal of Economic Literature*, *45*(3), 595–628.

MacLeod, W. B., & Urquiola, M. (2015). Reputation and School Competition. *American Economic Review*, *105*(11), 3471–3488.

Mizala, A., & Urquiola, M. (2013). School markets: The impact of information approximating schools' effectiveness. *Journal of Development Economics*, *103*(1), 313–335.

Nunes, L. C., Reis, A. B., & Seabra, C. (2015). The publication of school rankings: A step toward increased accountability? *Economics of Education Review*, *49*, 15–23.

Nusche, D., Miron, G., Santiago, P., & Teese, R. (2015). Oecd reviews of school resources - Flemish community of Belgium 2015. *OECD Reviews of School Resources.*

Penninckx, M. (2015). *Inspecting School Inspections.* Universiteit Antwerpen.

Rosenthal, L. (2004). Do school inspections improve school quality? Ofsted inspections and school examination results in the UK. *Economics of Education Review*, *23*(2), 143–151.

Schutz, G., West, M. R., & Wossmann, L. (2007). *School accountability, autonomy, choice, and the equity of student achievement: International evidence from PISA 2003. Organisation for Economic Co-operation and Development.*

Woessmann, L., Lüdemann, E., Schütz, G., & West, M. R. (2007). *Autonomy , Choice , and the Level of Student Achievement. International Evidence from PISA 2003.* (No. 13).

Wooldridge, J. (2007). What's New in Econometrics? *Lecture 10*, 1–31.

Appendix

Table A 1: Summary statistics per city center.

City	Size	Students#		Schools*					Inspection	
		All	Primary	All	P	K	S	KPS	Positive	Negative
1	83.709	20.889	5.526	66	33	91%	97%	88%	69,2%	30,8%
2	513.570	42.614	15.235	137	81	93%	70%	63%	44,4%	55,6%
3	117.886	11.060	1.588	27	10	90%	100%	90%	66,7%	33,3%
4	65.463	16.722	5.536	58	33	94%	55%	52%	74,1%	25,9%
5	253.266	50.513	13.336	157	71	89%	82%	75%	52,3%	47,7%
6	76.331	22.864	4.732	67	31	94%	45%	45%	15,4%	84,6%
7	75.219	12.612	2.685	35	14	93%	0%	0%	50%	50%
8	98.376	8.386	1.393	24	7	71%	29%	14%	0%	100%
9	83.975	13.956	2.950	36	16	88%	94%	81%	66,7%	33,3%
10	70.460	14.164	4.218	46	27	96%	59%	59%	37,5%	62,5%
11	60.386	20.199	5.464	74	38	92%	50%	50%	71,4%	28,6%
12	74.289	16.991	3.926	39	19	95%	53%	53%	NA	NA
13	42.637	15.600	3.718	59	27	93%	96%	89%	75%	25%
Cities	1.478.542	266.570	70307	825	407	91,7%	68,1%	62,9%	58,1%	41,9%
Total	6.160.592	1.121.479	385.354	3.585	2.173	92,1%	40,1%	36,5%	59,5%	40,5%

Note: For the sake of presentation, only values are shown calculated using 2012 data. #: 'Students' indicates the number of students in compulsory education and kindergarten. *: P: Primary; K: Kindergarten; S: Secondary. 'Kindergarten' indicates the percentage of elementary schools that also provide kindergarten, secondary, and all three together: KPS. NA: non-available data.

Table A 2: Random effects regression results (1).

Rural	Enrollment (Students)		SES composition (% low)	
Variable	(6)	(7)	(6)	(7)
School inspection in t	2.993**		-0.000811	
	(1.313)		(0.00160)	
Positive evaluation		3.654**		0.00176
		(1.595)		(0.00212)
Negative evaluation		2.076		-0.00437**
		(2.019)		(0.00220)
Constant	269.2***	269.2***	0.184***	0.184***
	(0.797)	(0.797)	(0.000780)	(0.000779)
Year controls	YES	YES	YES	YES
School FE	YES	YES	YES	YES
Observations	9,492	9,492	9,489	9,489
Number of schools	1,582	1,582	1,582	1,582

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table A 3: Random effects regression results (2).

Urban	Enrollment (Students)		SES composition (% low)	
Variable	(6)	(7)	(6)	(7)
School inspection in t	-0.940		-0.00810**	
	(2.148)		(0.00358)	
Positive evaluation		-0.894		-0.0129***
		(2.505)		(0.00463)
Negative evaluation		-1.012		-0.000755
		(3.135)		(0.00447)
Constant	285.5***	285.5***	0.299***	0.299***
	(4.450)	(4.450)	(0.00851)	(0.00851)
Year controls	YES	YES	YES	YES
School FE	YES	YES	YES	YES
Observations	3,978	3,978	3,978	3,978
Number of schools	663	663	663	663

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1