Decomposing the rural-urban differential in student achievement in Colombia using PISA microdata

Raul Ramos (AQR-IREA, UB)
Juan Carlos Duque (RiSE-group, EAFIT University)
Sandra Nieto (UOC & AQR-IREA)

Abstract: Despite the large number of studies that draw on Programme for International Student Assessment (PISA) microdata in their analyses of the determinants of educational outcomes, no more than a few consider the relevance of geographical location. This paper examines the differences in educational outcomes between students attending schools in rural areas and those enrolled in urban schools. We use microdata from the 2006, 2009 and 2012 PISA survey waves for Colombia. The Colombian case is particularly interesting in this regard due to the structural changes suffered by the country in recent years, both in terms of its political stability and of the educational reform measures introduced. Our descriptive analysis of the data shows that the educational outcomes of rural students are worse than those of urban students. In order to identify the factors underpinning this differential, we use the Oaxaca-Blinder decomposition and then exploit the time variation in the data using the methodology proposed by Juhn-Murphy-Pierce. Our results show that most of the differential is attributable to family characteristics as opposed to those of the school. From a policy perspective, our evidence supports the need to complement measures of positive discrimination of rural schools with actions addressed at improving household conditions.

Keywords: Educational outcomes, rural-urban differences, decomposition methods.

JEL codes: J24, I25, R58.

1. INTRODUCTION

One of the branches of the economics of education that has aroused greatest interest among researchers in recent decades has been the analysis of the factors influencing students' educational outcomes. The greater availability of statistical information has facilitated the analysis of this question in a greater number of countries and, more particularly, for a wider number of developing countries. A key concern in this regard is the analysis of possible differences in student performance at schools in rural and urban areas and the factors that account for this differential. In fact, educational policies are usually seen as one of the most relevant tools to promote development in rural areas (Oyarzún Méndez and Miranda Escolar, 2011).

In the various studies conducted to date numerous factors have been identified as determinants of students' educational outcomes[1] and, according to their nature, they can be categorised into three groups. The first group is made up of individual characteristic, among which, variables related to the student's nationality and main language stand out. It has been reported that the educational outcomes of immigrants are worse than those of native students (Meunier 2011, Chiswick and DebBurman 2004) and it is argued that this effect is related to the different home environments of each of the groups under analysis (Ammermueller, 2007a and Entorf and Lauk, 2008). In the case of languages, there is evidence that immigrants improve their academic outcomes when they speak the official language of the country in their home domain (Entorf and Minoiu, 2005).

The second group of variables refers to the family background. Coleman et al. (1966) was one of the earliest studies to show the impact of family variables on students' educational attainment. A number of studies, including Haveman and Wolfe (1995) and Feinstein and Symons (1999), claim that variables of this type have the greatest impact on educational performance. It is found that students whose parents have a high educational level obtain better outcomes than students whose parents have a lower level of education (Häkkinen et al. 2003, Woßmann 2003). In addition, the families' socio-economic level is also related to a student's academic performance – the outcomes improving the higher the parents' social and economic level. The genetic transmission of cognitive skills is one of the most frequently presented arguments for explaining the better performance of those students whose parents have a high level of education. Moreover, the presence of a good cultural environment and a stable family environment also contribute to enhance students' academic outcomes. In fact, there is usually a positive correlation between the parents' level of education and the family's socio-economic and cultural levels.

[1] Hanushek and Woßmann (2011) provide an up-to-date survey of recent literature on the topic.

Finally, the third group of variables is related with different characteristics of the school attended by the students including, for example, its urban or rural location, the type of school – public or private, the teacher-student ratio or school size.

The studies typically coincide in identifying the influence of individual characteristics and of family background on educational outcomes. However, this consensus is not so broad in studies that analyse the influence of variables relating to the schools attended by the students. Studies undertaken by Heyneman and Loxley (1983), Harbison and Hanushek (1992), Fuller and Clarke (1994), Gamoran and Long (2006), Banerjee et al. (2007) and Behrman (2010) found that the characteristics of the school have an important impact on academic performance in developing countries. Studies such as Coleman and Hoffer (1987), Hanushek (1986), Stevans and Sessions (2000), Vandernberghe and Robin (2004) and Opdenakker and Van Damme (2006) among others, find that students attain better outcomes in private than in public schools. Yet, other studies including, for example, Noell (1982), Sander (1996), Fertig (2003), Somers et al. (2004) and Smith and Naylor (2005), report no effect of school type on student outcomes. Likewise, the effect of school size on student outcomes is unclear. While Barnett et al. (2002) and Howely (2003) find a positive relation between school size and educational attainment, Hanushek and Luque (2003) do not observe any significant impact of this variable in the majority of countries analysed. Results regarding the impact of the number of students per teacher are similarly inconclusive. Arum (2000) and Krueger (2003) show that students perform better in small classes, while Hanushek (2003) and Rivkin et al. (2005) fail to find a statistically significant effect of this variable on students' educational outcomes.

Few studies have examined the impact of a school's rural or urban location on students' educational outcomes. The first were conducted in the United States in the mid-80s and to date there would appear to be no consensus on the significance of this characteristic. Thus, Edington and Martellaro (1984) and Ward and Murray (1985) find no significant differences in the outcomes of students at urban and rural schools in the state of New Mexico; similar findings are reported by Monk and Haller (1986) for the state of New York. Williams (2005) finds that, after controlling for the International Socio-Economic Index of Occupational Status (SES), the urban/rural location variable remains a statistically significant predictor of mathematics scores in only four of a sample of 24 countries. By contrast, Kleinfeld et al. (1985), in Alaska; Young (1998), in Western Australia; and Blackwell and McLaughlin (1999) and Roscigno and Crowley (2001), for the whole of the United States, do find the rural-urban location variable to be significant in explaining performance. The debate on the impact of this variable centres on the possibility that the differences in the performance of students in rural and urban schools are not due to the location per se, but rather to the fact that the characteristics of the students, their families and the schools differ in these two groups. Students in rural zones typically belong to families with few financial resources, their parents have low levels of education and the schools they attend are usually poorly endowed in terms of facilities and they are, generally, smaller than urban schools. Studies such as Hannaway and Talbert (1993) and Tayyaba (2012) claim that, rather than the location variable itself, it is these differences in the characteristics of urban and rural areas that account for most of the differences in the performance of students at rural and urban schools. The question is, therefore, in which cases (regions or countries), the location variable continues to be significant when it is studied in conjunction with other situational variables.

Over the last decade, attention has turned to emerging countries, due to the greater availability of data and the importance of the rural sector in these countries. Table 1 summarises the studies conducted in South American countries. The obtained evidence highlights the existence of differences in the location variable in most of the considered studies. Similar results are obtained in those studies focusing on other countries, mainly located in the Asian continent: for instance, Othman and Muijs (2013), for Malaysia, Lounkaew (2013) for Thailand and Karopady (2014) for India.

To the best of our knowledge, only two studies have been undertaken for the Colombian case examining the rural-urban differential in student attainment: Woßmann (2010) and Deutsch et al. (2013). Woßmann (2010)'s study was based on the test results of the Progress in International Reading Literacy Study of 2001 and he found that students living in settlements with more than three thousand inhabitants obtain outcomes that are 26 points higher than those obtained by students in rural zones[2]. However, Deutsch et al. (2013) include a control for location in the school in their analysis of individual efficiency using PISA 2006 data and they found no significant differences between rural and urban areas. Taking into account that the results from the previous studies are so different, it is interesting to analyse which has been the trend in the rural-urban gap in educational outcomes in Colombia using more recent data.

[2] Other studies of educational attainment undertaken in Colombia indicate that the main factors accounting for academic performance are socioeconomic level and the school's resources (Piñeros and Rodríguez, 1998; Gaviria and Barrientos, 2001a and Rangel and Lleras, 2010). The level of education of the parents also has a significant impact on the students' performance (Gaviria and Barrientos 2001b). As their main source of information, these studies use results from ICFES tests taken by all students in the final year of secondary schooling.

Table 1. Studies on rural-urban educational differences in South American countries

Study	Country	Data	Subject areas	Method & Characteristics
Harbison and Hanushek (1992)*	Brazil	EDURURAL data collection 1981, 1983, 1985. Second and fourth grades	Portuguese and mathematics	Longitudinal value added**; individual, family, teachers, peers, infrastructure, study materials
Mizala and Romaguera (2000)	Chile	SIMCE Educational Quality Measurement System (average) 1996 fourth and eighth grade	Mathematics and Spanish	OLS. family, personal, teachers, SIMCE 1994
Abdul_Hamid (2004)	Argentina	PISA 2000	Mathematics, reading and science	GLS; family, individual, school
Cueto et al. (2005)	Peru	Project: "Young Lives" 2002 (children between the ages of 7.5 and 8.5)	Reading, writing and mathematics	OLS, family, individual, home and community social capital
Santos (2007)***	Argentina	PISA 2000	Reading and mathematics	OLS, family, individual, school
Woßmann (2010)	Argentina, Colombia	Progress in International Reading Literacy Study (PIRLS). Fourth-grade students in 2001	Reading	WLS, CLRL; family, individual, school, test score in the previous period
Deutsch et al. (2013)****	Brazil, Chile, Colombia, Mexico, Uruguay	PISA 2006	Mathematics, Reading and Science	Individual efficiency related to educational means at home, school inputs and educational inputs

* Study centred on rural areas only (no urban-rural comparison undertaken).
** The aggregate value takes the variation in the student's score between the two periods as the endogenous variable. The remaining models take the score obtained by the student on a single test as the endogenous variable.
*** Rural-urban differential significant for mathematics but not for reading.
**** Rural-urban differential significant only for Mexico.

In fact, the 2009 PISA report, compiled by the OECD, analyses in part the importance of a school's location in accounting for differences in the results obtained on the reading test after controlling for the socioeconomic characteristics of the students' families (see Table II. 2.4 of OECD, 2010). The results show that while for the OECD as a whole the mean difference in the scores obtained by students in the least and most populated zones differed by around 4%, in Colombia this difference was over 8% (although it is true that in other countries, such as Panama, Peru and Argentina, the differentials were even more marked).

Moreover, an important aspect to take into account in the case of Colombia is that the study of differences in students' outcomes as a function of the rural-urban location of the school that they attend takes on special relevance if we consider the enormous gap between these two environments resulting from the armed conflict that for more than forty years has affected rural communities above all[3]. Forced migration of the population; the recruitment of minors by guerrilla groups[4]; confrontations between the army, guerrilla and paramilitary groups; attacks on school premises and the use of the schools as centres for military operations and recruitment; numerous murders of teachers[5]; among others, have constituted an obstacle to the normal development of schooling in the rural zones of Colombia.

According to Inter-American Commission on Human Rights (2006) and the UN Educational, Scientific and Cultural Organisation (2011) countries experiencing internal armed conflicts have poor performance in terms of education because of two main reasons: (1) the need to allocate significant resources into military spending, which reduces the available budget for education, and (2) the normal development of educational skills in youth is hampered because family income and cohesion is deeply affected by the conflict. This last reason is particularly evident in Colombian rural areas, where the actors in an armed, and dehumanized, conflict have used family disintegration as a strategy to gain control over the territory. According to Ibáñez and Vélez (2008), 29.1% of the Colombian rural population has been victim of forced migration.

Although both schools and families have been affected during the armed conflict, the Colombian government has decided to implement strategies on the supply-side (i.e., investing in schools) rather than on the demand-side (subsidies to families). As discussed in the Box IV.4.3 of OECD (2012), Colombia is one of the countries that have achieved a higher improvement between 2006 and 2012.

---

[3] For a review of the impact of violence on education investments in Latin America, see Ospina Plaza and Giménez Estaban (2009).

[4] According to War Child (2007), one in every four members of the illegal armed groups is under the age of fifteen; many of whom have been recruited in villages and rural schools.

[5] Colombia, together with Iraq, Nepal and Thailand, appears among the countries with the highest numbers of killings of teachers (O'Malley, 2010).

Improvement was related to different policies trying to improve access and quality of schooling. One of the most successful initiatives within this strategy was the "Rural Education Project" (PER) implemented in 2002 by the National Government and the World Bank. This program included pedagogical models and teaching material designed for the specific needs of students in rural areas, as well as specialized training for teachers. Rodriguez et al. (2009) measured the impact of the PER program and they found it to be a very successful project: it increased the passing rates, lowered the dropout rates and increased the quality of education. Other programmes such as PACES (Programa de Ampliación de la Cobertura y Mejoramiento de la Calidad de la Educación Secundaria) and Escuela Nueva have also contributed to this improvement.

However, although the supply-side intervention proved to be effective, there still exist a gap between urban-rural students' outcomes. Is this gap the reflection of the need for an intervention on the demand side? Or, even if we discount the effects attributable to differences in student and family profiles and the characteristics of the schools, may the location variable well be a determinant of differences in student performance?

In order to analyse the possible existence of differences in educational outcomes for students attending schools in the rural and urban areas of Colombia, we draw on data from the 2006, 2009 and 2012 PISA survey waves to examine the results obtained in the subject areas of mathematics, science and reading. To do so, we apply methods of decomposition of the rural-urban differential by estimating an educational production function that includes explanatory variables related to the characteristics of the students, their families and the schools they attend. The application of the decomposition proposed by Oaxaca (1973) and Blinder (1973), which has been widely used in the framework of labour economics, for example, to try to explain the causes of wage differentials between men and women, should enable us to identify which variables contribute most to explain the differences in educational outcomes between rural and urban areas. Additionally, the extension of this methodology as proposed by Juhn, Murphy and Pierce (1993) allows us to determine the factors that explain the changes in the differential between rural and urban areas over time, thereby providing the ideal framework for exploiting the time dimension in the data[6].

The rest of the paper is organized as follows. Section 2 presents the database and defines the variables of interest for the study. Then, section 3 describes the

[6] These techniques have been rarely used in this context. Some exceptions include Burger (2011), Zhang & Lee (2011) and Ammermueller (2007b). Of these three studies, the only one to examine the rural-urban differential is Burger (2011) who uses data on educational performance in Zambia obtained from a survey that is distinct to that of PISA. Her results suggest that both the characteristics of students as well as the outcomes obtained are important in explaining the rural-urban differential.

methodological approach used and the results obtained. Finally, we summarize the main findings and policy conclusions.

2 EDUCATIONAL PERFORMANCE AND STUDENTS' AND SCHOOL'S CHARACTERISTICS

The data source drawn on in this study is the Programme for International Student Assessment (PISA), coordinated by the OECD, which aims to assess students on reaching the end of compulsory education, at the age of 15, in the subject areas of mathematics, science and reading, providing, in addition, information about the students themselves, their family background and the school as a learning environment. It is a triennial survey that currently provides data for five waves: 2000, 2003, 2006, 2009 and 2012. The set of countries analyzed in each of the years has grown over time to include 65 countries in 2012. Colombia is one of the countries included in the latest waves. Specifically, data are available for 2006, 2009 and 2012, which are the sources we use here.

As mentioned above, the main objective of PISA is to assess student attainment on reaching the end of compulsory education in the subject areas of mathematics, science and reading. To this end, the survey provides five plausible values for each subject area. Plausible values are not the students' actual test scores and should not, therefore, be treated as such; rather, they are random numbers drawn from the distribution of scores that could be reasonably assigned to each individual. This methodology was developed by Mislevy and Sheehan (1987, 1989) and is based on Rubin's theory for imputing missing or lost values (1987). The idea is that each individual responds to a limited number of test questions, and, for this reason, it is necessary to estimate their behaviour as if they had answered all the questions on the test. To do this, their results are predicted using the responses to the questions they have actually answered and other variables obtained from the context questionnaire. Instead of predicting a single score, a distribution of values is generated for each individual with their associated probabilities and five plausible values are obtained randomly for each individual. In this way, the bias introduced when estimating the outcomes from a small number of test questions is avoided. Plausible values contain random error variance components and are not optimal as individual test scores. Thus, while unsuitable for the diagnosis of subjects they are well suited to the consistent estimation of population parameters. In this analysis, we use these values to conduct our proposed empirical analysis; however, in the descriptive statistics shown below the mean values are used. We have also used, in all cases, the weighting factors provided by the survey itself both for 2006, 2009 and for 2012. The results are shown for all students and for boys and girls, separately as explanatory factors of educational outcomes can be different according to gender (see, for instance, González de San Román and de la Rica, 2016).

As regards the other variables of interest, the individual characteristics provided in the survey and considered in our analysis are gender and age. Regarding family characteristics, we have been unable to control for the type of family structure (nuclear, single parent and mixed race), as this information was not included in the 2006 questionnaire, but, as in other studies, we include the educational level of the parents (Meunier 2011, Martins and Veiga, 2010, among others), the economic, social and cultural status (ESCS) of the family, the availability of computer in the household and the students' cultural background proxied by the number of books found in the home. Finally, we also included variables related to the school including its location in urban or rural areas (the key variable in this study)[7], school size, the number of students per teacher, the proportion of public funding received by the school and three indicators related to school quality: the index of school responsibility for curriculum and assessment, the index of quality of school's educational resources and teacher shortage. Tables A1, A2 and A3 in the appendix show the main descriptive statistics for the variables described above.

The figure 1 shows the evolution of the average scores in Mathematics, Reading and Science between 2006 and 2012 in Colombia. The improvement in educational outcomes by Colombian students is clearly appreciated in this figure, although it is worth mentioning that most of the improvement was achieved between 2006 and 2009, while between 2009 and 2012, results are very similar.

Figure 2 shows the gap in average scores in Mathematics, Reading and Science between students in rural and urban areas for the three considered years. Differences along the distribution are shown in figures A1, A2 and A3 in the appendix. From these results, it is clear that the educational achievement of students in rural areas is worse than that of students in urban areas, both for boys and girls. This marked differential is approximately 30 points in both periods, although when we compare the evolution in outcomes between 2006 and 2009 we find an increase in the gap, that is partially reduced between 2009 and 2012, particularly for girls.

In the next section, we apply statistical and econometric techniques to analyse the influence of these variables on the differences in educational performance recorded between students in rural schools and those in urban schools.

[7] An urban school is defined as a school located in a city or large city. A rural school is defined as a school located in a village, small town or town.

Figure 1. Educational outcomes in Colombia (2006, 2009 and 2012)

Figure 2. Raw gap in scores between rural and urban students in Colombia

3 METHODOLOGY AND RESULTS

Thus, the first step in determining whether the differences observed in the educational outcomes of students attending schools in rural and urban areas of Colombia are related to individual factors or to characteristics of the family or school environment, we specify and estimate an educational production function which includes various controls at the individual, family and school levels. Specifically, the educational production function for each of the subject areas used in this study is based on the following expression:

$$RTest_i = \alpha + \beta \cdot Z_i + e_i \tag{1}$$

where $RTest_i$ refers to the five plausible values of the test results in each subject area for student i, Z_i is a vector of control variables related to the characteristics of the individuals, their family backgrounds and school environment, while e_i is a random error term.

The results of estimating model (1) for the main variables of interest using 2006, 2009 and 2012 PISA microdata and for the three subject areas tested (Mathematics, Reading and Science) for all students in rural and urban areas and disaggregated by gender are shown in tables A.4, A.5 and A.6 in the appendix. Given the nature of the dependent variable (described in detail above), in order to estimate this model we need a method that will allow us to make multiple estimations of the dependent variable[8], which refers to the five plausible values of the educational outcomes in each subject area. Additionally, and due to the complex sample design used in PISA, a replication procedure has to be applied to calculate the variance of the estimators. For data of this type, the OECD (2009) recommends the Fay-modified balanced repeated replication (BRR) method (Fay, 1989), which improves the accuracy of the variance estimator without modifying the coefficients. This was the procedure adopted in this study. For space limitations, we cannot here describe in detail the estimates of all the models, but the obtained results are quite similar to those obtained in other studies using PISA microdata. Specifically, and as expected, a student's gender has a statistically significant effect on his or her academic outcomes, although the sign differs depending on the subject area under analysis. Girls record poorer academic outcomes than boys in Mathematics and Science, but present better results in Reading. The age of the students, around 15 years and 9 months with small

[8] To do so we employed the Stata module for performing estimations with plausible values. http://ideas.repec.org/c/boc/bocode/s456951.html

variations either way of 3 months, has a positive impact as it increases in all three subject areas. In the case of the set of variables related to a student's family background, we see that the dummy variables referring to the number of books in the family home, included as an indicator of the cultural environment, have a positive effect on the student's educational performance, which improves as the number of books in the home increases. Likewise, the mother's educational level has a positive effect on the academic performance of her children. However, the same does not hold for the father's educational level, although it is worth mentioning that educational levels are also related to the ESCS index which is positive and statistically significant in nearly all models. As for the variables related to the characteristics of the school, it can be seen that none of the usual characteristics (public/private, size and student-teacher ratio) is statistically significant. In fact, the most relevant variables are associated to the degree of school autonomy, the quality of educational resources and in a more limited number of models to the teacher shortage.

Returning to the main focus of this study, the analysis of differences between rural and urban areas, the rest of this section involves a decomposition of the differences in educational outcomes between students attending schools in rural areas and those enrolled in schools in urban areas by applying the Oaxaca-Blinder methodology followed by the Juhn-Murphy-Pierce method.

As discussed in the introduction, the wage decomposition methodology of Oaxaca-Blinder has been widely used to analyze employment discrimination on grounds of gender, race or other worker characteristics. As is well known, the technique allows us to decompose the difference between two groups in the mean level for a given variable into a part that is explained by group differences in the observed characteristic and a part caused by differences in the outcomes associated with these characteristics. The Juhn-Murphy-Pierce extension of this methodology represents an important advance in these decomposition techniques, to the extent that it enables us to decompose the changes in the differences over time between the two groups studied.

Based on the educational production function estimated jointly for students in rural and urban areas as the reference structure in the decomposition, the difference in the educational performance of both groups can be expressed as:

$$\overline{RTest}_R - \overline{RTest}_U = (\bar{Z}_R - \bar{Z}_U) \cdot \beta_R + \bar{Z}_U \cdot (\beta_R - \beta_U) + (\bar{e}_R - \bar{e}_U) \qquad (2)$$

where the subindices R and U correspond to rural and urban areas respectively. Equation (2) enables us to quantify the extent to which the cause of the differences between students in rural and urban areas is related to differences observed in individual factors or in characteristics of the family or the school environment, or to the influence of unobserved factors. More specifically, the first term on the right-hand side of the equation corresponds to that part of the differential in educational performance attributable to the group differences in the observed characteristics, coinciding with the "explained" component of the Oaxaca-Blinder decomposition, while the second and third terms correspond to the difference in coefficients and differences in unobservable skills and capture, basically, the discriminatory or "unexplained" component of this decomposition.

The results obtained when applying the Oaxaca-Blinder decomposition[9] for the 2006, 2009 and 2012 PISA waves, using as our reference structure the estimation of the educational production function for the whole of the sample, are presented in Table 2.[10] As it can be seen from this table, much of the gap between rural and urban studies can be attributed to the poorer characteristics of students in rural areas. In fact, the “unexplained” part is not statistically significant at the usual level in most models. In all three subject areas and for the different years and groups of students, the explained accounts for over 90% of the “explained” part of the differential. Is it worth mentioning that individual and family characteristics play a major role in explaining these differences (more than 60% of the gap), although school characteristics are becoming more important. In the three subjects, and particularly for boys, the relative importance of schools to explain the gap between rural and urban students is higher in 2012 than in 2006.

The Juhn-Murphy-Pierce decomposition assumes that the contribution of the individual characteristics is the same for both groups. Thus, the starting point for this decomposition is the following:

$$\overline{RTest}_R - \overline{RTest}_U = (\bar{Z}_R - \bar{Z}_U) \cdot \beta_R - \bar{e}_U = (\bar{Z}_R - \bar{Z}_U) \cdot \beta_R - \bar{\theta}_U \cdot \sigma_R \quad (3)$$

[9] To do so we employed the Stata module to compute the Blinder-Oaxaca decomposition, http://ideas.repec.org/c/boc/bocode/s456936.html

[10] Various tests of robustness were conducted on different regressions but the results remained largely unchanged. The advantage of working with the whole sample rather than with the information as it relates separately to students in urban and rural areas is that our results are directly comparable with those obtained when conducting the Juhn-Murphy-Pierce decomposition. Detailed results of the decomposition are available from the authors on request.

where σ_R is the standard deviation of the residuals (e_R) and $\theta U = e_U/\sigma_R$. The interpretation of both terms is similar to that described above in the Oaxaca-Blinder decomposition. If, on the basis of this equation, we compare the changes in the educational performance differential between two different points in time (for example, *t* and *t'*), we obtain the following expression:

$$D_{t'} - D_t = (\Delta Z_{t'} - \Delta Z_t) \cdot \beta_{Rt} + \Delta Z_{t'}(\beta_{Rt'} - \beta_{Rt}) +$$
$$+(\Delta\theta_{t'} - \Delta\theta_t) \cdot \sigma_{Rt} + \Delta\theta_{t'}(\sigma_{Rt'} - \sigma_{Rt}) \qquad (4)$$

where $D_{t'}$ represents the differential in the mean educational performance of students in rural and urban areas at time *t'*, D_t represents the same differential but at time t and the symbol Δ denotes the variation between rural and urban areas for each of the associated variables or parameters. The rest of the elements follow exactly the same notation as in (3). The first term in (4) corresponds to the change observed in the characteristics (quantity effect); the second term is related to changes in the coefficients and, therefore, with variations in prices (price effect); the third is related to the interaction between the two; while, the last term captures the variation not explained by the previous ones.

Table 3 shows the results of applying this methodology[11] in order to explain the variations in the educational performance differential between rural and urban areas in 2006 and 2009 and in 2009 and 2012 in each of the subject areas (Mathematics, Reading and Science). The table only presents the relative contribution of the first term (quantity term) while the relative contribution of the other three terms has been grouped into one term[12]. As can be seen from this table, between 2006 and 2009 the gap increased in all three subjects for boys and girls (with the only exception of Reading for boys where the change is not statistically significant). In all cases, the variation in the "explained" part has contributed to increase the differential. In particular, in the case of boys, the gap has increased between 2006 and 2009 due to the worst characteristics of rural schools compared to urban ones. In the case of girls, this negative contribution of schools is also associated to a relative worsening of individual and family characteristics.

[11] To do so we employed the Stata module JMPIERCE2 to compute trend decomposition of outcome differentials, http://ideas.repec.org/c/boc/bocode/s448804.html
[12] Detailed results are available from the authors on request.

Table 2. Oaxaca-Blinder decomposition of the rural-urban differential for Colombian students

All students	Mathematics			Reading			Science		
	2006	2009	2012	2006	2009	2012	2006	2009	2012
Raw gap (Rural – Urban)	-22.0***	-29.1***	-24.4***	-27.1***	-35.8***	-32.3***	-17.0***	-26.3***	-21.3***
“Explained” part	-23.7***	-26.7***	-22.7***	-25.0***	-30.7***	-28.4***	-19.1***	-25.6***	-22.3***
Individual and family characteristics	-17.7***	-16.8***	-14.8***	-20.4***	-19.3***	-18.6***	-15.9***	-17.3***	-14.4***
School characteristics	-6.0***	-9.9***	-7.9***	-4.6*	-11.4***	-9.8***	-3.2*	-8.3***	-7.9***
“Unexplained” part	1.7	-2.4	-1.7	-2.1	-5.1	-3.9	2.1	-0.7	1.0

Boys	Mathematics			Reading			Science		
	2006	2009	2012	2006	2009	2012	2006	2009	2012
Raw gap (Rural – Urban)	-23.7***	-30.9***	-31.3***	-34.1***	-35.7***	-38.4***	-18.4***	-28.7***	-27.1***
“Explained” part	-25.8***	-32.0***	-26.3***	-24.9***	-33.1***	-31.1***	-20.6***	-28.2***	-26.6***
Individual and family characteristics	-18.7***	-17.4***	-16.6***	-20.7***	-16.3***	-19.8***	-16.7***	-15.9***	-16.6***
School characteristics	-7.1**	-14.6***	-9.7***	-4.2	-16.8***	-11.3***	-3.9	-12.3***	-10.0***
“Unexplained” part	2.1	1.1	-5.0	-9.2	-2.6	-7.3	2.2	-0.5	-0.5

Girls	Mathematics			Reading			Science		
	2006	2009	2012	2006	2009	2012	2006	2009	2012
Raw gap (Rural – Urban)	-19.3***	-29.4***	-19.4***	-22.3***	-35.3***	-25.7***	-15.5***	-25.3***	-17.0***
“Explained” part	-20.6***	-24.1***	-21.2***	-26.1***	-28.1***	-25.3***	-17.1***	-25.9***	-19.9***
Individual and family characteristics	-15.1***	-18.2***	-14.8***	-20.9***	-21.2***	-16.8***	-14.2***	-20.5***	-13.7***
School characteristics	-5.5**	-5.9***	-6.4***	-5.2*	-6.9***	-8.5***	-2.9	-5.4***	-6.2***
“Unexplained” part	1.3	-5.3	1.8	3.8	-7.2	-0.4	1.6	0.6	2.9

* p<0,1; ** p<0,05; *** p<0,01

Table 3. Juhn-Murphy-Pierce decomposition of the rural-urban differential for Colombia

All students	Mathematics		Reading		Science	
	2006-2009	2009-2012	2006-2009	2009-2012	2006-2009	2009-2012
Change in the raw gap (Rural – Urban)	7.1***	-4.7***	8.8***	-3.6***	9.3***	-5.0***
"Explained" part	8.5***	-8.0***	14.9***	-5.4***	11.0***	-6.4***
Individual and family characteristics	0.9	-2.6*	1.3	-0.7	2.9**	-4.2***
School characteristics	7.6***	-5.4***	13.6***	-4.7***	8.1***	-2.2*
"Unexplained" part	-1.4	3.3*	-6.1**	1.8	-1.7	1.4

Boys	Mathematics		Reading		Science	
	2006-2009	2009-2012	2006-2009	2009-2012	2006-2009	2009-2012
Change in the raw gap (Rural – Urban)	7.2***	0.4	1.7	2.7	10.3***	-1.5
"Explained" part	5.1***	-7.1***	12.1***	-2.9	9.8***	-3.8**
Individual and family characteristics	-1.1	-2.7	-4.0**	2.5	0.4	0.0
School characteristics	6.2***	-4.4***	16.1***	-5.4**	9.4***	-3.8**
"Unexplained" part	2.1	7.5***	-10.4***	5.6**	0.5	2.3*

Girls	Mathematics		Reading		Science	
	2006-2009	2009-2012	2006-2009	2009-2012	2006-2009	2009-2012
Change in the raw gap (Rural – Urban)	10.0***	-9.9***	13.0***	-9.6***	9.9***	-8.4***
"Explained" part	14.0***	-7.7***	13.5***	-6.5***	14.1***	-8.1***
Individual and family characteristics	5.6***	-2.5	3.5	-3.6***	7.9***	-7.9***
School characteristics	8.4***	-5.2***	10.0***	-2.9**	6.2***	-0.2
"Unexplained" part	-4.0*	-2.2	-0.5	-3.1*	-4.2**	-0.3

* p<0,1; ** p<0,05; *** p<0,01

Results for the period 2009-2012 are very different. While in the case of boys the gap has remained stable, for girls there has been a clear reduction of the gap. This reduction is mainly associated to a relative improvement of school characteristics (a result that is also observed for boys, although compensated by the evolution of the unexplained component), but also to a clear improvement in the family and school characteristics in rural areas when compared to urban. In fact, taking together the results from the Oaxaca and the Juhn-Murphy-Pierce decomposition, the obtained evidence shows that there is room to improve educational outcomes by improving school quality, but policies addressed to families will also be a proper instrument to fight against the rural-urban gap. The fact that the unexplained term is not statistically significant can be understood as evidence that the end of the armed conflict has not played an important role to explain the trend in educational outcomes. However, it is possible that some of the educational reforms that have taken place in Colombia in the last years would have not been possible if the violence had not ceased.

4 CONCLUSIONS

This paper has analysed the possible existence of differences in the educational performance of students in rural and urban areas of Colombia in the subject areas of mathematics, science and reading. To do so, we have used data from the 2006, 2009 and 2012 PISA survey waves and we have specified and estimated an education production function that includes variables related to the location of the school and to the typical controls at the individual and family levels. Additionally, and so as to identify the factors that account for any differences, we have used Oaxaca-Blinder decomposition and the Juhn-Murphy-Pierce decomposition method to analyse the time variation in these differences. The results obtained from the application of both methods show that most of the rural-urban school differential is related to family characteristics and not so much to those of the school, although the analysis of the time dimension has enabled us to highlight the role of the improvement in school quality in more recent times to explain the observed reduction of the rural-urban gap in educational outcomes.

From the perspective of educational policy, the evidence obtained reinforces the suitability of adopting measures aimed at improving the general educational situation and conditions in the family and, perhaps, as opposed at only adopting measures of positive discrimination in rural schools as a means to improve educational performance. Our evidence supports that both ways should be explored.

Based on these results, several future paths of research are opened up. However, such studies will require a richer database as regards the information needed to capture the characteristics of the areas in which the students are resident. Such data would enable us to analyse the mechanisms via which the geographical environment can have an impact on a student's educational outcomes and the extent to which this fails to capture the importance of other variables that we have been unable to control for adequately in this study (omission of relevant variables). These might be found to include the institutional improvements that have occurred in Colombia in the period under review and which may have had a greater impact on rural than they have had on urban zones.

REFERENCES

ABDUL-HAMID, H. (2004): "Assessing Argentina's preparedness for the knowledge economy: Measuring student knowledge and skills in reading, mathematical and scientific literacy. Evidence from PISA 2000", mimeo.

AMMERMUELLER, A. (2007a): "Poor background or low returns? Why immigrant students in Germany perform so poorly in the programme for international student assessment", *Education Economics* 15(2), 215–230.

AMMERMUELLER, A. (2007b): "PISA: What makes the difference?: Explaining the gap in test scores between Finland and Germany", *Empirical Economics* 33 (2), 263-287.

ARUM, R. (2000): "Schools and communities: Ecological and institutional dimensions", *Annual Review of Sociology* 26, 395–418.

BANERJEE, A., COLE, S., DUFLO, E., LINDEN, L. (2007): "Remedying education: Evidence from two randomized experiments in India", *The Quarterly Journal of Economics* 122(3), 1235–1264.

BARNETT, R. R., GLASS, J. C., SNOWDON, R. I., STRINGER, K. S. (2002): "Size, performance and effectiveness: cost-constrained measures of best-practice performance and secondary-school size", *Education Economics* 10 (3), 291-311.

BEHRMAN, J.R. (2010): "Investment in education: Inputs and incentives", In: Rodrik, D. and Rosenzweig, M. R. (Eds.), *Handbook of development economics: The economics of development policy*. Vol. 5, pp. 4883–4975. Amsterdam: North-Holland Publishing.

BLACKWELL, D.L., MCLAUGHLIN, D.K. (1999): "Do Rural Youth Attain Their Educational Goals?". *Rural Development Perspectives* 13(3), 37-44

BLINDER, A. S. (1973), "Wage Discrimination: Reduced Form and Structural Variables", *Journal of Human Resources* 8, 436-455.

BURGER, R. (2011): "School effectiveness in Zambia: The origins of differences between rural and urban outcomes", *Development Southern Africa* 28 (2), 157-176.

CHISWICK, B. R., DEBBURMAN, N. (2004), "Educational attainment: analysis by immigrant generation", *Economics of Education Review* 23, 361-379.

COLEMAN, J. S., HOFFER, T. (1987): *Public and private high schools. The impact of communities*. New York: Basic Books.
COLEMAN, J. S., CAMPBELL, J. S., CAMPBELL, E. Q., HOBSON, C. J., MCPARTLAND, J., MOOD, A. M., WEINFELD, F. D., YORK, R. L. (1966): *Equality of Educational Opportunity*. Washington, D.C.: US Department of Health, Education, and Welfare, US Government Printing Office.
CUETO, S., GUERRERO, G., LEON, J., DE SILVA, M., HUTTLY, S., PENNY, M. E., LANATA, C. F., VILLAR, E. (2005): *Social capital and education outcomes in urban and rural Peru*. Working paper # 28, Young Lives, London.
DEUTSCH, J., DUMAS, A., SILBER,A., (2013): “Estimating an educational production function for five countries of Latin America on the basis of the PISA data”, *Economics of Education Review*, 36, 245-262.
EDINGTON, E.D., MARTELLARO, H.C. (1984): “Variables affecting academic achievement in New Mexico schools. Las Cruces”, *New Mexico Center for Rural Education*. (Eric Document Reproduction Service No. 271 267)
ENTORF, H., LAUK, M. (2008): “Peer effects, social multipliers and migrants at school: An international comparison” *Journal of Ethnic and Migration Studies* 34(4), 633–654.
ENTORF, H., MINOIU, N. (2005): “What a difference immigration policy makes: A comparison of PISA scores in Europe and traditional countries of immigration”, *German Economic Review* 3, 355–376.
FAY, R. E. (1989): “Theoretical application of weighting for variance calculation”, *Proceedings of the Section on Survey Research Methods of the American Statistical Association*, 212-217.
FULLER, B., CLARKE, P. (1994): “Raising school effects while ignoring culture? Local conditions and the influence of classroom tools, rules and pedagogy”, *Review of Educational Research* 64(1), 119–157.
FEINSTEIN, L., SYMONS, J. (1999): “Attainment in secondary education”, *Oxford Economic Papers* 51, 300-321.
FERTIG, M. (2003): “Who’s to Blame? The Determinants of German Students’ Achievement in the PISA 2000 Study”, *IZA Discussion Paper Series* 739.
GAMORAN, A., LONG, D. A. (2007): “Equality of educational opportunity: A 40-year retrospective. International Studies in Educational Inequality” In: Teese, R., Lamb, S. and Duru-Bellat, M., Editors. *International studies in educational inequality: Theory and policy*, Dordrecht, Springer, pp. 23–47.
GAVIRIA, A., BARRIENTOS, J. H. (2001a): “Determinantes de la calidad de la educación en Colombia”, *Archivos de Economía* 159.
GAVIRIA, A., BARRIENTOS, J. H. (2001b): “Calidad de la educación y rendimiento académico en Bogotá”, *Coyuntura Social* 24, 111–127.
GONZÁLEZ DE SAN ROMÁN, A., DE LA RICA, S. (2016), “Gender Gaps in PISA Test Scores: The Impact of Social Norms and the Mother’s Transmission of Role Attitudes”, *Estudios de Economía Aplicada*, 34 (1), 79-108.
HÄKKINEN, I., KIRJAVAINEN, T., UUSITALO, R. (2003): “School resources and student achievement revisited: new evidence from panel data”, *Economics of Education Review* 22, 329-335.
HANNAWAY, J., TALBERT, J.E. (1993): “Bringing context into effective school research: Urban-suburban differences”, *Educational Administration Quarterly* 29, 164–186.
HANUSHEK, E. A. (1986): “The economics of schooling”, *Journal of Economic Literature* 24, 1141-1177.
HANUSHEK, E. A. (2003): “The failure of input based schooling policies”, *The Economic Journal* 113, 64-98.
HANUSHEK, E. A., LUQUE, J. A. (2003): “Efficiency and equity in schools around the world” *Economics of Education Review* 22, 481-502.

HANUSHEK, E. A., WOESSMANN, L. (2011), "The Economics of International Differences in Educational Achievement" in Hanushek, E. A., Machin, S., Woessman, L. (eds.), *Handbook of the Economics of Education*, Volume 3, chapter 2, pp. 89-200.

HARBISON, R. W., HANUSHEK, E. A. (1992): *Educational performance of the poor: Lessons from rural northeast Brazil*, Oxford: Published for the World Bank, Oxford University Press.

HAVEMAN, R., WOLFE, B. (1995): "The determinants of children's attainment: A review of methods and findings", *Journal of Economics Literature* 33(4), 1829–1878

HEYNEMAN S. P., LOXLEY, W. A. (1983), "The effect of primary school quality on academic achievement across 29 high and low income countries", *American Journal of Sociology* 88(6), 1162–1194.

HOWLEY, C. (2003): "Understanding mathematics education in rural context", *Educational Forum* 67(3), 215-224.

IBÁÑEZ, A.M., VÉLEZ, C.E. (2008): "Civil Conflict and Forced Migration: The Micro Determinants and Welfare Losses of Displacement in Colombia", *World Development* 36(4), 659–676.

INTER-AMERICAN COMMISSION ON HUMAN RIGHTS (2006): *Violence and Discrimination Against Women in the Armed Conflict in Colombia*. OEA/Ser.L/V/II. Doc 67, October 18.

JUHN, C., MURPHY, K., PIERCE, B. (1993): "Wage inequality and the rise in returns to skill" *Journal of Political Economy* 101 (31), 410-442.

KAROPADY, D.D. (2014), Does school choice help rural children from disadvantaged sections?, Economic and Political Weekly, 49 (51), pp. 46-53.

KLEINFELD, J.S., MCDIARMID, G.W., HAGSTROM, D. (1985): "Alaska's small rural high schools: Are they working?" *ISER Report Series* No. 57, Alaska: University of Alaska.

KRUEGER, A. B. (2003): "Economics considerations and class size" *Economic Journal* 113, 34-63.

LOUNKAEW, K. (2013), "Explaining urban-rural differences in educational achievement in Thailand: Evidence from PISA literacy data", *Economics of Education Review*, 37, 213-225.

MARTINS, L., VEIGA, V. (2010): "Do inequalities in parents' education play an important role in PISA students' mathematics achievement test score disparities?" *Economics of Education Review* 29 (6), 1016-1033.

MEUNIER, M. (2011), "Immigration and student achievement: Evidence from Switzerland", *Economics of Education Review* 30 (1), 16-38.

MISLEVY, R. J., SHEEHAN, K. M. (1987): "Marginal estimation procedures". In: Beaton, A.E., Editor, 1987. The NAEP 1983-84 technical report, National Assessment of Educational Progress, Educational Testing Service, Princeton, 293-360.

MISLEVY, R. J., SHEEHAN, K. M. (1989): "Information matrices in latent-variable models", *Journal of Educational Statistics* 14, 335-350.

MIZALA, A., ROMAGUERA, P. (2000): "School performance and Choice: The Chilean Experience" *Journal of Human Resources* 35(2), 392-417.

MONK, D. H., HALLER, E. J. (1986): *Organizational alternatives for small/rural schools: final report to the New York state legislature*. New York, NY: Cornell University, 1986. ED 281 694.

NOELL, J. (1982): "Public and Catholic schools: A re-analysis of public and private schools", *Sociology of Education* 55, 123-132.

OAXACA, R. (1973): "Male-Female Wage Differentials in Urban Labor Markets", *International Economic Review* 14 (3), 139-148.

OECD (2009): *Technical report 2006*. OECD, Paris.

OECD.(2010): *PISA 2009 Results: Overcoming Social Background: Equity in Learning Opportunities and Outcomes (Volume II)*, OECD, Paris.

OECD (2012), *PISA 2012 Results What makes schools successful? Resources, policies and practices. Volume IV*, OECD Paris,

O'MALLEY, B. (2010): “Education under attack”, UNESCO.
OPDENAKKER, M. C., VAN DAMME, J. (2006): “Differences between secondary schools: A study about school context, group composition, school practice, and school effects with special attention to public and Catholic schools and types of schools”, *School Effectiveness and School Improvement* 17(1), 87-117.
OSPINA PLAZA, N., GIMÉNEZ ESTEBAN, G. (2009), “La violencia en Latinoamérica y sus efectos sobre la inversión y la educación”, *Estudios de Economía Aplicada*, 27-3, 1-22.
OTHMAN, M., MUIJS, D. (2013), “Educational quality differences in a middle-income country: the urban-rural gap in Malaysian primary schools”, *School Effectiveness and School Improvement*, 24 (1), 104-121.
OYARZÚN MÉNDEZ, E. G., MIRANDA ESCOLAR, B. (2011), "La economía rural en Chile: entre la pobreza y el desarrollo", *Estudios de Economía Aplicada*, 29 (1), 31-56.
PIÑEROS, L. J., RODRÍGUEZ, A. (1999): “Los insumos escolares en la educación secundaria y su efecto sobre el rendimiento académico de los estudiantes: Un estudio en Colombia”, *LCSHD Paper Series* 36. Washington, DC: World Bank.
RANGEL, C., LLERAS, C. (2010): “Educational inequality in Colombia: family background, school quality and student achievement in Cartagena”, *International Studies in Sociology of Education* 20(4), 291-317
RIVKIN, S. G., HANUSHEK, E. A., KAIN, J. F. (2005): “Teachers, Schools and Academic Achievement”, *Econometrica* 73(2), 417-458.
RODRIGUEZ, C., SANCHEZ, F., ARMENTA, A., (2010): “Do Interventions at School Level Improve Educational outcomes? Evidence from a Rural Program in Colombia” *World Development* 38(3), 415-428.
ROSCIGNO, V. J., M. L. CROWLEY. (2001): “Rurality, Institutional Disadvantage, and Achievement/ Attainment”, *Rural Sociology* 66(2), 268-293.
RUBIN, D. B. (1987): *Multiple imputation for nonresponse in surveys*, New York, Wiley.
SANDER, W. (1996): “Catholic grade schools and academic achievement”, *The Journal of Human Resources* 31 (3), 540-548.
SANTOS, M.E. (2007): “Calidad de la educación en Argentina: Determinantes y distribución utilizando los resultados de PISA 2000”, *Bienestar y política social* 3(1), 79-109.
SOMERS, M. A., MCEWAN, P. J., WILLMS, J. D. (2004): “How effective are private schools in Latin America?” *Comparative Education Review* 48(1), 48–69.
SMITH, J., NAYLOR, R. A. (2005): “Schooling effects on subsequent university performance: evidence for the UK university population”, *Economics of Education Review* 24, 549-562.
STEVANS, L. K., SESSIONS, D. N. (2000): “Private/public school choice and student performance revisited” *Education Economics* 8 (2), 169-184.
TAYYABA, S. (2012): “Rural-urban gaps in academic achievement, schooling conditions, student, and teachers' characteristics in Pakistan”, *International Journal of Educational Management* 26(1), 6 – 26.
UNESCO (2011): *The hidden crisis: Armed conflict and education*, ISBN 978–92–3–104191–4.
VANDERNBERGHE, V., ROBIN, S. (2004): “Evaluating the effectiveness of private education across countries: A comparison of methods”, *Labor Economics* 11(4), 487-506.
WARD, A.P., MURRAY, L.W. (1985): “Factors affecting performance of New Mexico high school students”, *Meeting of the Rocky Mountain Educational Research Association*, Las Cruces. (Eric Document Reproduction Service No. 271 266)
WAR CHILD. (2007): “Child Soldiers: The Shadow of Their Existence”, mimeo.
WILLIAMS, J.H. (2005): “Cross-National Variations in Rural Mathematics Achievement: A Descriptive Overview”, *Journal of Research in Rural Education* 20(5), 1-18.
WOßMANN, L. (2003): “Schooling resources, educational institutions and student performance: the international evidence”, *Oxford Bulletin of Economics and Statistics* 65 (2), 117-170.

WOßMANN, L. (2010): “Families, schools and primary-school learning: evidence for Argentina and Colombia in an international perspective”, *Applied Economics* 42, 2645–2665.

YOUNG, D.J. (1998): “Rural and urban differences in student achievement in science and mathematics: A multilevel analysis” *School Effectiveness and School Improvement* 9(4), 386-418.

ZHANG, L., LEE, K.A. (2011): “Decomposing achievement gaps among OECD countries”. *Asia Pacific Education Review* 12 (3), 463-474.

APPENDIX. Table A.1. Summary statistics – 2006 sample

	All students		Boys		Girls	
	Urban	Rural	Urban	Rural	Urban	Rural
Mathematics	385.27	363.275	396.3966	372.677	375.1823	355.8683
	[83.7826]	[72.4515]	[84.2777]	[72.751]	[82.0868]	[71.3888]
Reading	405.0316	377.9651	398.9743	364.915	410.5233	388.2452
	[102.598	[90.1352]	[103.1059]	[90.030	[101.8872	[88.9294]
Science	400.7836	383.7561	405.0225	386.646	396.9404	381.4794
	[79.0615]	[73.2607]	[81.4013]	[75.839	[76.7253]	[71.1246]
Female	0.5245	0.5594	0	0	1	1
	[.4996]	[.4966]	[0]	[0]	[0]	[0]
Age	15.8392	15.8577	15.8464	15.8423	15.8327	15.8698
	[.2919]	[.2869]	[.2907]	[.2894]	[.293]	[.2846]
Computer	0.4025	0.2127	0.4119	0.1856	0.394	0.234
	[.4906]	[.4093]	[.4925]	[.3891]	[.4889]	[.4236]
Between 0 and 10 books	0.2444	0.4132	0.2659	0.4523	0.2248	0.3823
	[.4299]	[.4926]	[.4421]	[.4981]	[.4177]	[.4862]
Between 11 and 25 books	0.3031	0.2938	0.2883	0.2952	0.3165	0.2928
	[.4597]	[.4557]	[.4533]	[.4564]	[.4654]	[.4553]
Between 26 and 100 books	0.3113	0.213	0.3011	0.1845	0.3205	0.2355
	[.4632]	[.4096]	[.459]	[.3881]	[.467]	[.4246]
More than 100 books	0.1412	0.08	0.1447	0.0681	0.1381	0.0893
	[.3484]	[.2713]	[.352]	[.2521]	[.3452]	[.2854]
Mother's education	10.1746	8.2042	10.4097	8.2448	9.9615	8.1723
	[4.4232]	[4.4422]	[4.4582]	[4.4048]	[4.383]	[4.4737]
Father's education	10.5546	8.2056	10.6643	8.1067	10.4552	8.2835
	[4.6873]	[4.6322]	[4.6485]	[4.59]	[4.7228]	[4.6663]
Economic, social and cultural status index	-0.6101	-1.3256	-0.5991	-1.387	-0.62	-1.2771
	[1.1117]	[1.1519]	[1.1141]	[1.1237]	[1.11]	[1.172]
Urban location	1	0	1	0	1	0
	[0]	[0]	[0]	[0]	[0]	[0]
Proportion of public funding	41.917	43.3376	43.5606	49.3401	40.4269	38.6092
	[36.4284]	[38.8392]	[36.3033]	[38.955	[36.4991]	[38.1087]
School size	2030.703	1385.661	2114.988	1370.47	1954.288	1397.623
	[1376.11	[1202.804]	[1357.067]	[1254.1	[1389.557	[1161.31
Student-teacher ratio	26.2253	22.4739	26.655	21.4593	25.8358	23.2731
	[9.3036]	[12.0797]	[8.7448]	[12.754	[9.7715]	[11.4648]
School resp. for curriculum and assessment	-0.1207	-0.4229	-0.1281	-0.459	-0.1139	-0.3944
	[.8317]	[.7622]	[.8556]	[.7328]	[.8099]	[.7839]
Quality of school's educational resources	-0.8916	-1.4367	-0.9371	-1.3875	-0.8502	-1.4754
	[1.068]	[.8826]	[1.0986]	[.9058]	[1.0385]	[.8624]
Teacher shortage	-0.1471	0.6911	-0.1027	0.7191	-0.1873	0.669
	[.9963]	[1.0872]	[.9843]	[1.1336]	[1.0059]	[1.0493]
Observations	1600	1550	751	698	849	852

Average values. Standard deviation in brackets. Source: Based on 2006 PISA data.

Table A.2. Summary statistics – 2009 sample

	All students		Boys		Girls	
	Urban	Rural	Urban	Rural	Urban	Rural
Mathematics	398.4698	369.3615	415.9158	384.9922	383.0289	353.6735
	[70.6169]	[64.8112]	[69.0654]	[66.7878]	[68.3595]	[58.7524]
Reading	434.676	398.8309	428.9826	393.2499	439.715	404.4324
	[81.7007]	[75.1616]	[80.7324]	[75.8623]	[82.2403]	[74.0679]
Science	418.9041	392.6064	430.9264	402.2538	408.2634	382.9237
	[76.2291]	[70.3217]	[73.8926]	[71.6185]	[76.7006]	[67.6612]
Female	0.5305	0.4991	0	0	1	1
	[.4991]	[.5001]	[0]	[0]	[0]	[0]
Age	15.8328	15.8603	15.8465	15.8672	15.8206	15.8534
	[.2783]	[.2862]	[.2843]	[.291]	[.2723]	[.2812]
Computer	0.6032	0.3408	0.6246	0.3622	0.5843	0.3193
	[.4893]	[.4741]	[.4844]	[.4809]	[.493]	[.4665]
Between 0 and 10 books	0.2298	0.4096	0.2104	0.4325	0.247	0.3867
	[.4208]	[.4919]	[.4077]	[.4957]	[.4314]	[.4872]
Between 11 and 25 books	0.3062	0.3314	0.3184	0.3142	0.2954	0.3487
	[.461]	[.4709]	[.466]	[.4644]	[.4563]	[.4768]
Between 26 and 100 books	0.3365	0.1898	0.3544	0.1786	0.3208	0.2011
	[.4726]	[.3922]	[.4785]	[.3832]	[.4669]	[.401]
More than 100 books	0.1274	0.0691	0.1169	0.0747	0.1368	0.0635
	[.3335]	[.2537]	[.3214]	[.263]	[.3437]	[.244]
Mother's education	10.7556	8.2895	10.8897	8.6616	10.6368	7.9161
	[4.4282]	[4.5104]	[4.3333]	[4.5182]	[4.5083]	[4.4738]
Father's education	10.8122	8.1851	10.9156	8.3382	10.7207	8.0314
	[4.5615]	[4.6334]	[4.5127]	[4.6882]	[4.6034]	[4.575]
Economic, social and cultural status index	-0.6546	-1.4893	-0.6196	-1.4184	-0.6856	-1.5605
	[1.1571]	[1.2084]	[1.1304]	[1.2084]	[1.1796]	[1.2048]
Urban location	1	0	1	0	1	0
	[0]	[0]	[0]	[0]	[0]	[0]
Proportion of public funding	55.5739	71.5332	55.5697	71.962	55.5776	71.1029
	[36.0847]	[30.5149]	[35.4149]	[30.1526]	[36.6763]	[30.8835]
School size	1642.016	1094.338	1701.08	1084.349	1589.741	1104.363
	[1175.307]	[757.9476]	[1224.088]	[759.5361]	[1128.114]	[756.6004]
Student-teacher ratio	28.2574	25.2572	27.93	25.0447	28.5473	25.4704
	[10.7301]	[8.143]	[10.9725]	[8.2287]	[10.5052]	[8.0547]
School resp. for curriculum and assessment	-0.1708	-0.2695	-0.1706	-0.2952	-0.171	-0.2437
	[.8422]	[.825]	[.8304]	[.7922]	[.8527]	[.8563]
Quality of school's educational resources	-0.7705	-1.5443	-0.772	-1.4818	-0.7692	-1.607
	[1.0027]	[1.1048]	[.9969]	[1.092]	[1.008]	[1.1146]
Teacher shortage	-0.0475	0.3275	-0.0978	0.3272	-0.0029	0.3278
	[1.0477]	[1.005]	[.9595]	[.9936]	[1.1182]	[1.0168]
Observations	3729	1987	1693	995	2036	992

Average values. Standard deviation in brackets. Source: Based on 2009 PISA data.

Table A.3. Summary statistics – 2012 sample

	All students		Boys		Girls	
	Urban	Rural	Urban	Rural	Urban	Rural
Mathematics	393.1097	368.7276	409.9854	378.6738	378.5795	359.1551
	[71.2468]	[64.0657]	[72.2569]	[66.5023]	[67.0571]	[60.1407]
Reading	424.0225	391.7693	417.6257	379.1883	429.5303	403.8775
	[74.8297]	[76.0721]	[77.7765]	[78.6295]	[71.7547]	[71.5191]
Science	413.2202	391.89	425.1271	397.9957	402.9683	386.0137
	[70.3473]	[68.4036]	[72.5588]	[70.926]	[66.7172]	[65.398]
Female	0.5373	0.5096	0	0	1	1
	[.4987]	[.5001]	[0]	[0]	[0]	[0]
Age	15.8524	15.8586	15.855	15.8555	15.8501	15.8615
	[.2854]	[.291]	[.2821]	[.2863]	[.2882]	[.2957]
Computer	0.7521	0.4979	0.7616	0.4738	0.7439	0.521
	[.4318]	[.5002]	[.4262]	[.4997]	[.4366]	[.4999]
Between 0 and 10 books	0.2871	0.4409	0.2896	0.458	0.285	0.4246
	[.4525]	[.4967]	[.4537]	[.4986]	[.4515]	[.4946]
Between 11 and 25 books	0.3245	0.3214	0.3083	0.3021	0.3385	0.34
	[.4682]	[.4672]	[.4619]	[.4595]	[.4733]	[.474]
Between 26 and 100 books	0.2795	0.1866	0.2831	0.2044	0.2764	0.1695
	[.4488]	[.3897]	[.4506]	[.4036]	[.4473]	[.3754]
More than 100 books	0.1089	0.051	0.119	0.0355	0.1001	0.066
	[.3115]	[.2202]	[.3239]	[.1851]	[.3002]	[.2485]
Mother's education	10.8922	8.5923	11.2261	9.0865	10.6048	8.1166
	[4.0216]	[4.4747]	[3.8588]	[4.4931]	[4.1359]	[4.4079]
Father's education	10.821	8.5783	11.1995	8.7799	10.4951	8.3842
	[4.2302]	[4.6184]	[4.1145]	[4.5596]	[4.3016]	[4.6691]
Economic, social and cultural status index	-0.8473	-1.6244	-0.7659	-1.5728	-0.9174	-1.674
	[1.0312]	[1.1859]	[1.026]	[1.1796]	[1.0307]	[1.1907]
Urban location	1	0	1	0	1	0
	[0]	[0]	[0]	[0]	[0]	[0]
Proportion of public funding	78.4346	90.1249	78.3649	90.5593	78.4946	89.7068
	[36.6835]	[24.2836]	[37.0394]	[23.0462]	[36.3808]	[25.4267]
School size	1668.58	1197.01	1686.906	1173.607	1652.801	1219.534
	[1154.222]	[949.4633]	[1138.999]	[913.7401]	[1167.144]	[982.7256]
Student-teacher ratio	28.6614	25.6875	28.6073	25.8706	28.708	25.5114
	[9.1732]	[8.2957]	[8.9738]	[8.1252]	[9.3429]	[8.4585]
School resp. for curriculum and assessment	-0.037	-0.1111	-0.0309	-0.1346	-0.0423	-0.0885
	[.8311]	[.8912]	[.8218]	[.8943]	[.8391]	[.8882]
Quality of school's educational resources	-1.1082	-1.5613	-1.1297	-1.6332	-1.0898	-1.4921
	[1.1899]	[.9814]	[1.1719]	[.9695]	[1.2051]	[.9885]
Teacher shortage	0.4417	1.005	0.4367	1.0168	0.446	0.9936
	[1.2956]	[1.4078]	[1.2701]	[1.4196]	[1.3175]	[1.3972]
Observations	5147	1416	2376	666	2771	750

Average values. Standard deviation in brackets. Source: Based on 2012 PISA data.

Table A.4. Estimates of the educational production function for Mathematics

Maths 2006	All students		Boys		Girls	
	Urban	Rural	Urban	Rural	Urban	Rural
Female	-21.062***	-25.149***				
Age	24.853**	24.250**	15.931	18.022	34.748***	28.681*
Computer	23.856**	15.154	25.662***	34.161***	22.233*	1.203
Between 11 and 25 books	23.649**	17.472***	20.206*	21.058***	29.284**	13.432
Between 26 and 100 books	24.619**	20.962**	25.511*	20.196*	27.395**	17.602
More than 100 books	31.328***	10.321	32.988**	17.758	34.932**	2.717
Mother's education	1.004	-0.861	1.397	-0.293	0.539	-1.119
Father's education	0.782	0.098	0.666	-0.400	0.965	0.200
ESCS	10.025*	13.024*	4.609	10.032	14.968*	15.776**
Proportion of public funding	-0.079	-0.211*	-0.073	-0.069	-0.062	-0.312**
School size	0.025	-0.001	0.034	-0.006	0.018	-0.001
School size^2	-0.000	-0.000	-0.000	0.000	-0.000	-0.000
Student-teacher ratio	-1.237	-1.517	-2.061	0.502	-1.213	-2.641
Student-teacher ratio^2	0.017	0.068	0.059	0.006	-0.001	0.106*
School's responsibility	13.518**	1.458	14.487**	4.262	10.838*	-0.680
Quality ed. Resources	13.091***	-1.793	15.956***	2.206	9.655*	-3.587
Teacher shortage	1.554	-5.435	-0.281	0.499	1.647	-9.319*
Intercept	-26.728	12.576	95.714	90.966	-189.661	-65.193
R^2	0.2464	0.1498	0.2340	0.1287	0.2678	0.1755
Observations	1600	1538	751	694	849	844

Maths 2009	All students		Boys		Girls	
	Urban	Rural	Urban	Rural	Urban	Rural
Female	-30.452***	-29.016***				
Age	10.750*	25.540***	18.141**	18.442**	3.076	32.959***
Computer	7.939	11.903*	3.100	8.888	11.646**	15.458**
Between 11 and 25 books	16.592***	11.769**	17.872**	16.816*	16.334***	7.084
Between 26 and 100 books	33.781***	29.069***	38.487***	27.574***	29.331***	31.499***
More than 100 books	31.892***	18.461**	36.370***	21.720	26.644***	14.893*
Mother's education	0.747	1.165	0.893	1.157	0.618	1.279
Father's education	0.281	-0.311	0.604	-0.132	-0.061	-0.498
ESCS	10.931***	7.328	7.663	8.042	13.851***	6.430
Proportion of public funding	-0.356***	-0.123	-0.414***	-0.220	-0.299***	-0.024
School size	0.003	-0.002	0.009	0.003	-0.004	-0.009
School size^2	0.000	0.000	-0.000	-0.000	0.000	0.000
Student-teacher ratio	1.089	0.386	0.552	1.411	1.372	-1.061
Student-teacher ratio^2	-0.021	-0.007	-0.015	-0.026	-0.025	0.020
School's responsibility	-1.045	0.360	-3.785	4.464	1.639	-3.394
Quality ed. Resources	9.337***	5.125	9.792***	7.954*	8.268**	1.869
Teacher shortage	3.747	-3.504	1.577	-0.950	4.528	-6.544
Intercept	224.215**	-18.925	104.251	90.213	321.921***	-157.781
R^2	0.3349	0.2161	0.3019	0.1987	0.3169	0.1772
Observations	3673	1987	1668	995	2005	992

Table A.4. Estimates of the educational production function for Mathematics (cont.)

Maths 2012	All students		Boys		Girls	
	Urban	Rural	Urban	Rural	Urban	Rural
Female	-28.775***	-20.401***				
Age	13.178***	10.766	17.930**	-1.751	9.282	24.940***
Computer	22.051***	16.287***	15.986***	19.804***	26.707***	14.586**
Between 11 and 25 books	7.949***	7.876*	13.616**	7.193	3.380	8.898
Between 26 and 100 books	16.955***	13.266	24.195***	14.440	10.840*	13.443
More than 100 books	21.847***	12.312	28.648***	32.333	16.147**	2.294
Mother’s education	-0.358	0.214	-1.292*	0.087	0.471	0.321
Father’s education	-0.955**	-1.771***	-1.268**	-2.838***	-0.701	-0.332
ESCS	14.439***	13.090***	17.644***	13.476***	11.731***	10.941**
Proportion of public funding	-0.528***	-0.188	-0.523***	-0.169	-0.517***	-0.209
School size	0.019***	-0.003	0.018**	-0.003	0.020***	-0.000
School size^2	-0.000**	0.000	-0.000*	0.000	-0.000**	0.000
Student-teacher ratio	-1.765*	1.124	-0.762	1.725	-2.531**	0.452
Student-teacher ratio^2	0.011	-0.011	0.000	-0.020	0.019	-0.002
School’s responsibility	-2.532	-0.239	-0.192	-3.984	-4.448	2.999
Quality ed. Resources	1.578	5.373	3.625	8.936**	-0.373	1.309
Teacher shortage	-3.815	6.418**	-4.250	8.126***	-3.870	4.129
Intercept	262.621***	227.987**	187.615	428.062***	292.967***	-25.903
R^2	0.2905	0.2004	0.2535	0.2052	0.2752	0.1970
Observations	5147	1416	2376	666	2771	750

Table A.5. Estimates of the educational production function for Reading

Reading 2006	All students		Boys		Girls	
	Urban	Rural	Urban	Rural	Urban	Rural
Female	11.498	14.377**				
Age	30.317**	11.069	31.973*	11.992	29.746**	9.874
Computer	27.197**	16.467	21.110	23.799*	31.434**	9.961
Between 11 and 25 books	39.686***	25.004***	47.481***	32.560**	34.748*	19.256*
Between 26 and 100 books	51.631***	34.397***	68.080***	42.506***	41.291**	24.122
More than 100 books	54.847***	19.911	78.151***	38.553**	39.937*	5.060
Mother’s education	-1.655	-0.767	-1.346	1.213	-1.792	-2.443*
Father’s education	-0.087	0.791	0.619	-1.579	-0.925	2.456
ESCS	20.778***	11.355	7.592	7.885	32.745***	16.081*
Proportion of public funding	0.003	-0.303**	-0.103	-0.193	0.119	-0.343**
School size	0.034	0.028	0.033	0.032	0.039	0.021
School size^2	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
Student-teacher ratio	-1.302	-3.677*	-2.327	-2.626	-1.155	-3.735
Student-teacher ratio^2	0.014	0.081	0.073	0.029	-0.017	0.106
School’s responsibility	15.042*	-2.645	18.228*	0.381	10.685	-3.546
Quality ed. Resources	9.572	-11.415**	11.352	-7.482	7.225	-13.030**
Teacher shortage	2.924	-10.567	-0.747	-5.858	4.694	-13.545*
Intercept	-102.887	207.336	-161.966	186.844	-54.318	245.614
R^2	0.2063	0.1338	0.2121	0.1146	0.2311	0.1615
Observations	1600	1538	751	694	849	844

Reading 2009	All students		Boys		Girls	
	Urban	Rural	Urban	Rural	Urban	Rural
Female	12.989***	13.620***				
Age	4.501	17.617***	8.041	10.682	-0.632	25.282***
Computer	14.166**	19.031***	15.730*	17.022**	12.376*	21.865***
Between 11 and 25 books	20.039***	11.126*	25.042***	14.848*	16.088**	7.788
Between 26 and 100 books	25.466***	26.871***	29.794***	20.954**	20.775**	34.385***
More than 100 books	26.215***	21.405**	30.059***	27.327**	19.702**	16.212
Mother’s education	0.149	1.579**	0.418	2.042**	-0.138	1.306
Father’s education	-1.005	-0.395	-0.810	-0.654	-1.401	-0.060
ESCS	17.046***	4.906	10.430**	2.290	23.856***	6.225
Proportion of public funding	-0.377***	-0.169	-0.423***	-0.373**	-0.338***	0.035
School size	-0.002	0.018	0.007	0.032*	-0.012	0.004
School size^2	0.000	-0.000	-0.000	-0.000	0.000	0.000
Student-teacher ratio	0.759	1.606	-0.473	2.065	1.682	0.724
Student-teacher ratio^2	-0.017	-0.028	-0.002	-0.036	-0.028*	-0.013
School’s responsibility	-1.990	-1.596	-3.115	2.159	-1.016	-4.909
Quality ed. Resources	12.126***	6.635	15.063***	8.760	8.696**	4.318
Teacher shortage	2.069	-4.659	1.903	-3.157	0.947	-7.101
Intercept	375.245***	81.654	320.103*	190.702	475.970***	-24.646
R^2	0.2777	0.1480	0.2669	0.1630	0.2961	0.1517
Observations	3673	1987	1668	995	2005	992

Table A.5. Estimates of the educational production function for Reading (cont.)

Reading 2012	All students		Boys		Girls	
	Urban	Rural	Urban	Rural	Urban	Rural
Female	14.913***	23.509***				
Age	15.119**	6.713	23.302**	-8.677	7.822	23.092**
Computer	22.170***	17.667***	19.879**	19.360**	23.555***	16.787*
Between 11 and 25 books	10.191***	11.545*	16.412***	9.162	5.195	14.296*
Between 26 and 100 books	16.236***	21.219**	21.376***	23.129*	11.652**	20.410**
More than 100 books	23.026***	9.359	28.546***	22.109	18.239**	2.829
Mother's education	-0.155	-0.021	-1.006	0.033	0.567	-0.233
Father's education	-1.375***	-2.249***	-1.793*	-2.777**	-1.013	-1.371
ESCS	18.371***	17.500***	21.506***	17.080***	15.871***	17.025***
Proportion of public funding	-0.535***	-0.293	-0.488***	-0.259	-0.561***	-0.370*
School size	0.018**	-0.024	0.013	-0.027	0.022***	-0.019
School size^2	-0.000*	0.000	-0.000	0.000	-0.000**	0.000
Student-teacher ratio	-0.302	2.380	0.422	4.114*	-0.804	1.229
Student-teacher ratio^2	-0.004	-0.026	-0.011	-0.048*	0.001	-0.010
School's responsibility	-2.727	-3.479	-1.283	-8.490	-3.945	0.796
Quality ed. Resources	1.870	9.094*	4.088	11.616**	-0.114	5.476
Teacher shortage	-3.238	7.070*	-2.984	8.582**	-3.870	4.810
Intercept	216.948**	308.017**	91.131	526.158***	341.388***	85.196
R^2	0.2339	0.2431	0.2249	0.2481	0.2452	0.2293
Observations	5147	1416	2376	666	2771	750

Table A.6. Estimates of the educational production function for Science

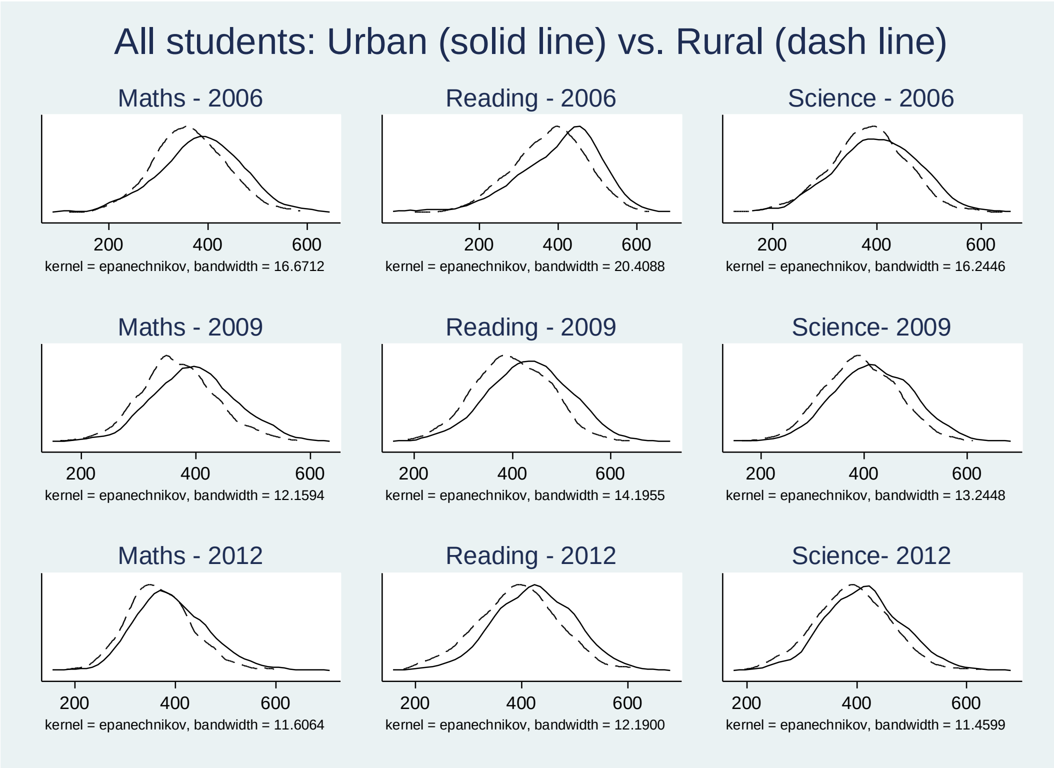
Science 2006	All students		Boys		Girls	
	Urban	Rural	Urban	Rural	Urban	Rural
Female	-8.958	-11.676**				
Age	19.269	12.415	17.985	8.019	21.951**	14.780
Computer	22.002***	11.521	20.453**	18.957	21.732**	3.610
Between 11 and 25 books	19.300**	16.084**	19.524**	31.619***	20.284*	3.099
Between 26 and 100 books	23.829***	27.051***	27.648**	34.675***	21.413*	16.987
More than 100 books	36.365***	25.036**	37.047***	38.043**	37.318***	11.324
Mother's education	-0.208	-1.152	0.045	-1.064	-0.510	-1.176
Father's education	0.047	0.884	0.098	1.040	-0.084	0.577
ESCS	12.185***	9.047*	10.596**	3.167	14.743**	15.049***
Proportion of public funding	-0.263*	-0.199	-0.285	-0.108	-0.230	-0.253*
School size	0.010	0.006	0.020	0.006	0.002	0.007
School size^2	-0.000	-0.000	-0.000	-0.000	0.000	-0.000
Student-teacher ratio	-0.031	-2.651	-2.187	-1.865	1.239	-3.088
Student-teacher ratio^2	-0.010	0.075	0.056	0.033	-0.048	0.101*
School's responsibility	12.422*	-0.075	10.265	7.147	13.873*	-4.074
Quality ed. Resources	6.776	-3.224	6.612	-1.377	7.248	-3.396
Teacher shortage	-0.080	-4.239	-4.937	-2.273	2.309	-5.528
Intercept	96.203	210.128	109.675	267.591	52.819	175.691
R^2	0.2192	0.1081	0.1994	0.0914	0.2558	0.1574
Observations	1600	1538	751	694	849	844

Science 2009	All students		Boys		Girls	
	Urban	Rural	Urban	Rural	Urban	Rural
Female	-20.225***	-17.455***				
Age	-2.067	8.330	7.373	3.307	-12.058	13.859
Computer	13.462**	22.517***	13.571	19.813**	12.959***	26.013***
Between 11 and 25 books	26.378***	16.086**	33.528***	23.419**	19.983**	8.822
Between 26 and 100 books	33.130***	26.978***	42.144***	25.311***	24.737**	29.439***
More than 100 books	32.709***	22.710**	40.434***	25.185**	24.250***	21.915*
Mother's education	0.340	1.520*	0.926	1.982*	-0.183	1.199
Father's education	-0.086	-0.633	0.208	-0.630	-0.473	-0.531
ESCS	10.412***	4.273	2.485	0.893	17.987***	6.659
Proportion of public funding	-0.323***	-0.266*	-0.359**	-0.432***	-0.297***	-0.100
School size	-0.008	-0.009	-0.004	0.002	-0.012	-0.020
School size^2	0.000	0.000	0.000	-0.000	0.000	0.000
Student-teacher ratio	0.712	2.037	0.176	2.752	1.100	0.887
Student-teacher ratio^2	-0.022	-0.033	-0.015	-0.048	-0.027	-0.009
School's responsibility	-0.076	-2.476	0.059	0.137	-0.673	-4.618
Quality ed. Resources	12.201***	3.296	12.429***	5.986	11.205***	0.433
Teacher shortage	1.783	-2.619	1.226	-2.149	1.245	-4.050
Intercept	467.544***	250.578**	302.769**	322.292**	622.422***	154.077
R^2	0.3147	0.1672	0.2787	0.1750	0.3332	0.1571
Observations	3673	1987	1668	995	2005	992

Table A.6. Estimates of the educational production function for Science (cont.)

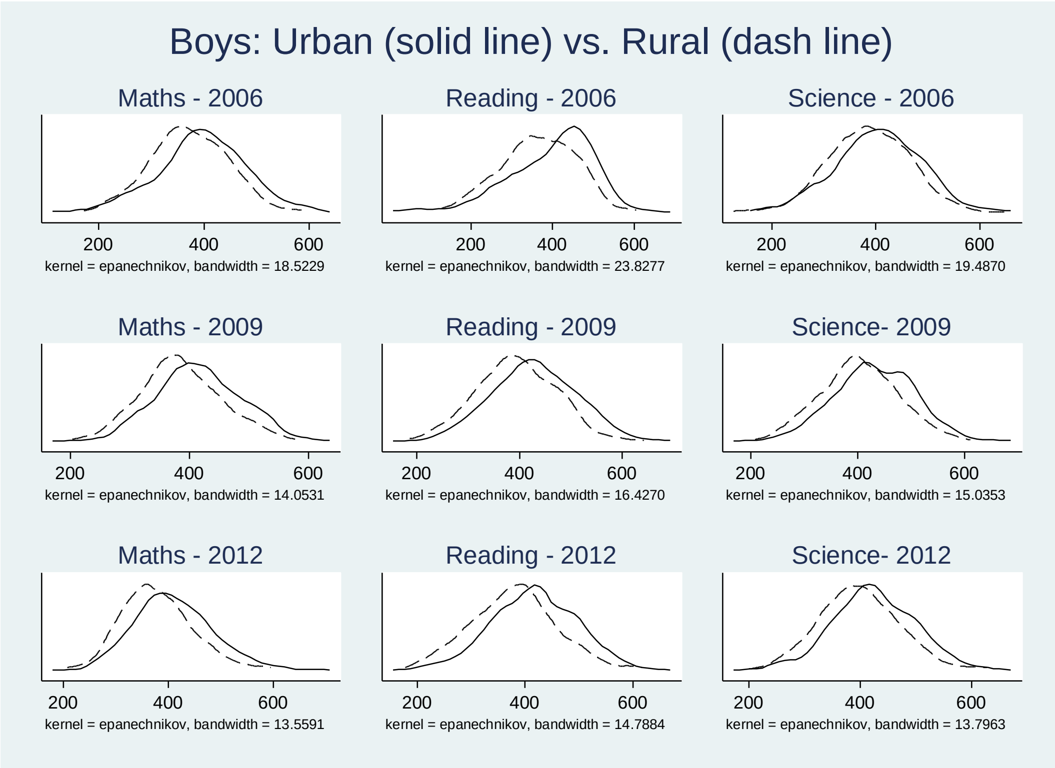
Science 2012	All students		Boys		Girls	
	Urban	Rural	Urban	Rural	Urban	Rural
Female	-19.928***	-13.009**				
Age	10.939*	11.714*	16.575*	-0.186	6.228	24.986**
Computer	20.569***	11.198*	14.429**	13.773*	25.103***	9.889
Between 11 and 25 books	9.008***	9.714	13.474**	4.199	5.537	15.590**
Between 26 and 100 books	14.460***	10.597	20.315***	13.956	9.502*	8.191
More than 100 books	24.364***	8.191	31.792***	25.222	17.900**	1.111
Mother's education	-1.107	-0.257	-2.148**	-0.875	-0.195	0.349
Father's education	-1.073***	-2.483***	-1.465**	-2.717**	-0.751	-1.797*
ESCS	16.576***	17.798***	19.764***	19.282***	13.980***	14.552***
Proportion of public funding	-0.359***	-0.120	-0.368***	-0.146	-0.334***	-0.113
School size	0.018**	-0.017	0.017*	-0.020	0.019**	-0.015
School size^2	-0.000	0.000	-0.000	0.000	-0.000	0.000
Student-teacher ratio	-2.168**	2.336	-1.130	3.726	-2.981***	1.349
Student-teacher ratio^2	0.016	-0.031	0.006	-0.050*	0.024*	-0.015
School's responsibility	-1.866	-5.313	0.287	-10.042*	-3.647	-1.572
Quality ed. Resources	3.153	5.197	5.036	9.218	1.353	0.853
Teacher shortage	-2.087	5.533*	-2.808	7.084**	-1.988	3.251
Intercept	320.350***	234.759**	234.476	420.674***	369.991***	2.164
R^2	0.2176	0.1924	0.1988	0.2254	0.2161	0.1832
Observations	5147	1416	2376	666	2771	750

Figure A1. Distribution of students' educational performance in rural and urban schools – All



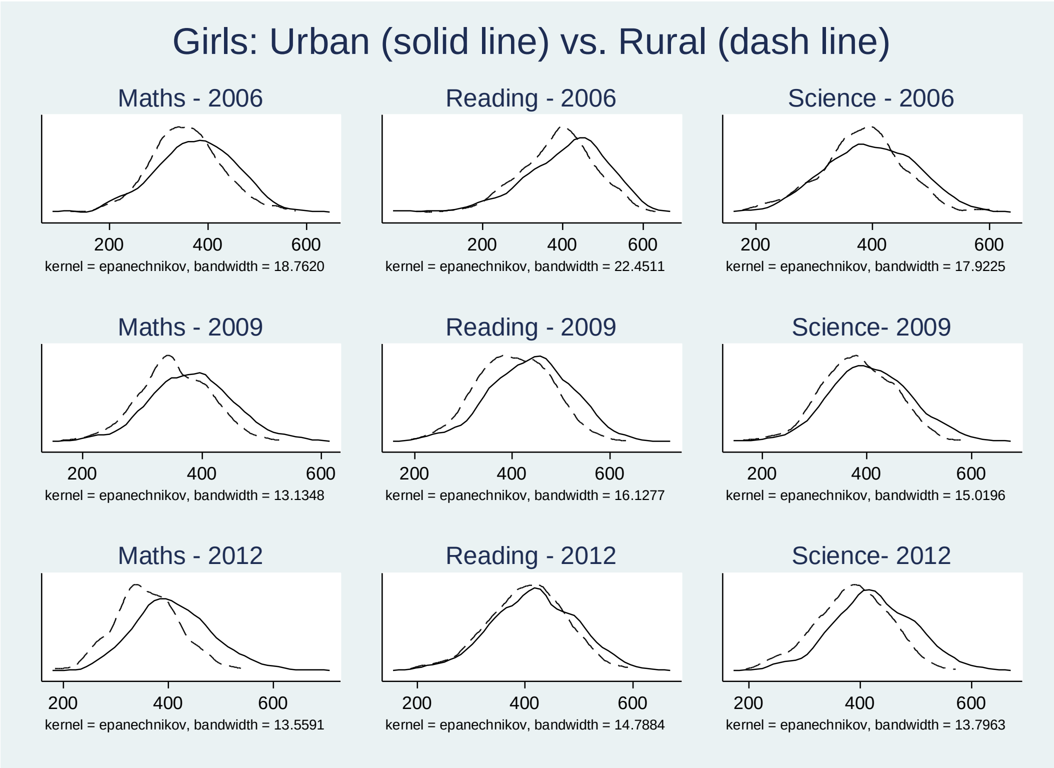
Source: Based on 2006, 2009 and 2012 PISA data.

Figure A2. Distribution of students' educational performance in rural and urban schools - Boys



Source: Based on 2006, 2009 and 2012 PISA data.

Figure A3. Distribution of students' educational performance in rural and urban schools - Girls



Source: Based on 2006, 2009 and 2012 PISA data.