Do Better Schools Raise Housing Prices? Evidence from Paris School Zoning*

Gabrielle Fack
(Harvard University and PSE)

Julien Grenet
(PSE)

October 2007

Abstract

This paper provides the first empirical evidence on the impact of middle school quality on housing prices in Paris, using comprehensive data on both school zoning and real estate transactions over the period 1997-2003. Because it is closely linked to spatial mobility, the willingness to pay for better schools is a crucial parameter for the calibration of optimal school zoning policy simulations. Building on Black's (1999) approach, we develop a matching framework to carefully compare sales across school attendance district boundaries as a way to deal with the endogeneity of school quality. We find that a standard deviation increase in middle school quality raises prices by about 2%, which implies that the fraction of housing price differentials across school zones that can be explained by school quality differential amounts to about 7% in Paris. In addition, we test the prediction of general equilibrium models (Nechyba (1999; 2000; 2003)) that the availability of private schools should tend to mitigate the impact of public school quality on residential segregation. In line with this prediction, we find evidence that the presence of private schools in certain Parisian neighborhoods, by offering an outside option to parents, tends to attenuate the capitalization of public school quality in the price of real estate.

JEL Classification H41, I21, I28, R21
Keywords School zoning policies, housing markets, residential segregation

*Mailing address: Paris School of Economics, Bureau B030, 48 bd Jourdan, 75014 Paris, France. Email: julien.grenet(at)ens.fr et gabrielle.fack(at)ens.fr.

We wish to thank Sandra Black, Esther Duflo, Steve Gibbons, Caroline Hoxby, Philip Oreopoulos and Thomas Piketty from their helpful comments and suggestions. We are also grateful to Florence Defresne, Françoise Œuvrard, Fabienne Rosenwald from the *Direction des Études et de la Prospective* of the French ministry of Education, Jean-Paul Caille and Alain Mariani from the *Rectorat de Paris* for their help in collecting the educational data and to the Notaries' Chamber of Paris for the data on house prices. This work was financially supported by CEPREMAP (Centre pour Recherche Économique et ses Applications).

1 Introduction

Motivation In France, as in the United States, there is an ongoing debate on how school zoning affects both educational and residential segregation. When French policy makers implemented strict school zoning in 1963 (a unique assigned middle school per residential location), they believed that it would serve as an efficient way of preventing school segregation. What they did not take into account, however, is that parents might care so much about school quality that they can choose to "vote with their feet" and change residence to make sure that their children attend the school of their choice. Theoretical models (including Benabou (1993), Fernandez and Rogerson (1996)) have shown that the existence of peer effects in education yields income and residential sorting in equilibrium.

One of the key parameters underlying the residential sorting that arises in the presence of school zoning is the willingness of parents to pay for school quality. Because it plays an important role in the location that parents choose in equilibrium, this parameter is an essential ingredient of any model aiming at deriving the welfare implications of alternative schemes of pupil allocation (strict zoning, soft zoning, school choice, etc.). In this paper, we propose an empirical framework to estimate the willingness to pay for better public middle schools in Paris

Related literature Several empirical papers have sought to test the theoretical prediction that housing prices should be higher in districts with good school quality than in districts with lower school quality. The main estimation problem is that measuring the effect of school quality on housing prices is a complicated task, since better schools tend to be located in wealthier neighborhoods because of the higher educational attainment of pupils drawn from privileged socio-economic backgrounds. If the estimation strategy does not correct for observable and unobservable neighborhood characteristics, potentially correlated both with housing prices and school quality, then the estimation of the marginal willingness to pay for a better school might suffer from severe biases. In an attempt to address this endogeneity problem, Black (1999) first suggested to compare the prices of houses located on opposite sides of a common primary schooling attendance district boundary. The identifying assumption is that changes in school quality are discrete at boundaries, while changes in neighborhood characteristics are smooth. The difference in mean housing prices located on opposite side of attendance district boundaries can therefore be related to differences in school exam scores only. When restricting the sample to the set of sales located within 0.15 mile of a boundary, Black finds that a 5% increase in primary schools' test scores (approximately one standard deviation) is associated with a 2.1% increase in housing prices, which is half the value of the "naive" OLS estimate. On UK data and using an alternative estimation strategy, Gibbons and Machin (2003; 2006) find an effect of the same order of magnitude for primary schooling. Bayer, Ferreira and McMillan (2003) also find similar results with a discrete choice model instead of a hedonic model. Furthermore, they provide evidence of significant heterogeneity in the marginal willingness of households to pay for school quality.

The paper's focus In respect to this existing literature, our paper innovates in four different directions. Firstly, the data sets that we use are both large and of very high quality: on the one hand, our sample of about 200,000 real estate transactions contains extensive information on almost every single sale that has taken place in the city of Paris during the period 1997-2003 (price, characteristics of the flat, precise location, etc.); this sample of sales can be matched with the exact middle school zoning scheme that was enforced every year. On the other hand, we are able to rank all public and private middle schools according to alternative and standardized measures of school quality.

Secondly, we focus on middle schools, whereas preceding papers have studied the impact of primary schools. We can therefore assess whether the effects found for the primary level also exist in secondary education. Moreover, middle schooling is a key stage in the Parisian educational system, as it will become clear below.

Thirdly, we improve the estimation strategy originally proposed by Black (1999) to deal with the endogeneity of school quality in the housing price equation. We do so by developing a matching framework consisting in the careful comparison of sales across middle school attendance boundaries.

Finally, we exploit a particular feature of the French private school system to investigate whether school choice reduces the capitalization of public school quality in housing prices. Several theoretical models (in particular Epple and Romano(2003), Nechyba (1999; 2000; 2003)) have shown that both open enrollment and private school vouchers tend to decrease residential sorting by lowering housing price differentials. Yet a well developed and almost entirely publicly funded private school system exists in Paris, enrolling about one third of all middle school pupils. This private school system is very close to a system of private school voucher theoretically available for all parents, but rationed in practice, because of restrictions set by the

State on the amount of public funding available for new private schools. This feature of the French system allows us to study how school choice affects the willingness to pay for better public schools.

Results We find that a standard deviation increase in school quality (as measured by exam scores) triggers a 2% increase in housing prices. The size of this effect, which is only a measure of the *average* willingness to pay for a better school, is similar to existing estimates in the US and UK contexts for primary schools (see above) and can explain roughly 7 % of observed inter school zones housing price differentials. In line with the theoretical predictions of school choice models, we also find evidence that the presence of good private schools in certain neighborhoods tends to attenuate the capitalization of public school quality in housing prices, by offering an outside option to parents.

The remainder of this paper is as follows: section (2) gives some background on the French educational system and the school zoning policy in Paris; section (3) describes the data; section (4) presents the results obtained when comparing housing prices across school attendance boundaries; section (5) performs some robustness checks and section (6) discusses how our estimates change when we allow the effect to vary with the availability of local private schools.

2 Middle schooling in Paris

2.1 The French educational system

In France, primary and secondary education is organized as a 12-year curriculum, divided into three stages: children spend 5 years in primary school (age 6 to 10), 4 years in middle school or *Collège* (age 11 to 14) and 3 years in high school or *Lycée* (age 15 to 17). The school year starts early in September and ends early in July[1]. While the curriculum is the same for all students, many specialized sections (including technical studies) and a variety of options can be chosen starting from the third year of middle school. Education is predominantly public, but there exists a large network of private middle schools, enrolling about 13% of all pupils in France. The French territory is divided into 35 local school board called *Académies*, which

[1] For simplicity, we denote each school year using one single number: for instance, "school year 1997" stands for "school year 1997-1998".

are in charge of managing human and financial resources and of implementing the official educational programs produced by the Ministry of Education.

In 1963, it was decided that children in primary and secondary education would be allocated to the different local public schools on the basis of a strict zoning scheme, which did not apply to private schools. According to this allocation rule, every pupil about to enter either primary or middle school is sent to a specific public school depending on his or her address. School zones are usually contiguous and centered around the corresponding establishment.

In the past 40 years, this system has not always been strictly enforced: in the mid-1980s, some *Académies* were authorized to relax the strict school zoning policy for middle schools only, thus giving to parents the opportunity of choosing where to send their children out of a given number of schools.

2.2 Middle school zoning in the *Académie de Paris*

The strict zoning scheme was enforced by the Paris *Académie* between 1963 and the beginning of school year 1987-1988. The zoning policy was then temporarily and partially relaxed until the beginning of school year 1997-1998: in certain *arrondissements* of the city[2], parents of children enrolled in the last year of primary schooling were allowed to rank in order of preference three different middle schools out of all those located in their *arrondissement*. Pupils were then chosen by the school principals on the basis of their school results. This experiment was definitely abandoned after September 1997, when the system moved back to strict zoning[3]: the spatial organization of the school attendance districts has remained remarkably stable ever since [4]. We therefore restrict our study to post-1997 school years.

There are two ways parents can get round the zoning system without actually changing residence. First, they can ask the *Académie* officials for a dispensation that entitles them to send their children to a school located outside their attendance zone. These dispensations can be granted on several grounds: if specific options or "rare" languages (e.g. Russian or Chinese) are not offered in the local school, if the child has a brother or a sister in a different school or if the local school is located much

[2] The city of Paris comprises 20 *arrondissements*, corresponding to administrative subdivisions.

[3] The reason for this reversal is not perfectly clear. Anecdotical evidence seems to point out that part of the explanation comes from the protests of the many parents whose children were allocated to the worse and sometimes distant public middle schools because of their poor academic performance.

[4] Every year, a small number of street sections (corresponding to less than 5% of all residence locations) are reassigned in case the number of students that are about to get enrolled in the first year of middle school exceeds the capacity of the local school

further away from home than a school belonging to an adjacent zone[5]. Every year, dispensations are granted to about 8% of pupils entering middle schools, the rate of rejection being around 40%. While a substantial proportion of these dispensations have true practical justifications, some parents use them to avoid what they perceive as low quality local middle schools. Another way of getting round the zoning system is to use the outside option provided by the extensive network of highly subsidized private middle schools.

2.3 Private middle schools in Paris

Several institutional specificities of the French private school system explain its relatively large size compared with the UK and the US.

In France, there are two types of private schools, depending on their degree of independence with the State and the Ministry of Education: the *"Sous Contrat"* and the *"Hors Contrat"* schools. On the one hand, *"Sous Contrat"* schools are subject to State supervision : they have to follow the same curriculum as public schools (same subjects, same rules...) and appoint qualified teachers who are paid by the State[6]. Part of their expenses are covered by Local Governments. Schools set fees, but only to pay the costs that are not publicly funded, such as optional subjects, in particular religious education. As a result, private education is usually not free, but the annual fee in a private middle school is not very high, between 500 and 2000 euros, plus small additional costs. *"Sous Contrat"* schools usually offer substantial rebates for high-performing pupils from low-income families. On the other hand, *"Hors Contrat"* private schools are not monitored nor financed by the State or Local Governement and can freely design their own curriculum. Nevertheless, the vast majority of private schools, especially in primary and secondary education, are under contract with the State: out of the 71 private middle schools in Paris, only 3 are *"Hors Contrat"*[7].

In any case, private schools are not subject to any zoning scheme and can select their students from anywhere in the city, offering an outside option to those parents who are willing to avoid the constraints of strict school zoning: as a result, about one third of middle school pupils in Paris are enrolled in private schools. This particular feature of the French school system makes it possible to investigate how the housing

[5] This might happen when the local school is not located in the centre of its zone.

[6] There is a specific diploma for private school teaching which is similar to the diploma for public school teaching.

[7] In our analysis, we therefore focus only on Private Schools under contract with the State.

market incorporates public school quality when there exists some degree of school choice.

3 Data and summary statistics

To estimate the impact of school quality on housing sales in Paris, we need data on school zones, school quality, individual property sales as well as information on local sociodemographic characteristics for school years 1997-2003.

3.1 School zones

Data on school zones was provided by the local Education Board of Paris (*Rectorat de l'académie de Paris*). During the period under study, the Board was in charge of drawing the assignment zones of primary and middle schools[8]. We chose to focus on middle schools because it is reasonable to assume that in the specific Parisian context, parents care more about middle school quality than about primary school quality. Primary and middle public schools in Paris are subject to strict zoning since 1997, but the high school regime is different, in the sense that it combines zoning and choice: when entering high school, a pupil can choose any particular school located within a broadly defined zone[9]. There are some very good, over-subscribed public high schools within each zone, as there are less popular, under-subscribed ones. Competition to enter very good institutions is quite vivid, and parents are conscious of the importance of sending their children to a good middle school as a means to increase their chances of being admitted into a good high school. In line with this, sociological work tends to show that parents care more about middle school choice than about primary school choice.

Using the precise location of each property in the housing sales data set (street name and street number), we match each transaction with its assigned middle school for every school year between 1997 and 2003. In order to perform our estimations, we need information on the quality of each of these schools.

[8] As from school year 2005-2006, the task of implementing school zoning was transferred to the Paris City council. This was part of a series of reforms aimed at giving more decision power to local authorities. However, this modification does not affect the 1997-2003 school years.

[9] Paris is divided into 4 different high school zones (West, South, East and North).

3.2 School quality

School quality measures come from two different data sets, both provided by the Statistical department of the French Ministry of Education[10]. Our first source of information on school quality is the *Scolarité* data set which is available every year over the period 1997-2003. It contains individual information on all students in middle and high school public and private education in France. We know each student's age, gender, citizenship, social status of the head of the family, *arrondissement* of residence, school attended in the current (n) and previous ($n-1$) years as well as current and previous educational level. However, this very rich data set suffers from two limitations. First, we do not have access to the students' identifer so we cannot use the panel dimension of the data. Secondly, it does not contain the students' results to examinations. Exam results are collected in a separate data set called *Océan* (which cannot be matched to the first one due to legal reasons). To measure school quality, we exploit pupils' results to the *Diplôme National du Brevet* (*DNB*), the first national anonymous exam taken at the end of *Troisième* (French equivalent of grade 9). Every pupil's score is a combination of continuous assessment (1/2) and a final national examination (1/2) consisting in three parts: Math, French and History & Geography. Each section is scored out of 20. To make sure that our measure of quality is comparable across schools, we use only the national exam component of the *DNB* score: the quality of any particular school is therefore given by the average Math, French and History & Geography score obtained by its 9$^{\text{th}}$ grade pupils. The mean score at the school level was 9.47 out of 20 in school year 2003-2004, with a standard deviation of 1.55. Figure 1 shows how public middle school quality (divided into 6 groups of equal size) is distributed across the city of Paris. Table 3 compares the enrolment and average test scores of public and private middle schools in 2004.

Unfortunately, the database *Océan* is not available for years prior to 2003-2004. This excludes the possibility of computing a medium-run measure of school quality by averaging the exam scores for each school over the period under study. This could be a problem, as property values may respond differently to short-run and medium-run changes in school quality. For instance Kane, Staiger and Samms (2003) have shown that in the US, long run averages of school quality have a significant impact on housing prices, but not year-to-year fluctuations. To deal with this issue, we use the *Scolarité* data set to compute an alternative index of school quality that

[10] *Direction des Études et de la Prospective.*

can be averaged over the period 1998-2004: the percentage of 9$^{\text{th}}$-graders who enter *Seconde générale*, which is the first year of pre-university high school (as opposed to a vocational curriculum). Given the French context, this variable can be considered a good approximation of school quality, since it is closely linked to educational attainment and varies greatly across schools. For each middle school and each year, we know the number of 9$^{\text{th}}$-graders who will be enrolled in pre-university high schools the following year. We then calculate the average value of this index for each school over the 1998-2003 period. Table 1 shows that taken at the school level, the mean fraction of 9$^{\text{th}}$-graders entering *Seconde générale* is 66.8 percent with a standard deviation of 12 percentage points. Unsurprisingly, this measure is highly correlated with the 2004 *DNB* exam score (the Pearson correlation coefficient being equal to 0.83)[11]. Moreover, the relative ranking of middle schools that can be derived from the fraction of students entering pre-university high schools is relatively stable over time: the rank correlation coefficient between this measure of school quality in 1998 and 2003 is 0.74.

The stability of school quality during the period under study is an important feature in the Parisian context: because parents do not have direct access to exam scores (the Ministry of Education being unwilling to publish league tables for middle schools[12]), school quality will tend be capitalized in housing prices only if it is not too volatile across school years. Parents who care about school quality make efforts to gather information on different local public schools through real estate agencies, parents' associations or simply through neighbors, the reliability of which is conditioned by the persistence of school quality in the span of 7 years. Provided that it is the case and that parents have a good idea of the medium-run relative ranking of public schools in their neighborhood, short-run (*DNB* exam score) and medium-run (fraction of 9$^{\text{th}}$-graders entering pre-university high schools averaged over the period 1998-2003) measures of school quality should be appropriate and yield similar results.

[11] It should nonetheless be noted that admission into pre-university high schools is not tied to any particular threshold in the *DNB* exam score, but rather to a global assessment of individual performance by the teachers.

[12] The publication of league tables for public would indeed contradict the fact that parents are not supposed to choose between different public middle schools.

3.3 Housing prices and neighborhood characteristics

Data on property sales come from the Notary Chamber of Paris and the Île-de-France. In France, all property sales have to be registered by a Notary, who collects the realty transfer fee to be paid to the Inland Revenue. The Notary Chamber has gathered the data for Paris and the Île-de-France since the mid-1990's. The data set is almost comprehensive, since it contains between 80 and 90% of all the transactions that took place since 1997[13]. For each transaction, we have information on the price for which the property was sold, along with its detailed characteristics (size, number of bedrooms and bathrooms, parking, date of construction...) and its precise geographical location (Lambert grid coordinates[14]). Our sample is restricted to all arm's-length sales of second hand Parisian flats[15] that took place between September 1997 and August 2004. We further exclude sales in the top and bottom percentiles of housing prices per square meter each year. We are left with a sample of around 200,000 transactions. The mean flat price in our sample is 185,509 euros with a standard deviation of 183,237 euros (table 1). The mean flat size is 52.24 m^2 with a standard deviation of 35.37 m^2. Figure 2 reports the average price per square meter (in 2004 euros) for each of Paris' 80 administrative districts (4 per *arrondissement*) during the period 1997-2004. This map shows that the city is clearly subdivided between the wealthy center and Western side and the less wealthy North-Eastern side. Unsurprisingly, the distribution of housing prices across the city almost perfectly matches the distribution of school quality displayed in figure 1, although this feature cannot receive any causal interpretation.

Information on neighborhood characteristics comes from the 1999 French Census. Our controls include the average number of persons per flat, the proportion of low-rent apartments, the proportion of owners, of single-parent families, of foreigners, of graduates as well as occupation (self-employed, managers, employed, manual workers) and unemployment. An important feature of our census controls is that, when possible, we calculate these variables not only for all households in a given

[13] In the early 1990s, when the Notary Chamber started to collect the data, not all Notary offices succeeded in transmitting their own data. However, the coverage rate has continuously improved since.

[14] The geographical precision of a particular location is of about 0.06 mile.

[15] We dropped newly-built property sales because their price differs greatly from the price of second-hand sales and because new properties represent only a very small share of all property sold in Paris. For the same reasons, we excluded the few houses contained in the sample. We also dropped transactions when the price or the number of rooms was missing. We finally decided to exclude the Islands (Île Saint Louis and Île de la Cité), because of their very specific location and pattern of housing prices.

area, but also for the subpopulation of families (*i.e.* households with at least one child).

Descriptive statistics on census characteristics are reported in table 2: 6% percent of households are female-headed families. Among the individuals aged 15 or more who have finished their studies, 39% hold a graduate degree. 6.2% of Parisians are manual workers, 4% are self-employed, 22% are managers, 15% are employees, 14% hold an intermediary occupation and 11.6% are unemployed. These summary statistics show that the city of Paris concentrates a relatively wealthy population.

4 Comparing transactions across school attendance boundaries

In this section, we present the empirical framework used for the analysis of the effect of school quality on housing prices. We start by presenting the naive estimates obtained with a simple hedonic model; we then turn to the "natural experiment" approach that consists in comparing sales across common middle school attendance boundaries.

4.1 Naive estimates of the impact of school quality on housing prices

The hedonic function model The standard hedonic housing price function describes the sale price as a function of the location of the flat and its intrinsic characteristics (floor, age of the building, number of rooms, number of bathrooms, etc.). The coefficient that is associated with each characteristic is supposed to measure the marginal purchaser's willingness to pay for that specific characteristic. The basic hedonic function for housing prices can therefore be written:

$$\ln p_{i,c,s,t} = \alpha + \beta.z_s + X'_{i,c,s,t}\gamma + N'_{i,c,s}\phi + L'_t\theta + \epsilon_{i,c,s,t} \qquad (1)$$

where $p_{i,c,s,t}$ is the price of sale i, located in census block c, in school attendance zone s during school year t; z_s is the quality index of school s, $X_{i,c,s,t}$ the vector of flat i's characteristics, $N_{i,c,s}$ the vector of neighborhood characteristics (at the census block level, during year 1999) and L_t a vector of time dummies (year and quarter). The OLS estimate of the parameter of interest β is supposed to measure the marginal willingness to pay for a school of better quality.

As pointed out by Black (1999), this methodology will produce upwardly biased estimates if there are unobserved neighborhood characteristics varying both across and within school districts that are correlated with school quality and are likely to influence housing prices.

Results The first two columns of table 4 show the results of estimating equation (1), with and without controlling for the characteristics of the census block. Column (1) indicates that a naive estimation of the impact of school quality on housing prices yields a strong apparent impact: for every standard deviation increase in the local school's average exam score, the housing price per square meter goes up by 16.7%. Column (2) shows that the coefficient falls to 2.7% when one adds controls for the detailed neighborhood characteristics (including controls for the subpopulation of families).

A way of solving the endogeneity problem of the school quality variable in the housing price equation is to compare sales that can be assumed to share the same unobserved characteristics corresponding to "neighborhood effects". To infer the value parents place on school quality, Black (1999) uses a first difference approach: she compares the prices of sales located on both sides of a common attendance district boundary with the difference in the quality of their respective schools.

In the following subsections, we start by applying Black's original estimation strategy to the Paris school zoning system. In order to correct for some of the potential pitfalls associated with her method, we subsequently develop a slightly more sophisticated version, based on a direct matching approach.

4.2 Including boundary dummies

The model Black's estimation strategy consists in replacing the vector of observed characteristics in the traditional hedonic hedonic equation by a full set of boundary dummies indicating sales that share (on either side) a common attendance district boundary:

$$\ln p_{i,s,b,t} = \alpha + \beta.z_s + X'_{i,s,b,t}\gamma + K'_b\phi + L'_t\theta + \epsilon_{i,s,b,t} \qquad (2)$$

where $p_{i,s,b,t}$ is the price of sale i, located in school attendance zone s, next to boundary b, during school year t; z_s is the quality index of school s (middle school *DNB* exam score in 2004), $X_{i,s,b,t}$ the vector of flat i's characteristics, K_b a full set of boundary dummies (only the boundaries that have remained unchanged during the 1997-2003 period are taken into account) and L_t a set of time dummies (year and quarter).

We use two different estimation samples, based on the distance of each observation to the boundary of its school attendance boundary: the first sample includes all the sales that are located within 0.20 mile of a boundary; the second sample includes the sales located within 0.15 mile of a boundary. Figure 3 shows the location of the sample of sales that are located within 0.15 mile of a boundary between 1997 and 2003.

Results Columns (3) and (4) of table 4 show the results of the estimation of the naive hedonic equation (1), for sales located within close distance from a boundary, when school quality is approximated by the average *DNB* exam score in 2004. Restricting the estimation sample does not change the coefficient on middle school exam scores: it remains close to that obtained with the full sample, 2.8% and 2.9% respectively for sales located within 0.20 and 0.15 mile of the nearest boundary.

Columns (5) and (6) of table 4 show the results of the estimation of equation (2), which includes boundary fixed effects. The crucial assumption here is that houses located on opposite sides of a boundary and within a given distance (0.20 or 0.15 mile) share the same unobserved neighborhood characteristics except for the quality of their local school. The coefficients on middle school exam scores are only slightly smaller than the coefficients estimated using the naive hedonic housing price regression: 2.3% for the 0.20 mile range (significant at the 1% level), 2.0% for the 0.15 mile range (significant at the 1% level).

4.3 Matching sales across boundaries

The estimation framework developed above relies on two relatively strong assumptions that can be considered unrealistic. First, the characteristics of flats are supposed to have the same impact on prices throughout the city of Paris. Yet there is some evidence[16] that the influence of such features as the age of a building or the number of rooms in a flat is likely to differ noticeably across the various areas of the city. The second implicit assumption underlying the inclusion of boundary fixed effects is that whatever the length of a particular border, it must be the case that houses located on both sides but at opposite ends of the line segment of a common boundary share the same unobservable characteristics. This may not be true in the case of very long boundaries passing through very different neighborhoods (as it

[16] See Laferrère (2005).

appears to be the case in Paris, judging from figure 3).

In order to circumvent these two potential limitations, we adopt a matching framework that compares each transaction with a carefully constructed counterfactual transaction. First, we correct the prices of all sales for *arrondissements*-specific flat characteristics effects. The prices of all sales are homogenized in the sense that there are now expressed in terms of the typical flat's characteristics[17]. Appendix 1 explains the exact steps that we have followed to compute these so-called "hedonic" prices.

We also modify the methodology developed by Black (1999) in a second direction: rather than including a full set of boundary fixed effects to estimate equation (4), we use a matching approach, which consists in restricting a transaction's comparison group to those sales that are located on the other side of the school boundary and within a given radius of that specific transaction. The principles of the approach we have followed are explained below and figure 4 shows how it is applies for a particular set of sale.

The model Any housing sale has four components: its location (indicated by its geographic coordinates x and y), the school year during which it has taken place, its hedonic price per square meter and the middle school zone it belongs to. We define a transaction's neighborhood as the area comprised within 0.20 (or 0.15) mile of that particular transaction[18].

We suppose that within each of the 20 *arrondissements* of Paris, the price per square meter (taken in log) of a transaction located in neighborhood n, belonging to school zone s during school year t can be written as the sum of a time-invariant neighborhood fixed effect θ_n, a time-invariant middle school fixed effet θ_s, a school year fixed effect θ_t and an error term $\epsilon_{n,s,t}$:

$$\ln p_{n,s,t} = \theta_n + \theta_s + \theta_t + \epsilon_{n,s,t} \quad (3)$$

where $\epsilon_{n,s,t}$ is clustered by school zone.

We make the assumption that the middle school fixed effect θ_s is a linear function

[17] The typical flat has two rooms, belongs to a building constructed between 1850 and 1913, is located on the ground floor, with one bathroom and average size rooms, without a maid's room or a garage and was sold during the fourth quarter.

[18] A transaction's neighborhood is thus a relative notion.

of our school quality index z_s and an error term η_s:

$$\theta_s = \alpha + \beta.z_s + \eta_s$$

The coefficient β, which measures the impact of school quality on housing prices, is our parameter of interest.

We call "reference sales" all housing transactions located within a distance of 0.20 (or 0.15) mile of a school attendance boundary and belonging to an area which has not been reassigned to a different school between 1997 and 2003. A reference sale that took place in neighborhood n, school zone s during school year t is denoted $H_{n,s,t}$. Every reference sale is associated to a fictive "counterfactual" sale, denoted $\widetilde{H}_{n,s',t}$. The price of this counterfactual sale is supposed to measure the amount for which the reference transaction would have been sold, had it been located in school zone s' rather than in school zone s, everything else being equal. The fixed effect associated with the counterfactual sale's middle school s' is a linear function of $z_{s'}$, the quality index of school s':

$$\theta_{s'} = \alpha + \beta.z_{s'} + \eta_{s'}$$

The price $\widetilde{p}_{n,s',t}$ of the counterfactual transaction $\widetilde{H}$ cannot be observed and has to be estimated. We do so by calculating the weighted geometric mean of the prices of all the transactions $h_{i,n,s',t}$ that took place in same neighborhood n and during the same school year t as the reference sale H, but belong to a different school zone s'. Transactions h_i are weighted by the inverse of their distance d_i to the reference transaction H, in order to give more importance to the sales that are located nearby relatively to the more distant ones. Hence, the estimated price per square meter $\hat{p}_{n,s',t}$ of the counterfactual sale $\widetilde{H}$ is calculated as:

$$\ln \hat{p}_{n,s',t} = \frac{1}{\sum_{i \in I} \frac{1}{d_i}} \sum_{i \in I} \frac{1}{d_i} \ln p_{i,n,s',t}$$

where (denoting $x(.)$ and $y(.)$ as the geographic coordinates of transactions H and h_i, expressed in the mile unit):

$$I = \left\{ h_{i,n,s',t} : \sqrt{\left(x(h_{i,n,s',t}) - x(H_{n,s,t})\right)^2 + \left(y(h_{i,n,s',t}) - y(H_{n,s,t})\right)^2} \leqslant d, d \in \{0.15, 0.20\}, s' \neq s \right\}$$

The estimated price $\hat{p}_{n,s',t}$ of counterfactual sale $\widetilde{H}$ is equal to the true counterfactual price $\widetilde{p}_{n,s',t}$ plus a measurement error $\nu_{n,s',t}$:

$$\ln \hat{p}_{n,s',t} = \ln \widetilde{p}_{n,s',t} + \nu_{n,s',t}$$

Note that a reference sale which is located next to different school zone boundaries[19] can have several counterfactuals. In this case, there will be as many counterfactuals as there are school attendance zones within a 0.20 (or 0.15) mile radius of the reference transaction.

The identification of the "middle school effect" relies on the crucial assumption that every couple (reference sale, counterfactual sale) shares the same neighborhood fixed effect and only differs through school attendance zones:

$$\ln \widetilde{p}_{n,s',t} = \theta_n + \theta_{s'} + \theta_t + \epsilon_{n,s',t}$$

Under this hypothesis, the housing price differential between the reference and the estimated counterfactual sales can be written:

$$\ln \widetilde{p}_{n,s,t} - \ln \hat{p}_{n,s',t} = \ln \widetilde{p}_{n,s,t} - \ln \widetilde{p}_{n,s',t} + \nu_{n,s',t} = \beta(z_s - \widetilde{z}_{s'}) + \mu_{n,s,s',t}$$

where $\mu_{n,s,s',t} = \eta_s - \eta_{s'} + \epsilon_{n,s,t} - \epsilon_{n,s',t} + \nu_{n,s',t}$ is an independent error term, clustered by school attendance boundary (s, s').

Parameter β can be estimated by running an OLS regression of the price differential between the reference and the counterfactual sale on the corresponding school quality differential. Note that the measurement error induced by the estimation of the counterfactual sale's price only affects the dependent variable and hence will not bias the regression coefficient $\hat{\beta}$.

Results Let us first compare the naive regressions using hedonic prices (the first two columns of table 5) with our previous naive estimates (Columns (3) and (4) of table 4). Reassuringly, results are very similar. Thus the existence of potentially different effects of flat characteristics across arrondissements does not seem to be the driving force explaining the magnitude of the naive estimates of the impact of school quality on housing prices.

Columns (3) and (4) of table 5 show the results of regressing the difference in prices across boundaries on the corresponding difference in school quality, calculated using the matching approach described above. The results are very close to those

[19] This will be the case for sales located in one of the "corners" of a school attendance zone.

obtained in table 4. In addition, the estimates obtained for both distances (0.20 and 0.15 mile) are almost exactly equal, so the change in the sample size does not seem to be driving the results. A point increase in the average exam score is estimated to raise the price per square meter by about 1.9% when we restrict the observations to those located within 0.15 mile of a boundary.

4.4 How large is the effect?

Our estimates of the impact of middle school quality on housing prices in Paris are of the same order of magnitude as existing estimates for primary schools in other countries: similar to the 2.1% effect for every standard deviation increase in school quality for Boston suburbs (Black (1999)); slightly smaller than the 3.7% effect obtained by Gibbons and Machin (2003) for the UK.

In order to interpret these results, it is important to get some sense of the magnitude of the effect. Firstly, we calculate that other things being equal, moving from the worst to the best public middle school (which corresponds to 4.8 times the standard error of the average *DNB* exam score at the school level) would imply a price premium of 9% (16,700 euros for the average flat price). Secondly, we estimate the fraction of the housing price differential between school zones that can be explained by differences in school quality. In order to do so, we calculate the observed difference in the average flat price (taken in logs) between each adjacent pair of school zones and relate it to the flat price differential predicted by the corresponding difference in school quality. We find that the difference in school quality explains roughly 7% of the observed difference in housing prices between adjacent school zones.

These calculations indicate that although school quality plays a non negligible role in the formation of housing prices, it is not the main driving force in the real estate market. However, this result does not imply that the way school quality determines parents' residential location should be neglected when school zone policies are designed.

5 Robustness checks

In this section, we perform several robustness checks. First, we test whether the results are sensitive to our measure of school quality. Secondly, we investigate the validity of the identifying assumption that flats located on either sides of attendance boundaries are in the same neighborhoods.

5.1 Results' sensivity to the measure of school quality

We start by testing the sensivity of the results to our definition of school quality. The main concern is that the average *DNB* exam score of 9$^{\text{th}}$-graders, which is measured at the end of the period under study, might not be a good measure of the middle-run school quality of schools during the period 1997 to 2003. To deal with this potential problem, we have investigated whether using an alternative mid-run index of school quality affects our results. As explained earlier, the fraction of 9$^{\text{th}}$-graders who are admitted into *Seconde générale* is a relatively good candidate: available since 1997, this index can be averaged at the school level over the period 1998-2004. Using this measure of school quality, we perform the same regressions as previously. Table 6 displays the naive estimates and the boundary fixed effects estimates, while table 7 shows the matching across boundaries estimates. The results are remarkably similar to those obtained when school quality is evaluated by the *DNB* exam score. The impact of a standard deviation increase in the proportion of 9$^{\text{th}}$graders admitted into *Seconde générale* on housing prices is significantly positive: about 2% when we directly match sales across school attendance boundaries. Hence, our results do not appear to be dependent on the particular definition of school quality that we have adopted in this paper.

5.2 The validity of the identifying assumption

Our estimation strategy relies on the assumption that flats located on either sides of attendance boundaries share the same neighborhood characteristics, so that their price differential is on average purely attributable to the difference in the quality of their respective schools. Yet this hypothesis might be violated if a particular side of the boundary display certain characteristics (e.g. more housing units with garages) which are valued by buyers, independently from the quality of the local school. If such characteristics tend to attract wealthier people, whose children educational attainment will mechanically drive up the quality of the local school, then one might obtain biased estimates of the effect of school quality on housing prices by comparing sales across attendance district boundaries.

First, we test whether the observable characteristics of flats, such as the age of building, the number and size of rooms, the number of bathrooms, the number of garages or the presence of a maid's room, are similarly distributed on either sides of boundaries. Table 8 compares the characteristics of flats located on either side of school attendance boundaries. Although the means are generally statistically different because of the large size of the sample, the distribution of characteristics on the "good" side and the "bad" side of boundaries are pretty similar. Moreover, the numbers show no obvious pattern that could explain that flats located on the "good" side of the boundary are more desirable than flats located on the "bad" side. In addition to these comparisons, we run the regressions without controlling for flat characteristics, and find the same results[20]. Thus our findings do not seem to be driven by differences in observable flat characteristics. One might still argue that some unobservable flat characteristics (e.g. one side gets more sun than the other) might bias the results, but there is no particular reason that such characteristics should be distributed differently across school zones.

A more serious issue is that household sorting might occur at boundaries, even if the houses are the same. Several papers on US data (Bayer, Ferreira and McMillan (2003) and (2004), Kane, Staiger and Riegg (2005)) have shown that not only school quality, but also several sociodemographic characteristics, such as income, could be discontinuous at boundaries. Estimations using boundaries fixed effects to control for neighborhood unobservables might then tend to overestimate the causal effect of school quality on housing prices.

To investigate the discontinuity at boundaries, we regress the log of housing prices, and several socioeconomic characteristics on a fourth order polynomial equation in distance to the boundary and a dummy variable indicating that transaction i is located on the "good side" of the boundary:

$$Logprice_i = \beta_0 + \beta_1 dist_i + \beta_2 dist_i^2 + \beta_3 dist_i^3 + \beta_4 dist_i^4 + \beta_5 Goodside_i + \epsilon_i \quad (4)$$

Where *dist* is an indicator of the distance to the frontier for sales located within 1000 feet of a boundary, and *Goodside* is a dummy which is equal to one if the sale is located in the zone of the school with the highest mean test score. We restrict our sample to the boundaries where the mean test score differential between schools is at least 0.8 (corresponding roughly to a 1/2 standard deviation at the school level, or a 1/4 standard deviation at the pupil level), to be able to visualize the discontinuity[21]. We graph the variables and the fitted form, and check wether the discontinuity is significant at the boundary. The results for the housing prices per square meter

[20] Since the results are basically the same than the initial regressions, we do not show them, but the table is available upon request.

[21] The mean school test score differential in this restricted sample is 1.8.

and the log of "hedonic" housing prices (already corrected to account for differences in flat characteristics) are displayed in figure 5 and present a clear and significant discontinuity at the boundary.

We now turn to sociodemographic characteristics. We do not know the average income in the neighborhood but we do have information on the professional status of individuals, so we use it as a proxy for well-to-do and less well-to-do neighborhoods. Some variables, such as the proportion of public housing in the neighborhood do not exhibit any discontinuity at the boundary (see figure 6).

Other variables, such as the proportion of households whose head is a manual worker and the proportion of households whose head is a manager do exhibit a significant discontinuity at the boundary (see figures 7 and 8). However, when we look more closely at the data, we see that the discontinuity is driven by families. The details of professional occupations is available by household size, so we use large households (3 individuals or more), as a proxy for families, and compare them to small households (1 or 2 individuals), used as a proxy for households without children[22].

The proportion of large households whose head is a manual worker severely shrinks on the good side of the boundary, and there is a clear discontinuous increase in the proportion of large households whose head is a manager. The discontinuity is much smaller for small households, and even insignificant for managers. Our interpretation is that the discontinuity in neighborhood characteristics occurring at boundaries is driven primarily by differences in school quality. Parents who care about schools tends to be high skilled workers, but high skilled workers without children do not seem to be ready to pay more to live in good school zones. We therefore are pretty confident that the effect measured by our regression will capture the effect of the school quality differential, that may generate sorting among families, but not a more general sorting effect.

Finally, private school choice provides us with an additional way of testing whether housing prices reflect sorting at boundaries. If we manage to show that the proximity of private schools in the neighborhood *lowers* the price premium for a house in a better public school zone, then we would conclude that this price premium is not driven by sorting.

[22]We actually have the information for families with children, but this figure is difficult to reconcile with the total number of households, so in order to be able to distinguish between families and other households we use the size of the household as a proxy. Results using families with children are very close to those obtained on large households.

6 The mitigating effect of private schools

Theoretical predictions of school choice models In section 2.2, we mentioned that parents who wish to avoid sending their children to the local public school but are not willing to change residence can either ask for a dispensation or send their children to a private school. Therefore, the presence of a good network of private schools should lower the price premium that parents are ready to pay for a flat located in a good local public school zone, or at least set an upper bound to that premium.

More precisely, the combination of strict public school zoning and private school choice in the Parisian context allows us to test some of the theoretical predictions of a general equilibrium model developed and calibrated by Nechyba (1999; 2000; 2003). This model of school finance includes multiple school districts (either state of locally financed), multiple neighborhoods within school districts with different housing qualities, and local public schools at district level that must admit every pupil in the district along with private schools that can choose their pupils. Using this particular set up, Nechyba studies the consequences of introducing private school choice on both school and residential stratification. The model is too complex to yield analytic solutions, and it is calibrated to predict the effect of alternative educational policies. In the benchmark equilibrium without private schools, peer effects yields substantial residential sorting[23]. Allowing for private school choice decreases residential stratification, while increasing peer stratification in schools [24]. The decrease in residential stratification is reflected in housing prices. In this section, we test whether the presence of private schools near the place of residence mitigates the effect of local public school quality on housing prices.

In Paris, 30% of middle school pupils were enrolled in a private school in 2004. Contrary to public schools, private schools are not evenly distributed throughout the city: the development of existing private schools and the creation of new ones is in fact limited by the State which controls the amount of resources that are allocated to the private sector. The scores obtained by pupils enrolled in private middle schools at the *DNB* exam are about one standard deviation above those obtained by pupils enrolled in public middle schools (see figure 9). Private schools thus offer

[23] There is only incomplete stratification because of the exogenous heterogeneity in neighborhoods the housing stock is fixed. These assumptions seem quite reasonable when applied to make prediction for a city like Paris.

[24] The overall effect on the quality of public schools varies with the type of school finance and whether per public spending increases enough or not to compensate for the decrease in peer quality.

a rather good outside option for parents reluctant to sending their children to a poor-performing local public school.

Under the assumption that parents don't usually want to enroll their children in a school which is located too far away from home, the impact of public school quality should depend on the density of the local private school offer. In residential areas where many private schools are available, we expect the quality of public schools to be less capitalized in housing prices than in areas where there are few.

The estimation strategy To test this hypothesis, we construct an index of the private middle school density in a given area. We want to take into account both the size (in terms of enrollment) and the proximity of private middle schools. Total enrollment in an important factor to take into account, since, unlike the size of public middle schools, it can vary a lot across private schools.

The index D_i is constructed as follows. For each transaction i, we calculate the average of the inverse of its distance to every private middle school in Paris, squared and weighted by the total enrollment of each of these private schools:

$$D_i = \frac{1}{Z} \sum_{j=1}^{N_j} z_j \frac{1}{d_{i,j}^2} \forall i$$

where Z denotes the total enrollment in all Parisian private middle schools, z_j the enrollment in private middle school j in 2004 and $d_{i,j}$ the distance between transaction i and private school j. The higher the value of this index, the higher the density of private schools in the transaction's neighborhood. Using the inverse of the distance squared allows us to give much more weight to the closest private schools.

We then split our indicator into four quarters, from the lowest density to the highest. Figure 10 shows the corresponding areas on the map of Paris. Reassuringly, our indicator of private school availability is well distributed among neighborhoods, and does not cut Paris into four geographically distinct zones, a feature which could bias our estimates. In the lower quartile, we find that the supply of private middle schools is scarce in some neighborhoods located in the center of the city as well as areas in the South East. At the other end of the distribution, the upper quartile is associated with two different types of neighborhoods. On the one hand, in the North-Eastern parts of Paris, parents who are confronted with low performing public schools have the opportunity of sending their children to one of the many private schools located in the area. On the other hand, in some of the South-Western wealthy neighborhoods where public middle schools are of high quality, the network of private school is also very dense. In both cases, we expect the willingness to pay of parents for a good public schools to be lower than in the neighborhoods where parents do not have this "outside option".

To estimate the mitigating effect of private school choice, we allow the parameter of interest to vary with the quartile of private middle school density :

$$\ln \widetilde{p}_{s,t} - \ln \hat{p}_{s',t} = \beta_1 Q_1.\Delta_{z_s,\widetilde{z}_{s'}} + \beta_2 Q_2.\Delta_{z_s,\widetilde{z}_{s'}} + \beta_3 Q_3.\Delta_{z_s,\widetilde{z}_{s'}} + \beta_4 Q_4.\Delta_{z_s,\widetilde{z}_{s'}} + \mu_{s,s',t}$$

where $\ln \widetilde{p}_{s,t} - \ln \hat{p}_{s',t}$ denotes the price differential between sale s and its counterfactual in time t, $\Delta_{z_s,\widetilde{z}_{s'}}$ the difference in school quality between the public school s assigned to the transaction and the public school s' of the counterfactual sale, Q_j ($j \in 1,2,3,4$) a dummy variable indicating the quartile of private school density in the sale's neighborhood and $\mu_{s,s',t}$ an independent term clustered at the school attendance zone.

Results Table 9 shows the results using the matching approach, for sales located within 0.20 and 0.15 mile of a boundary, using the *DNB* exam score in 2004 as our measure of public school quality. Columns (2) to (5) indicate that as the density of private middle schools increases, the impact of public school quality on housing prices becomes smaller: while for the lower quartile of private school density, the coefficient on public school quality lies between 3.9% (0.15 mile) and 3.4% (0.20 mile), it becomes small (0.8-0.9%) and insignificant for the upper quartile (column (5)).

The results are very similar when we use the percent of pupils going into *Seconde générale* instead of the average *DNB* exam score: table 10 indicates that the impact of public middle school quality declines as the density of private school increases, moving from 4.4-4.8% for the lower quartile to 0.4-0.8% for the upper quartile.

On the whole, these results support the theoretical predictions mentioned above by showing that the effect of public school quality on housing prices is heterogeneous: when parents have the possibility of sending their children to a good private school, then housing prices do not seem to depend on the quality of the local public middle school; on the contrary, when there is no good private school available in the neighborhood, then the local public middle school quality appears to be capitalized into housing prices.

Given these estimations, we can perform the simple exercise that consists in comparing the cost of attending a private school with the cost of moving into a better public school zone. Given that the average private school tuition fee is 1,500 euros per year Paris, four years of private middle schooling costs about 6,000 euros to the parents. In areas where good private schools are available, the quality of the private school is higher than that of the assigned public school by roughly 3 points on the *DNB* exam score (2 standard deviations), so the housing price premium to be paid in other areas for a similar increase in public school quality would be 4%, that is roughly 7,400 euros for the average flat price. This value being comparable to the individual cost of a private four-year individual tuition fee, it appears that private schooling might be more attractive to single-child families than to families with two or more kids, which have higher incentives to locate near good public schools.

7 Conclusion

Using an almost exhaustive data set on housing sales in Paris over the period 1997 to 2003, we find that the quality of public schools has a significant impact on housing prices by comparing price and school quality differentials across school attendance boundaries: a standard deviation increase in the average exam score at the school level raises prices by about 2%. Additionally, our estimates implies that roughly 7% of inter school zones price differentials are explained by school quality differentials.

We also find evidence that, following the prediction of theoretical models of school choice, private schools tend to attenuate the capitalization of public school quality in housing prices, by offering an outside option to parents.

References

Bayer, P., Ferreira, F., and McMillan, R. (2003). "An Unified Framework for Measuring Preferences for School and Neighborhoods". Yale University Economic Growth Center Discussion Paper No 872.

Bayer, P., Ferreira, F., and McMillan, R. (2004). "Tiebout Sorting, Social Multipliers and the Demand for School Quality". NBER Working Paper No 10871.

Benabou, R. (1993). "Workings of a City: Location, Education and production". *Quarterly Journal of Economics*, Vol.108(3).

Black, S. (1999). "Do Better Schools Matter? Parental Valuation of Elementary Education". *Quarterly Journal of Economics*, Vol.114(2).

Epple, D. and Romano, R. (2003). "Neighborhood Schools, Choice, and the distribution of Education Benefits". In Hoxby, C., editor, *The Economics of School Choice*. University of Chicago Press.

Fernandez, R. and Rogerson, R. (1996). "Income Distribution, Communities and the Quality of Public Education". *Quarterly Journal of Economics*, Vol.111(1).

Gibbons, S. and Machin, S. (2003). "Valuing English Primary Schools". *Journal of Urban Economics*, 53.

Gibbons, S. and Machin, S. (2006). "Paying for Primary Schools: Supply Constraints, School Popularity or Congestion". *The Economic Journal*, Vol.116(510).

Kane, T. J., Staiger, D. O., and Riegg, S. K. (2005). "School Quality, Neighborhoods and Housing Prices: The impact of School Desegregation". NBER Working Paper No 11347.

Kane, T. J., Staiger, D. O., and Samms, G. (2003). "School Accountability Ratings and Housing Values". *Brookings-Wharton Papers on Urban Affairs*.

Laferrère, A. (2005). "Hedonic Housing Price Indexes: the French Experience". *BIS Papers*, 21.

Nechyba, T. (1999). "School Finance Induced Migration and Stratification Patterns: The Impact of Private School Vouchers". *Journal of Public Economic Theory*, Vol.1(1).

Nechyba, T. (2000). "Mobility, Targeting and Private School Vouchers". *American Economic Review*, Vol.90(1).

Nechyba, T. (2003). "Introducing School Choice into Multi-District School Systems". In Hoxby, C., editor, *The Economics of School Choice*. University of Chicago Press.

Appendix: The computation of "hedonic" prices

To apply our estimation strategy, we need to compare the evolution of housing prices in adjacent zones. Unfortunately, it not possible to compare prices of transactions directly, as two flats rarely have the exact same characteristics. We do not have enough transactions taking place in a given year and within a given zone to be able to match housing units that share the same characteristics.

This is the reason why we adopt a hedonic method to construct series of comparable housing prices, following the methodology defined by the Insee, the French statistical agency, to compute a hedonic price index[25]. We first define large zones (18 for Paris[26]) where prices are fairly homogenous, and regress the log of housing prices on observable flat characteristics, inside each zone. We then define a "typical flat" and use the coefficient estimated with the hedonic regression to compute a "typical flat equivalent price" for each sale. The hedonic model is the following:

$$\ln p_i = \ln p_0 + \sum_a \alpha_a y_{a,i} + \sum_t \beta_t t_{t,i} + \sum_k \gamma_k x_{k,i} + \epsilon_i$$

with p_i, the price per square meter of sale i, $y_{a,i}$ and $t_{t,i}$ time dummies indicating respectively the year and quarter in which the flat was sold; $x_{k,i}$ is a full set of dummies indicating sale i's characteristics.

The regression is run separately for each zone, to allow the coefficients across the different areas of Paris. We show in table 11 an example of the hedonic regression for zone 1 (which groups *arrondissements* 1 to 4, corresponding to the right bank center of Paris). The "typical flat" is a two-room flat located on the ground floor, with one bathroom and medium-sized rooms without a parking, a terrace, a balcony nor a maid's room, constructed between 1850 and 1914 and sold during the fourth quarter of the year. This method gives us homogenized prices that can be compared from one sale to another, as the difference in prices that are related to structural differences in observable characteristics of the properties are corrected for. As for the hedonic model usually employed in the literature, in order to use the "typical apartment equivalent prices" in the matching estimation strategy, we need to assume that school and neighborhood fixed effects are uncorrelated with the structural characteristics

[25] The methodology is described in details in Laferrère (2005).

[26] We used geographic and price criteria to define the 18 zones, from the 80 administrative districts of Paris. Zones do not necessarily correspond to the 20 *arrondissements* of Paris.

of the properties. We therefore suppose that the "true model" is:

$$\ln p_i = \ln p_0 + \sum_a \alpha_a y_{a,i} + \sum_t \beta_t t_{t,i} + \sum_k \gamma_k x_{k,i} + \sum_j \delta_j q_{j,i} + \lambda c_i + \epsilon_i$$

where $q_{j,i}$ is the set of neighborhood dummies and c_i is an indicator of the local school quality.

In this model, school and neighborhood characteristics might be correlated with unobservable variables, but not with the flat's structural characteristics. Under this assumption, hedonic coefficients will not be biased and the price and neighborhood effects can be estimated later when matching sales across boundaries.

This first hedonic regression stage is necessary in our matching estimation strategy. If we wanted to control directly for the structural characteristics of the dwelling in the matching model, we would have to include the set of all the interactions between the control variables for each zone, because the matching model is estimated in differences. In addition, the coefficients would be identified only for the sample of houses close to school attendance boundaries.

Figure 1: *Quality of public middle schools as measured by the average Diplôme National du Brevet exam score in 2004.*

Figure 2: *Average price per square meter in the 80 administrative districts of the city of Paris (in 2004 euros). Period 1997-2004*

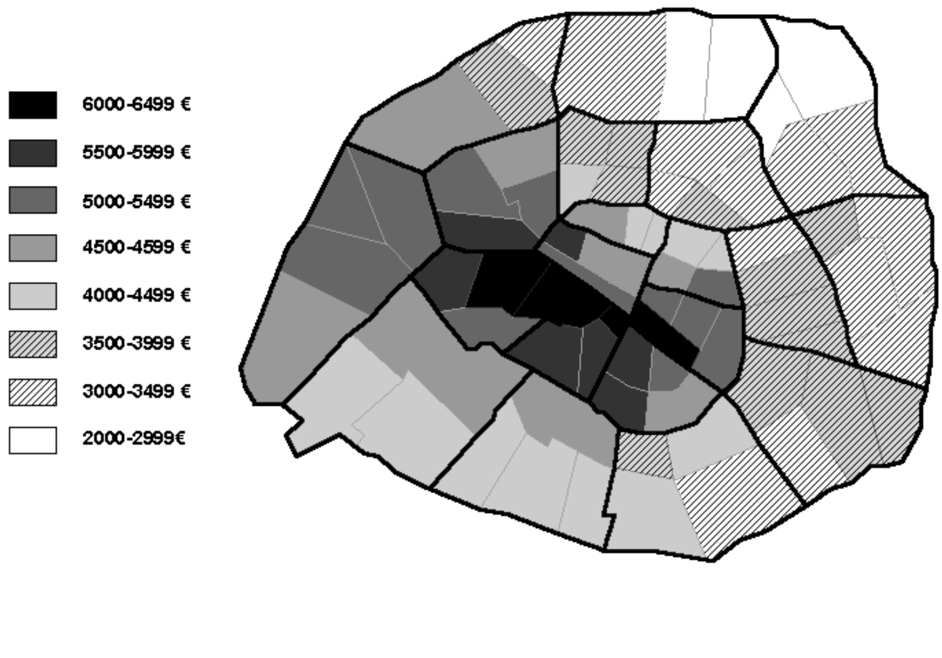

Figure 3: *Sales located within 0.15 mile of a boundary. School years 1997-2003.*

Figure 4: *Matching sales across boundaries: an example*

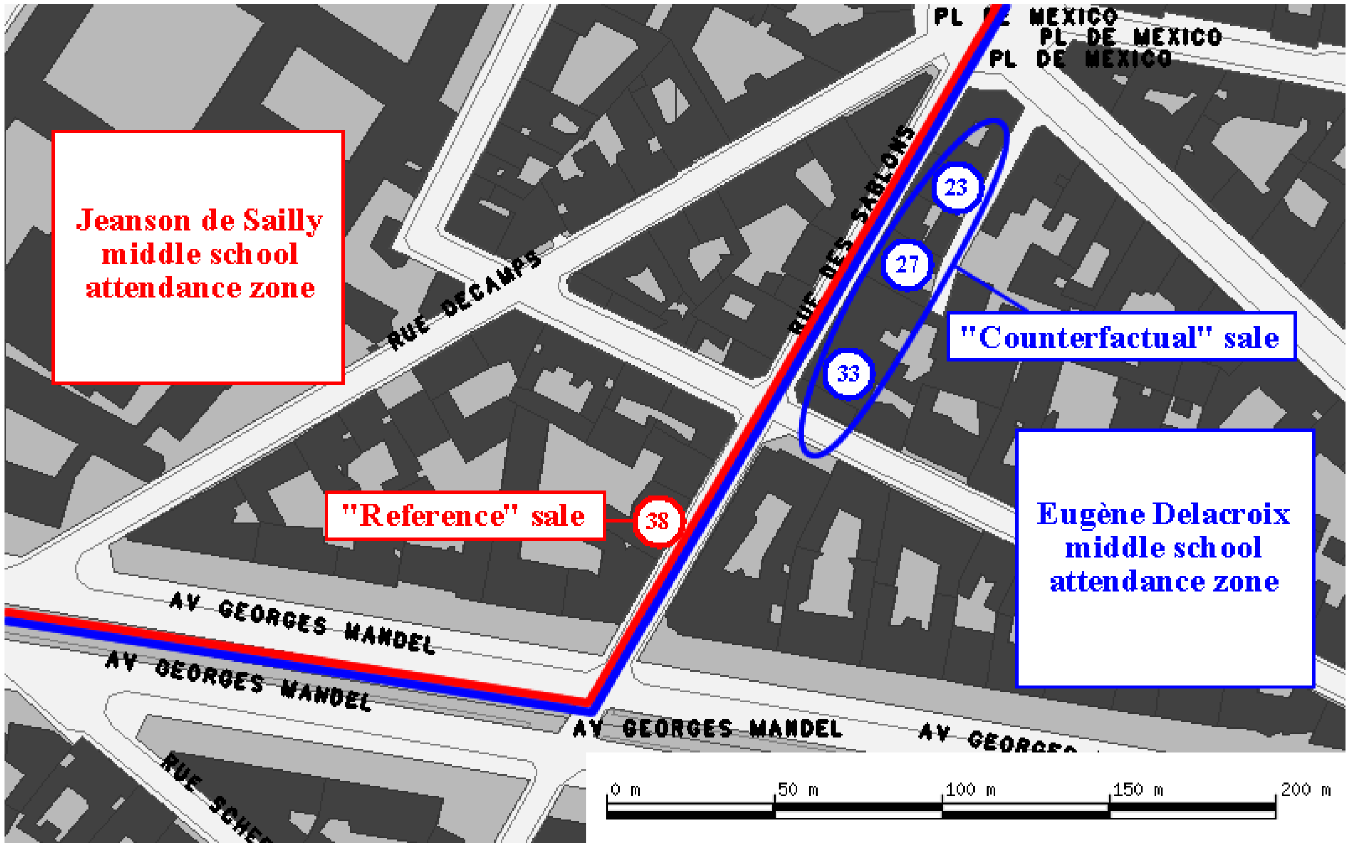

Figure 5: *The discontinuity in housing prices at school boundaries*

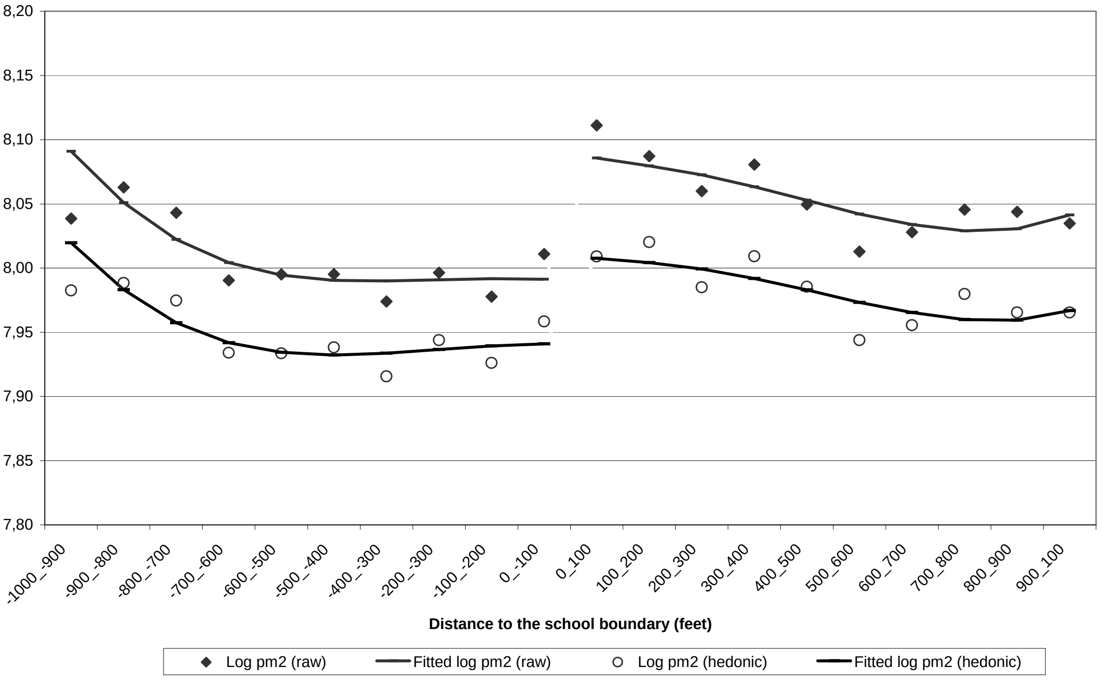

Figure 6: *The evolution of the proportion of public housing in the neighborhood near school boundaries*

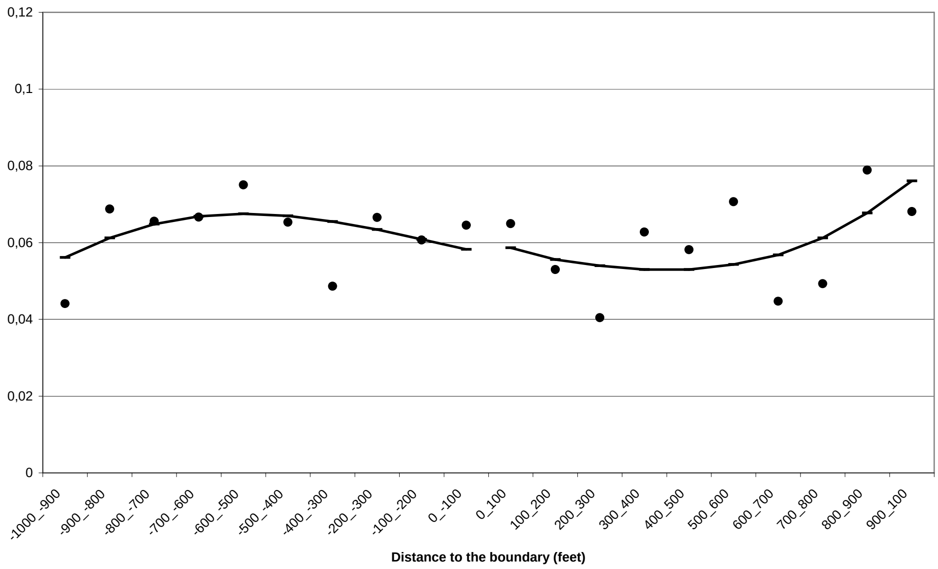

Figure 7: *The evolution of the proportion of high-skilled households at school boundaries*

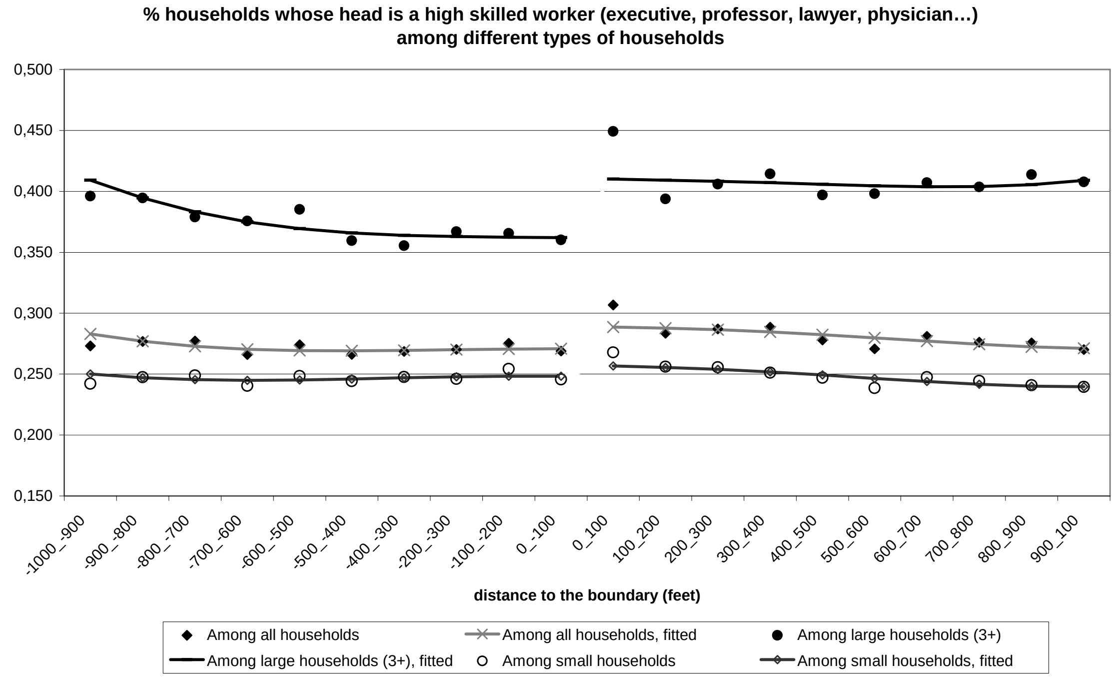

Figure 8: *The evolution of the proportion of low-skilled households at school boundaries*

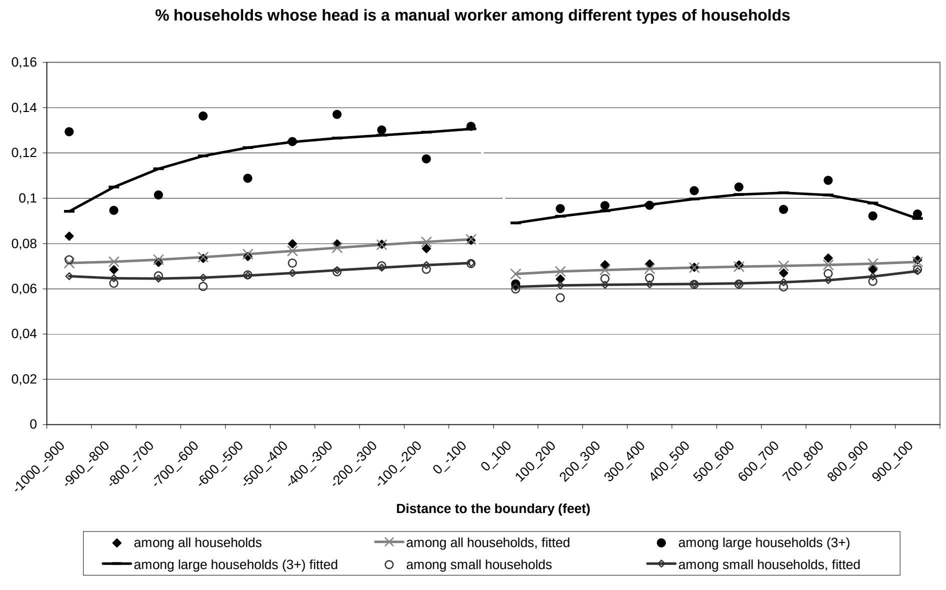

Figure 9: *Distribution of DNB exam scores at the school level for public and private middle schools in 2004.*

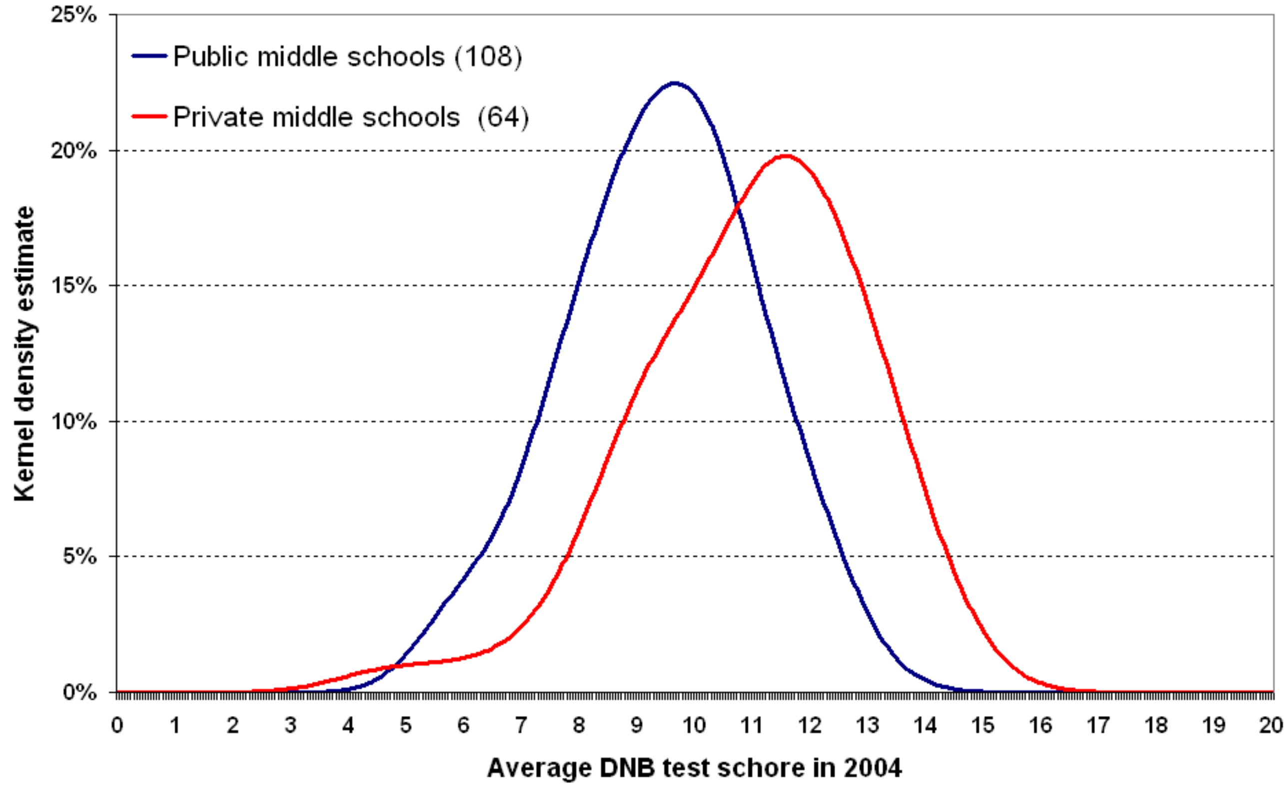

Figure 10: *Density of private middle schools in Paris.*

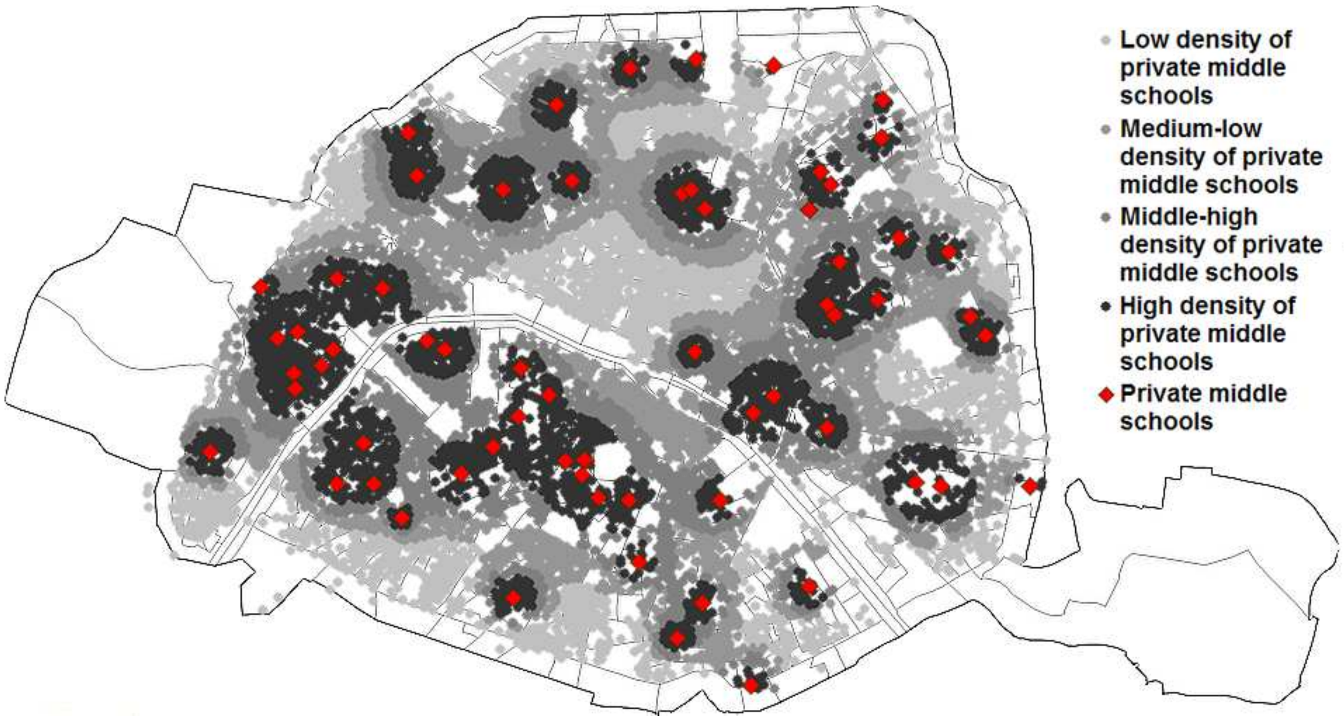

Table 1: *Summary statistics: characteristics of flats. School years 1997-2003.*

Distance from boundary	All sales		< 0.20 mile		< 0.15 mile	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Price (in 2004 euros)	185,509	183,237	184,333	173,865	183,748	170,787
Flat size (in square meters)	52.24	35.37	51.91	34.39	51.91	34.15
Price per m^2 (in 2004 euros)	3304	1295	3325	1265	3318	1249
FLAT CHARACTERISTICS						
Age of building (percent)						
Unknown	0.112		0.116		0.115	
Before 1850	0.054		0.054		0.051	
1850-1913	0.411		0.409		0.411	
1914-1947	0.152		0.159		0.161	
1948-1969	0.129		0.130		0.131	
1970-1980	0.111		0.105		0.104	
After 1981	0.030		0.027		0.026	
Floor (percent)						
Ground floor	0.103		0.107		0.108	
First	0.158		0.161		0.160	
Second	0.162		0.163		0.163	
Third	0.155		0.155		0.154	
Fourth	0.143		0.143		0.143	
Fifth	0.120		0.118		0.118	
Sixth or more	0.158		0.153		0.155	
Number of rooms (percent)						
One	0.235		0.236		0.234	
Two	0.367		0.365		0.365	
Three	0.222		0.223		0.225	
Four	0.105		0.105		0.106	
Five or more	0.072		0.071		0.071	
% without ind. bathroom	0.211		0.211		0.212	
% with maid's room	0.042		0.041		0.041	
% with lift	0.901		0.900		0.901	
% with garage	0.129		0.122		0.120	
Number of sales	180,522		110,453		87,653	

Table 2: *Summary statistics: census characteristics at the district level and school characteristics. School years 1997-2003.*

Distance from boundary	All sales		< 0.20 mile		< 0.15 mile	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
CENSUS VARIABLES						
Nb of census districts	902		771		705	
Nb of individuals per district	2291	789	2328	739	2336	730
Nb of households per district	1204	423	1236	403	1241	398
Nb of families per district	523	194	529	183	532	180
Nb of persons per flat	1.89	0.25	1.87	0.23	1.87	0.24
All households						
% families	0.440	0.08	0.434	0.08	0.434	0.08
% public housing	0.151	0.22	0.130	0.19	0.130	0.19
% owners	0.298	0.11	0.308	0.10	0.309	0.10
% with graduate degree	0.393	0.11	0.401	0.09	0.402	0.09
% female-headed households	0.061	0.03	0.058	0.03	0.058	0.03
% foreigners	0.227	0.06	0.223	0.06	0.222	0.06
% self-employed workers	0.039	0.02	0.040	0.02	0.039	0.02
% manual workers	0.062	0.04	0.060	0.04	0.059	0.04
% managers	0.220	0.07	0.225	0.06	0.227	0.06
% employees	0.149	0.05	0.145	0.04	0.144	0.04
% intermediary occupation	0.142	0.04	0.141	0.04	0.141	0.04
% unemployed	0.116	0.04	0.115	0.04	0.115	0.04
Families only						
% foreign families	0.217	0.12	0.215	0.11	0.213	0.11
% self-employed workers	0.098	0.06	0.099	0.06	0.099	0.05
% manual workers	0.131	0.08	0.128	0.08	0.127	0.08
% managers	0.380	0.16	0.388	0.15	0.392	0.15
% employees	0.151	0.09	0.146	0.09	0.144	0.09
% intermediary occupation	0.161	0.08	0.161	0.08	0.160	0.08
SCHOOL CHARACTERISTICS						
DNB exam score (out of 20)	9.40	1.62	9.47	1.56	9.47	1.56
% going into 2^{nde} *générale*	0.667	0.12	0.671	0.12	0.671	0.12
Nb of middle schools	108		105		105	
Number of sales	180,522		110,453		87,653	

Table 3: *Summary statistics: characteristics of public and private middle schools in 2004.*

	Mean	Standard deviation
PUBLIC SCHOOLS		
Total enrollment	56,711	
Average enrollment	520	(146)
DNB exam score (out of 20)	9.40	(1.62)
% going into *2nde générale*	0.67	(0.15)
Number of schools	108	
PRIVATE SCHOOLS		
Total enrollment	27,319	
Average enrollment	420	(246)
DNB exam score (out of 20)	10.97	(1.94)
% going into *2nde générale*	0.78	(0.18)
Number of schools	64	

Table 4: *Regressions results - naive estimates and estimates using boundary fixed effects. The school quality index is the average middle school exam score in 2004. School years 1997-2003.*

Dependent variable: log of price per square meter (in 2004 euros)

	(1)	(2)	(3)	(4)	(5)	(6)
	Naive estimates				Including boundary fixed effects	
Distance from boundary:	All sales	All sales	< 0.20 mile	< 0.15 mile	< 0.20 mile	< 0.15 mile
DNB exam score in 2004[a]	0.167**	0.027**	0.028**	0.029**	0.023**	0.020**
Adjusted standard error[b]	(0.013)	(0.008)	(0.008)	(0.009)	(0.005)	(0.005)
School year & quarter fixed effects	YES	YES	YES	YES	YES	YES
Flat characteristics[c]	YES	YES	YES	YES	YES	YES
Census variables[d]	NO	YES	YES	YES	NO	NO
Number of sales[e]	180,522	179,576	110,453	87,653	111,071	88,169
Number of boundaries					192	166

[a] The *Diplôme National du Brevet* exam score for each school is the average score in 2004 on the national exam that students take in their fourth year of middle high school (equivalent to ninth grade).

[b] Standard errors are adjusted for clustering at the attendance district level. *: significant at the 5% level; **: significant at the 1% level.

[c] Flat characteristics include a set of dummy variables for the age of the building (before 1850, 1850-1913, 1914-1947, 1948-1969, 1970-1980, after 1981), the number of bathrooms (1, 2 or more), the presence of a garage, the presence of a maid's room, the floor (first to fourth or more with and without a lift), the number of rooms (from 1 to 5 or more), average area per room (small, medium, large).

[d] Census variables include: average number of persons per flat, proportion of low-rent apartments, percent owners, percent families (*i.e.* household with at least one child), percent single-parent families, percent foreign , percent self-employed, percent managers, percent employed position, percent manual workers. The distribution of occupations is calculated separately for all households and for families only.

[e] The actual number of observations is higher than the number of sales, because flats that are close to several boundaries are used several times.

Table 5: *Regressions results - naive estimates and estimates using matching across boundaries. Housing prices (in 2004 euros) are "hedonic", i.e. they have been homogenized in terms of a reference flat. Period 1997-2003. The school quality index is the average middle school exam score in 2004. School years 1997-2003.*

	(1)	(2)	(3)	(4)
	Naive estimates		Matching across boundaries	
Distance from boundary:	< 0.20 mile	< 0.15 mile	< 0.20 mile	< 0.15 mile
Dependent variable:	Log of hedonic housing price		Log of hedonic housing price differential across boundaries	
DNB exam score in 2004[a]	0.022**	0.024**		
Adjusted standard error[b]	(0.009)	(0.010)		
Middle school exam score differential			0.020**	0.019**
Adjusted standard error[c]			(0.005)	(0.005)
School year & quarter fixed effects	YES	YES	NO	NO
Census variables[d]	YES	YES	NO	NO
Number of sales[e]	110,453	87,653	106,141	82,640

[a] The *Diplôme National du Brevet* exam score for each school is the average score in 2004 on the national exam that students take in their fourth year of middle high school (equivalent to ninth grade).

[b] Standard errors are adjusted for clustering at the attendance district level. *: significant at the 5% level; **: significant at the 1% level.

[c] Standard errors are adjusted for clustering at the attendance district boundary level. *: significant at the 5% level; **: significant at the 1% level.

[d] Census variables include: average number of persons per flat, proportion of low-rent apartments, percent owners, percent families (*i.e.* household with at least one child), percent single-parent families, percent foreign , percent self-employed, percent managers, percent employed position, percent manual workers. The distribution of occupations is calculated separately for all households and for families only.

[e] The actual number of observations is higher than the number of sales, because flats that are close to several boundaries are used several times.

Table 6: *Sensivity analysis - naive estimates and estimates using boundary fixed effects. Period 1997-2003. The school quality index is the fraction of 9^{th}-graders admitted into Seconde générale averaged over the school years 1997-2003.*

Dependent variable: log of price per square meter (in 2004 euros)

	(1)	(2)	(3)	(4)	(5)	(6)
	Naive estimates				Including boundary fixed effects	
Distance from boundary:	All sales	All sales	< 0.20 mile	< 0.15 mile	< 0.20 mile	<0.15 mile
Percent going into *2nde générale*[a]	0.145**	0.025**	0.025**	0.025**	0.028**	0.026**
Adjusted standard error[b]	(0.013)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)
School year & quarter fixed effects	YES	YES	YES	YES	YES	YES
Flat characteristics[c]	YES	YES	YES	YES	YES	YES
Census variables[d]	NO	YES	YES	YES	NO	NO
Number of sales[e]	180,522	179,576	110,453	87,653	111,071	88,169
Number of boundaries					192	166

[a] The fraction of 9^{th}-graders admitted into *Seconde générale* is averaged over the period 1998-2003 for each school.

[b] Standard errors are adjusted for clustering at the attendance district level. *: significant at the 5% level; **: significant at the 1% level.

[c] Flat characteristics include a set of dummy variables for the age of the building (before 1850, 1850-1913, 1914-1947, 1948-1969, 1970-1980, after 1981), the number of bathrooms (1, 2 or more), the presence of a garage, the presence of a maid's room, the floor (first to fourth or more with and without a lift), the number of rooms (from 1 to 5 or more), average area per room (small, medium, large).

[d] Census variables include: average number of persons per flat, proportion of low-rent apartments, percent owners, percent families (*i.e.* household with at least one child), percent single-parent families, percent foreign , percent self-employed, percent managers, percent employed position, percent manual workers. The distribution of occupations is calculated separately for all households and for families only.

[e] The actual number of observations is higher than the number of sales, because flats that are close to several boundaries are used several times.

Table 7: *Sensivity analysis - naive estimates and estimates using matching across boundaries. Housing prices (in 2004 euros) are "hedonic", i.e. they have been homogenized in terms of a reference flat. Period 1997-2003. The school quality index is the fraction of 9^{th}-graders admitted into Seconde générale averaged over the school years 1997-2003.*

	(1)	(2)	(3)	(4)
	Naive estimates		Matching across boundaries	
Distance from boundary:	< 0.20 mile	< 0.15 mile	< 0.20 mile	< 0.15 mile
Dependent variable:	Log of hedonic housing price		Log of hedonic housing price differential across boundaries	
Percent going into 2^{nde} *générale*[a]	0.024**	0.025**		
Adjusted standard error[b]	(0.006)	(0.006)		
Percent going into 2^{nde} *générale* differential			0.021**	0.020**
Adjusted standard error[c]			(0.005)	(0.005)
School year & quarter fixed effects	YES	YES	NO	NO
Census variables[d]	YES	YES	NO	NO
Number of sales[e]	110,453	87,653	106,141	82,640

[a] The fraction of 9^{th}-graders admitted into *Seconde générale* is averaged over the period 1998-2003 for each school.

[b] Standard errors are adjusted for clustering at the attendance district level. *: significant at the 5% level; **: significant at the 1% level.

[c] Standard errors are adjusted for clustering at the attendance district boundary level. *: significant at the 5% level; **: significant at the 1% level.

[d] Census variables include: average number of persons per flat, proportion of low-rent apartments, percent owners, percent families (*i.e.* household with at least one child), percent single-parent families, percent foreign , percent self-employed, percent managers, percent employed position, percent manual workers. The distribution of occupations is calculated separately for all households and for families only.

[e] The actual number of observations is higher than the number of sales, because flats that are close to several boundaries are used several times.

Table 8: *Comparison of the characteristics of the "good" and "bad" side (in terms of public middle school quality) of school attendance boundaries. School years 1997-2003.*

Distance from boundary	< 0.20 mile		< 0.15 mile	
	"Good" side	"Bad" side	"Good" side	"Bad" side
FLAT CHARACTERISTICS				
Age of building:				
Unknown	0.128	0.116	0.124	0.115
Before 1850	0.073	0.060	0.062	0.054
1850-1913	0.382	0.405	0.394	0.404
1914-1947	0.158	0.149	0.168	0.149
1948-1969	0.127	0.136	0.127	0.138
1970-1980	0.099	0.110	0.094	0.114
After 1981	0.033	0.024	0.030	0.02
Floor:				
Ground floor	0.106	0.101	0.108	0.104
First	0.161	0.157	0.163	0.158
Second	0.165	0.161	0.166	0.163
Third	0.161	0.154	0.158	0.152
Fourth	0.374	0.393	0.372	0.390
Fifth or more	0.321	0.330	0.033	0.032
Number of rooms:				
One	0.238	0.234	0.233	0.241
Two	0.348	0.362	0.359	0.363
Three	0.222	0.219	0.227	0.215
Four	0.108	0.107	0.105	0.106
Five or more	0.083	0.078	0.076	0.074
Bathrooms:				
No bathroom	0.202	0.212	0.209	0.209
One	0.739	0.735	0.739	0.740
Two	0.058	0.054	0.052	0.051
Garage:				
No garage	0.878	0.876	0.883	0.875
One	0.115	0.117	0.109	0.118
Two	0.007	0.007	0.007	0.007
Room size:				
Small	0.419	0.440	0.441	0.442
Average	0.342	0.338	0.340	0.340
Large	0.239	0.223	0.220	0.218
% with maid's room	0.050	0.045	0.045	0.041
School quality:				
DNB exam score (out of 20)	10.36	9.24	10.34	9.17
% going into *Seconde générale*	0.728	0.653	0.728	0.646

Table 9: *Regressions results - Differential Estimates for private middle school density using Matching across boundaries. Quality index: DNB exam score. School years 1997-2003.*

Dependent variable: log of price per square meter (in 2004 euros)

	(1)	(2)	(3)	(4)	(5)
		Quartiles of local private school density[a]			
	All sales	Lower Quartile	Middle Lower Quartile	Middle Upper Quartile	Upper Quartile
Panel A: Sales <0.20 mile of a boundary					
DNB exam score differential[b]	0.020**	0.039**	0.014**	0.018**	0.013**
Adjusted standard error[c]	(0.005)	(0.017)	(0.006)	(0.006)	(0.005)
Panel B: Sales <0.15 mile of a boundary					
DNB exam score differential	0.019**	0.034**	0.016**	0.018**	0.009
Adjusted standard error	(0.005)	(0.017)	(0.005)	(0.007)	(0.005)

[a] The indicator of the density of private middle schools is constructed as follows. For each transaction, we calculate the average of the inverse of its distance to every private middle school in Paris, squared and weighted by the total enrollment of each of these private schools. We then split our indicator into four quarters, from the lowest density of private schools to the highest.

[b] The *DNB* exam score for each school is the average score in 2004 on the *DNB* national exam that students take in their fourth (*i.e.* the last) year of middle high school.

[c] Standard errors are adjusted for clustering at the attendance district level. *: significant at the 5% level; **: significant at the 1% level.

Table 10: *Regressions results - Differential Estimates for private middle school density using Matching across boundaries. Quality index: Percent going into 2^{nde} générale. School years 1997-2003.*

Dependent variable: log of price per square meter (in 2004 euros)

	(1)	(2)	(3)	(4)	(5)
		Quartiles of local private school density[a]			
	All sales	Lower Quartile	Middle Lower Quartile	Middle Upper Quartile	Upper Quartile
Panel A: Sales <0.20 mile of a boundary					
Percent going into 2^{nde} *générale* differential[b]	0.021**	0.048**	0.014**	0.019**	0.008
Adjusted standard error[c]	(0.005)	(0.016)	(0.003)	(0.005)	(0.005)
Panel B: Sales <0.15 mile of a boundary					
Percent going into 2^{nde} *générale*differential[d]	0.020**	0.044**	0.015**	0.016**	0.004
Adjusted standard error	(0.005)	(0.017)	(0.003)	(0.006)	(0.005)

[a] The indicator of the density of private middle schools is constructed as follows. For each transaction, we calculate the average of the inverse of its distance to every private middle school in Paris, squared and weighted by the total enrollment of each of these private schools. We then split our indicator into four quarters, from the lowest density of private schools to the highest.

[b] The fraction of 9^{th}-graders admitted into *Seconde générale* is averaged over the period 1998-2003 for each school.

[c] Standard errors are adjusted for clustering at the attendance district level. *: significant at the 5% level; **: significant at the 1% level.

[d] The fraction of 9^{th}-graders admitted into *Seconde générale* is averaged over the period 1998-2003 for each school.

Table 11: *Regressions results - Example of hedonic regression for Zone 1 (*arrondissements *1 to 4). School years 1997-2003.*

Dependent variable: log of housing price[a]

Variable	Coefficient	(Std. Err)
Intercept	8.458	(0.019)
Age of Building		
Unknown	0.020	(0.009)
Before 1850	0.063	(0.007)
1850-1913	ref.	
1914-1947	-0.003	(0.012)
1948-1969	0.022	(0.015)
1970-1980	0.097	(0.023)
After 1981	0.061	(0.025)
Bathrooms		
No bathroom	-0.167	(0.007)
1 bathroom	ref.	
2 bathrooms	0.067	(0.017)
Garage		
No garage	ref.	
1 garage	0.182	(0.024)
2 garages	0.205	(0.085)
Floor		
Ground floor	ref.	
First	0.046	(0.015)
Second	0.077	(0.014)
Third	0.085	(0.014)
Fourth or more with lift	0.081	(0.013)
Fourth or more without lift	0.054	(0.018)
Number of rooms		
One	-0.064	(0.008)
Two	ref.	
Three	0.029	(0.009)
Four	0.073	(0.012)
Five or more	0.032	(0.015)
Room size		
Small	0.016	(0.007)
Average	ref.	
Large	0.036	(0.009)
Maid's room	0.068	(0.020)
Number of sales	10,338	

[a] Regressions include year and term fixed effects.