Neighborhood characteristics around schools and around students’ homes: Which one affects academic performance the most?

Weebob J. Requia (weeberb.requia@fgv.br)  
Fundação Getúlio Vargas Brasília

Luciano Moura Silva  
Fundação Getúlio Vargas Brasília

Article

Keywords: Neighborhood characteristics, academic performance, green spaces, roads, wildfire

Posted Date: May 25th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2920071/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract
Access to high-quality education is essential for achieving social and economic development, yet academic performance remains a challenge in many low- and middle-income countries, such as Brazil. The neighborhood environment is increasingly recognized as an important factor that can influence academic performance. The existing literature on this topic is limited. Most studies have focused either on the exposure surrounding a student’s school or their home. To address this gap, this study aimed to evaluate the association between neighborhood characteristics around both the student’s school and home in the Federal District (FD), Brazil, focusing on green spaces (NDVI and quantity of green spaces), roads (length of roads), and wildfres as potential predictors of academic performance.
We analyzed 344,175 students who were enrolled in the public schools (256 schools) of the FD from 2017 to 2020. We employed a mixed-effects regression model to investigate the relationship between individual-level academic performance and the neighborhood characteristics, while controlling for temporal, socioeconomic, and school-specific variables. Our results suggest that neighborhood characteristics represented by green spaces, roads, and wildfres, can have a significant impact on academic performance. We found that some neighborhood characteristics had a stronger association with academic performance when they were located around schools, compared to when they were located around students’ homes. For example, the coefficient for NDVI within the buffer of 250m and 500m surrounding schools was estimated to be 2.48 (95%CI: 2.33; 2.63) and 0.59 (95%CI: 0.47; 0.71), respectively, while for NDVI surrounding homes, it was 0.23 (95%CI: 0.15; 0.30) and 0.25 (95%CI: 0.17; 0.33) for the buffer of 250m and 500m, respectively. Our study provides valuable insights into the potential role of neighborhood characteristics in promoting academic performance, which could inform urban planning and public policy aimed at improving educational outcomes.

1. INTRODUCTION
Neighborhood characteristics have been associated with numerous impacts on quality of life, including the increase risk of morbidity and mortality\(^1\)–\(^3\), facilitation of social inclusion/cohesion\(^4\),\(^5\), and the development of physical and mental health\(^6\),\(^7\). A large body of studies in this field has focused on the connection between neighborhood environment and academic performance (as a proxy of cognitive functions). These studies have explored different aspects of neighborhood characteristics and reported various results, including environmental characteristics associated positively [e.g., the presence of green spaces in the neighborhood −\(^8\)] and negatively [e.g., the high density of roads in the neighborhood −\(^9\)] with academic performance.

It is important to note that the exposure to environmental factors of neighborhoods varies by several geographical aspects, including information on people’s location\(^10\). For example, the effects of road density (neighborhood characteristic as a proxy for air pollution) on children may differ depending on whether they are at home\(^11\), at school\(^12\), or during their daily commute to school\(^13\)–\(^15\).

Despite the importance of neighborhood characteristics in shaping academic performance, the existing literature on this topic is limited. Most studies have focused either on the exposure surrounding a student’s school or their home. Only one study, conducted by Dadvand et al. (2015), has examined both home and school exposure, but it only considered green spaces. To address this gap in the literature, our study aims to evaluate the association between neighborhood characteristics around both the student’s school and home in the Federal District, Brazil, focusing on green spaces, roads, and wildfres as potential predictors of academic performance.

2. RESULTS
We analyzed 344,175 students who were enrolled in the public schools (256 schools) of the FD from 2017 to 2020. The average grade for student-level academic performance was 6.29 points, with a standard deviation of 1.72 points. The academic performance of the students varied greatly, with grades ranging from 0 to 10. The descriptive statistics for academic performance, as well as the socioeconomic status (SES) variables and neighborhood characteristics surrounding schools and homes, are presented in Table 1.
Table 1
Descriptive statistics of the grades (student-level academic performance), SES variables, and neighborhood characteristics surrounding schools and homes.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Minimum</th>
<th>1st quartile</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>3rd quartile</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grades</td>
<td>0.00</td>
<td>5.54</td>
<td>6.29</td>
<td>1.72</td>
<td>7.30</td>
<td>10.00</td>
</tr>
<tr>
<td>Income</td>
<td>0.00</td>
<td>155.00</td>
<td>199.10</td>
<td>75.04</td>
<td>233.00</td>
<td>906.00</td>
</tr>
<tr>
<td>Education</td>
<td>0.00</td>
<td>3.00</td>
<td>8.75</td>
<td>10.35</td>
<td>10.00</td>
<td>108.00</td>
</tr>
<tr>
<td><strong>Surrounding schools</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDVI 250m</td>
<td>0.01</td>
<td>0.08</td>
<td>0.13</td>
<td>0.06</td>
<td>0.18</td>
<td>0.38</td>
</tr>
<tr>
<td>NDVI 500m</td>
<td>0.00</td>
<td>0.11</td>
<td>0.15</td>
<td>0.05</td>
<td>0.18</td>
<td>0.37</td>
</tr>
<tr>
<td>NDVI 750m</td>
<td>0.00</td>
<td>0.12</td>
<td>0.16</td>
<td>0.06</td>
<td>0.20</td>
<td>0.39</td>
</tr>
<tr>
<td>NDVI 1km</td>
<td>0.00</td>
<td>0.13</td>
<td>0.17</td>
<td>0.07</td>
<td>0.21</td>
<td>0.39</td>
</tr>
<tr>
<td>Quant.Green 250m</td>
<td>0.00</td>
<td>0.00</td>
<td>20,819.00</td>
<td>39,530.36</td>
<td>27,688.00</td>
<td>196,247.00</td>
</tr>
<tr>
<td>Quant.Green 500m</td>
<td>0.00</td>
<td>0.00</td>
<td>129,684.45</td>
<td>180,859.86</td>
<td>189,624.00</td>
<td>784,438.00</td>
</tr>
<tr>
<td>Quant.Green 750m</td>
<td>0.00</td>
<td>15,257.00</td>
<td>378,235.42</td>
<td>424,639.04</td>
<td>597,447.00</td>
<td>1,756,850.00</td>
</tr>
<tr>
<td>Quant.Green 1km</td>
<td>0.00</td>
<td>150,623.00</td>
<td>819,232.96</td>
<td>769,886.16</td>
<td>1,274,760.00</td>
<td>3,126,880.00</td>
</tr>
<tr>
<td>Roads 250m</td>
<td>0.00</td>
<td>2,332.00</td>
<td>4,596.00</td>
<td>3,042.34</td>
<td>6,800.00</td>
<td>12,343.00</td>
</tr>
<tr>
<td>Roads 500m</td>
<td>0.00</td>
<td>10,135.00</td>
<td>18,171.00</td>
<td>11,431.24</td>
<td>26,902.00</td>
<td>47,202.00</td>
</tr>
<tr>
<td>Roads 750m</td>
<td>0.00</td>
<td>18,817.00</td>
<td>38,424.00</td>
<td>24,410.02</td>
<td>57,884.00</td>
<td>101,384.00</td>
</tr>
<tr>
<td>Roads 1km</td>
<td>0.00</td>
<td>30,970.00</td>
<td>63,107.00</td>
<td>40,494.22</td>
<td>97,102.00</td>
<td>159,040.00</td>
</tr>
<tr>
<td>Wildfires 250m</td>
<td>0.00</td>
<td>0.00</td>
<td>0.03</td>
<td>0.19</td>
<td>0.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Wildfires 500m</td>
<td>0.00</td>
<td>0.00</td>
<td>0.13</td>
<td>0.43</td>
<td>0.00</td>
<td>4.00</td>
</tr>
<tr>
<td>Wildfires 750m</td>
<td>0.00</td>
<td>0.00</td>
<td>0.32</td>
<td>0.86</td>
<td>0.00</td>
<td>12.00</td>
</tr>
<tr>
<td>Wildfires 1km</td>
<td>0.00</td>
<td>0.00</td>
<td>0.72</td>
<td>1.53</td>
<td>1.00</td>
<td>20.00</td>
</tr>
<tr>
<td><strong>Surrounding homes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDVI 250m</td>
<td>0.06</td>
<td>0.15</td>
<td>0.20</td>
<td>0.07</td>
<td>0.24</td>
<td>0.59</td>
</tr>
<tr>
<td>NDVI 500m</td>
<td>0.06</td>
<td>0.15</td>
<td>0.22</td>
<td>0.07</td>
<td>0.26</td>
<td>0.56</td>
</tr>
<tr>
<td>NDVI 750m</td>
<td>0.01</td>
<td>0.18</td>
<td>0.23</td>
<td>0.07</td>
<td>0.28</td>
<td>0.54</td>
</tr>
<tr>
<td>NDVI 1km</td>
<td>0.00</td>
<td>0.20</td>
<td>0.25</td>
<td>0.06</td>
<td>0.29</td>
<td>0.54</td>
</tr>
<tr>
<td>Quant.Green 250m</td>
<td>0.00</td>
<td>0.00</td>
<td>25,754.00</td>
<td>46,371.74</td>
<td>33,080.00</td>
<td>196,619.00</td>
</tr>
<tr>
<td>Quant.Green 500m</td>
<td>0.00</td>
<td>0.00</td>
<td>153,913.00</td>
<td>195,698.40</td>
<td>249,916.00</td>
<td>786,323.00</td>
</tr>
</tbody>
</table>

Notes: 1 number of homes within the census tract with family income higher than three minimum wages; 2 number of people above 20 years old who is literate within the census tract; 3 NDVI values within the buffers of 500m, 750m, and 1km; 4 Area (m²) of green spaces within the buffers of 250m, 500m, 750m, and 1km; 5 Length of roads (m) within the buffers of 250m, 500m, 750m, and 1km; and 6 Number of wildfire records within the buffers of 250m, 500m, 750m, and 1km.
<table>
<thead>
<tr>
<th>Variables</th>
<th>Minimum</th>
<th>1st quartile</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>3rd quartile</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^4$ Quant. Green 750m</td>
<td>0.00</td>
<td>38,624.00</td>
<td>441,491.00</td>
<td>440,873.50</td>
<td>697,612.00</td>
<td>1,767,200.00</td>
</tr>
<tr>
<td>$^4$ Quant. Green 1km</td>
<td>0.00</td>
<td>227,699.00</td>
<td>923,499.00</td>
<td>781,236.20</td>
<td>1,494,720.00</td>
<td>3,140,220.00</td>
</tr>
<tr>
<td>$^5$ Roads 250m</td>
<td>0.00</td>
<td>472.70</td>
<td>4,032.20</td>
<td>3,517.44</td>
<td>6,982.20</td>
<td>14,462.00</td>
</tr>
<tr>
<td>$^5$ Roads 500m</td>
<td>0.00</td>
<td>1,982.00</td>
<td>14,965.00</td>
<td>12,571.57</td>
<td>25,649.00</td>
<td>48,908.00</td>
</tr>
<tr>
<td>$^5$ Roads 750m</td>
<td>0.00</td>
<td>5,050.00</td>
<td>31,531.00</td>
<td>25,946.65</td>
<td>51,291.00</td>
<td>106,573.00</td>
</tr>
<tr>
<td>$^5$ Roads 1km</td>
<td>0.00</td>
<td>24,023.00</td>
<td>67,530.00</td>
<td>49,177.80</td>
<td>102,657.00</td>
<td>209,662.00</td>
</tr>
<tr>
<td>$^6$ Wildfires 250m</td>
<td>0.00</td>
<td>0.00</td>
<td>0.02</td>
<td>0.15</td>
<td>0.00</td>
<td>4.00</td>
</tr>
<tr>
<td>$^6$ Wildfires 500m</td>
<td>0.00</td>
<td>0.00</td>
<td>0.11</td>
<td>0.36</td>
<td>0.00</td>
<td>7.00</td>
</tr>
<tr>
<td>$^6$ Wildfires 750m</td>
<td>0.00</td>
<td>0.00</td>
<td>0.27</td>
<td>0.60</td>
<td>0.00</td>
<td>9.00</td>
</tr>
<tr>
<td>$^6$ Wildfires 1km</td>
<td>0.00</td>
<td>0.00</td>
<td>0.39</td>
<td>0.73</td>
<td>1.00</td>
<td>12.00</td>
</tr>
</tbody>
</table>

Notes: $^1$ number of homes within the census tract with family income higher than three minimum wages; $^2$ number of people above 20 years old who is literate within the census tract; $^3$ NDVI values within the buffers of 500m, 750m, and 1km; $^4$ Area (m$^2$) of green spaces within the buffers of 250m, 500m, 750m, and 1km; $^5$ Length of roads (m) within the buffers of 250m, 500m, 750m, and 1km; and $^6$ Number of wildfire records within the buffers of 250m, 500m, 750m, and 1km.

Results of the Wilcoxon test indicate statistically significant difference between the neighborhood characteristics (all characteristics defined in our study across all buffers) around schools and homes. The p-value associated with the test was less than 0.05 in all analyses, suggesting the rejection of the null hypothesis that there is no difference between the two samples. Table S1 (supplementary materials) shows the “W” value and the respective p-values of the Wilcoxon test.

The effects of neighborhood characteristics on students’ academic performance varied significantly across type of characteristics, buffer size, and location of the exposure (home or school). After controlling for covariates, the primary analysis showed that higher values of NDVI were associated with better academic performance, except to the buffers of 750m and 1km around schools (Fig. 2). For NDVI within the buffers of 250m and 500m, our results suggest a higher positive association surrounding schools, with an estimated coefficient of 2.48 (95%CI: 2.33; 2.63) for the buffer of 250m and 0.59 (95%CI: 0.47; 0.71) for the buffer of 500m; while for NDVI surrounding homes, we estimated a coefficient of 0.23 (95%CI: 0.15; 0.30) for the buffer of 250m and 0.25 (95%CI: 0.17; 0.33) for the buffer of 500m (Fig. 2).

For the neighborhood characteristic represented by the area of green spaces within the buffers (Fig. 2, charts categorized as “VEG”), we found a positive association in all analyses. However, statistical significance for exposure around schools was only observed in the buffers of 250m, 500m, and 750m. For exposure around homes, a robust association was observed only in the first buffer, 250m. Similar to NDVI, the presence of this greenness index around schools had a greater positive impact on academic performance compared to this greenness index around homes (Fig. 2).

For the characteristic represented by the length of roads, considering the exposure around homes, only the buffer of 250m presented a significant association, suggesting a positive effect. In contrast, the presence of roads surrounding schools across all buffers is negatively associated with student-level academic performance. We found that the highest effects occur in the buffer of 250m, in which an increase of 1km of length of roads around schools was associated with a statistically significant decrease of 0.011 (95%CI: 0.008; 0.013) points in students’ grades (students’ academic performance varies from 0 to 10) (Fig. 2).

Finally, for wildfires around homes, we observed negative effects in the buffers of 250m, 500m, and 1km, but the results were not statistically significant. For wildfires around schools, we observed unexpected results, with coefficients indicating a positive association between the presence of wildfires around schools and academic performance in the buffers of 500m, 750m, and 1km. A negative association was observed only in the buffer of 250m, though the coefficient did not reach statistical significance (Fig. 2).
Table S2 in the supplementary materials presents all the coefficients, as well as the associated random effects (σ2), including the results from the primary and sensitivity analyses. In Figs. 3 and 4, we present the results of the stratified analyses by type of school (urban and rural) and school level (middle and high school). Overall, the type of school and the school level modify the effects of neighborhood characteristics on academic performance, either strengthening or weakening them.

3. DISCUSSION

Our study aimed to evaluate the association between neighborhood characteristics around students’ school and home and academic performance in the FD, Brazil. Our results suggest that neighborhood characteristics, particularly green spaces, roads, and wildfires, can have a significant impact on academic performance.

The presence of green spaces around schools and homes was positively associated with academic performance, with a stronger effect observed around schools than around homes. Our findings are supported by the literature showing that the presence of vegetation may contribute to improved air quality \(^{16,17}\), reduced noise pollution \(^{18}\), and enhanced psychological well-being \(^{6,7}\), all of which may facilitate learning and cognitive function \(^{19,20}\). Our results are in agreement with previous investigations that have also identified the beneficial effects of green spaces on academic performance \(^{21-24}\).

On the other hand, the length of roads surrounding schools was negatively associated with academic performance, particularly in the buffer of 250m around schools. The presence of roads may contribute to noise pollution \(^{25}\), air pollution \(^{26,27}\), and increased stress levels, which could have detrimental effects on learning and cognitive function \(^{9,28}\). Our findings are consistent with previous research that has shown a negative association between the presence of roads and academic performance \(^{29,30}\).

The association between wildfires and academic performance was mixed, with negative effects observed in the buffers around homes (but statistically insignificant) and positive effects in the buffers (500m, 750m, and 1km) around schools. While wildfires are known to have negative effects on air quality \(^{31}\), which could impact cognitive function and learning \(^{32}\), it is possible that the positive association observed around schools could be due to the fact that schools may have better fire protection measures in place and/or more resources to deal with the aftermath of a wildfire. Another possibility is that students enrolled in schools located in areas with more frequent wildfires are more likely to have access to educational resources and community support that could mitigate the negative effects of these events on academic outcomes.

We found that some neighborhood characteristics had a stronger association with academic performance when they were located around schools, compared to when they were located around students’ homes. This is consistent with previous research \(^{8}\). This finding suggests that the physical environment around schools may be particularly important for students' academic success. One possible explanation for this is that the physical environment around schools may influence the opportunities and resources available to students, both in terms of academic resources (e.g., well-equipped classrooms, libraries, and other educational materials) and extracurricular activities (e.g., sports fields, playgrounds, and other recreational facilities). Additionally, neighborhood characteristics around schools may also affect the quality of education that students receive, by influencing the quality of the teaching staff, the curriculum, and the overall school culture. For example, previous studies have shown that green spaces are associated with less crime \(^{33}\), and this is in line with research showing that schools located in neighborhoods with higher levels of crime may have a harder time attracting and retaining highly qualified teachers \(^{34}\). Regarding the curriculum, schools located in neighborhoods with poor environmental quality may be more likely to offer remedial classes and less likely to offer advanced classes \(^{35}\). Regarding school culture, students who attend schools located in more affluent neighborhoods (with environmental quality) may be more likely to participate in extracurricular activities and had higher levels of academic achievement \(^{36}\).

In contrast, neighborhood characteristics around students’ homes may have a more indirect effect on academic performance. This may be the reason that explains the number of insignificant associations when we accounted for the neighborhood characteristics around home. Our model may not be capturing these indirect associations for a variety of reasons, including omitted covariates and mediating variables. While students may spend a significant amount of time at home, the physical environment around their homes may not have as direct an impact on their academic success as the environment around their schools. Instead, the physical environment around homes may indirectly influence academic performance by affecting students’ health, well-being, and overall readiness to learn.
While our study provides valuable insights into the association between neighborhood characteristics and academic performance among students in the FD, Brazil, there are some limitations that should be acknowledged. First, our study was limited by the use of cross-sectional data, which prevents us from establishing causal relationships between neighborhood characteristics and academic performance. Second, our study only focused on three types of neighborhood characteristics, including green spaces, roads, and wildfires. Other neighborhood characteristics, such as noise and air pollution, could also impact academic performance but were not accounted for in our study. Third, our study was conducted in a specific geographical location and may not be generalizable to other regions or countries. Different regions and countries may have different neighborhood characteristics and cultural factors that impact academic performance. Fourth, while we included socioeconomic factors such as income and education level as covariates, there may be other factors that we did not account for that could impact academic performance, such as family structure or parental involvement. Finally, while our regression model with mixed effects took into account school-specific variables related to the school’s characteristics and location, there may be unmeasured school-level factors that could impact academic performance.

Despite the limitations, our study has several strengths. First, we employed a comprehensive approach that accounted for both the neighborhood characteristics around student’s school and home, as well as individual-level academic performance. This approach enabled us to control for potential confounding factors and to better understand the unique contributions of neighborhood characteristics to student performance. Second, we used mixed-effects regression models that account for the clustering of students within schools and the variability of neighborhood characteristics across students, thus improving the precision and accuracy of our estimates. Third, our study used objective measures of neighborhood characteristics, such as satellite data, to minimize the potential bias and measurement errors associated with self-reported data. Finally, our study provides valuable insights into the potential role of neighborhood characteristics, such as green spaces and roads, in promoting academic performance, which could have important implications for urban planning and public policy aimed at improving educational outcomes.

4. CONCLUSIONS

Overall, our results suggest that neighborhood characteristics play an important role in academic performance. Our study adds to the growing body of research that supports the need for policymakers and urban planners to consider the impact of neighborhood characteristics on academic performance when designing and implementing policies and programs aimed at improving educational outcomes. By creating environments that are conducive to learning, we may be able to provide students with the opportunities and resources they need to succeed academically.

5. MATERIALS AND METHODS

5.1. Study area

The study was conducted in the Federal District (FD) in Brazil, where the city of Brasilia, the capital of Brazil, is located. The FD has an estimated population of 3 million inhabitants (about 700,000 are children and adolescents), with a population density of more than 520 inhabitants/km². According to data from the Brazilian Institute of Geography and Statistics (IBGE), there were 684 schools in the Distrito Federal, Brazil as of 2021. In terms of the number of students, around 447,000 students were enrolled in schools in the Distrito Federal in 2020. This represents approximately 1.9% of the total number of students enrolled in Brazil, which was around 47.3 million in the same year.

5.2. Data

We accounted for three types of neighborhood characteristics around students’ schools and homes, including green spaces, roads, and wildfires. In the next sections, we describe these variables, the data representing academic performance, and the covariates used in the statistical analyses.

5.2.1. Academic performance

The Department of Education in the FD provided data on academic achievement, which consists of student-level data from 2017 to 2020. This information includes each student’s final grades, which are represented by a numeric variable ranging from 0–10, with 10 being the highest grade. The data also includes the school name, school level (middle school, high school, or adult education program), year, location of the school (represented by district, called as “RA”), and the type of school (urban or rural). It is important to note that
this data only includes students from public schools in the FD. The data covers 256 public schools in the FD and 344,175 students who were enrolled in these schools between 2017 and 2020.

All experimental protocols with the academic performance data were approved by the Department of Education in the FD. Administrative permissions were required to access the raw data used in this study. The Department of Education in the FD granted permission.

5.2.2. Green spaces

To evaluate the impacts of greenness on student-level academic performance, we utilized two different greenness metrics. Specifically, we analyzed the quantity of green spaces and NDVI. To obtain the first metric, we used land use data that were provided by the Brazilian Institute of Geography and Statistics (IBGE). The land use information was derived from images taken by Landsat 8, which is equipped with the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) instruments. The IBGE processed the data from Landsat 8 and created 14 different land use classes for Brazil in 2016, including agricultural areas, blue areas (such as rivers and lakes), urban areas, areas of deforestation, forests, and grasslands. For our study, we only focused on the classes that correspond to green spaces. To estimate the quantity of green spaces, we used Geographic Information System (GIS) techniques to measure the area of all green spaces within buffers of 250m, 500m, 750m, and 1km around each school and each student’s home.

For the second metric, we utilized the Green Space NDVI, which was accessed via satellite remote sensing. We used data from the Moderate Resolution Imaging Spectroradiometer (MODIS) NDVI, which was captured by the Terra and Aqua satellites (MODIS MOD13Q1 and MODIS MYD13Q1 products). This data was smoothed, and gaps caused by clouds or bad pixels were filled using the NASA Stennis Time Series Product Tool – TSPT. The vegetation index was calculated as the normalized ratio of the near-infrared (NIR) and red reflectance (RED) bands. The NDVI ranges from −1 to 1, where negative values represent water, values near zero represent barren areas of rock or sand, and positive values correspond to greener regions (with 1 representing the greenest area). To calculate the average NDVI for the study period (2017–2020), we used data from the months that correspond to the Brazilian school year (February–June, August–November). We estimated the average NDVI of all pixels (250m x 250m) in a buffer of 250m, 500m, 750m, and 1km around each school and each student’s home.

5.2.3. Roads

To investigate the potential impacts of traffic (a proxy for air quality, representing an environmental condition in the neighborhood) on the academic performance of students, we considered the total length of all roads within specific buffer zones surrounding schools and homes. The data regarding the road network was sourced from the Secretary of State for Habitation in the FD. Using GIS techniques, we calculated the total length of roads (in meters) within each buffer zone, including 250m, 500m, 750m, and 1km around each school and each student’s home.

5.2.4. Wildfires

The National Institute of Spatial Research of Brazil, Instituto Nacional de Pesquisas Espaciais - INPE, provided wildfire data, including the date of wildfire records and its geographical location. The data were collected from seven satellite remote sensing observations, processed by INPE, and based on the AQUA satellite as a reference. The data encompassed all wildfire records occurring in the FD between 2017 and 2020. Using GIS techniques, we aggregated the annual number of wildfires within buffers of 250m, 500m, 750m, and 1km around each school and each student’s home. Figure 1 illustrates the spatial distribution of schools in the FD, stratified by rural and urban locations and the three types of neighborhood characteristics considered in this study, including green areas, roads, and wildfires.

5.2.5. Covariates

We used income and education level as covariates. Income information was derived from the number of households within the census tract with a family income exceeding three minimum wages, while education level was assessed based on the number of literate individuals above 20 years old within the same census tract. The Brazilian Institute of Geography and Statistics provided the relevant census tract data, which was then spatially linked with the education and exposure variables (green spaces, roads, and wildfires) using GIS techniques based on the location of the school and census tract and then students’ home and census tract.

5.3. Statistical analyses
First, applied the Wilcoxon test to compare the environmental characteristics surrounding schools and homes. This test determines if the two samples (characteristics surrounding schools and homes) come from the same population or if one sample tends to have larger values than the other.

Then, we employed a regression model with mixed effects to investigate the relationship between academic performance at the individual level and neighborhood characteristics (green spaces, roads, and wildfires) around students’ schools and homes in the FD. Our model took into account temporal factors, socioeconomic factors, and school-specific variables related to the school’s characteristics and location. We hypothesized that neighborhood characteristics around schools and homes may be associated with individual students in a manner that produces varying residuals across students. To account for this, we incorporated a random intercept and slope that could differ across students. The statistical model used is described in Eq. 1.

\[ y_{i,z} = \alpha_{[i],z} + \beta_1[i] (UST_z) + s(\text{year}) + \beta_2 (\text{income}) + \beta_3 (\text{education}) + \beta_4 (\text{type}) + \beta_5 (\text{location}) + \beta_6 (\text{level}) + \epsilon_{i,z} \] (1)

where \( y \) is the student-level academic performance for student \( i \), considering the exposure to the neighborhood characteristic \( z \), \( \alpha \) is the intercept coefficient for neighborhood characteristic \( z \) varying by a student \( i \), \( \beta_1 \) is the slope coefficient for neighborhood characteristic \( z \), which varies by a student \( i \), \( s() \) is the smoothing spline function to characterize nonlinear relationships between student-level academic performance and year (2017–2020); \( \beta_2, \beta_3, \beta_4, \beta_5, \beta_6 \) are the vectors of coefficients that represent the \( \text{income} \) (variable describe above, item 2.2.3), \( \text{education} \) (variable describe above, item 2.2.3), \( \text{type} \) (type of school, including urban and rural), \( \text{location} \) (RA where the school is located), \( \text{level} \) (school level, including middle school and high school), respectively.

The model described by Eq. 1 was applied individually for each neighborhood characteristic (amount of green spaces, NDVI, roads, and wildfire), each buffer (250m, 500m, 750m, and 1km), and each place (schools and homes). Finally, we conducted stratified the analyses by type of school and school level. We used R software, version 4.1.3, including the R package "lme4" for the mixed effects regression model.

Declarations

ACKNOWLEDGEMENT

This work was supported by the Brazilian Agencies National Council for Scientific and Technological Development (CNPq) and by the Ministry of Science, Technology and Innovation (MCTI). This work was supported by the Getulio Vargas Foundation (FGV) in Brazil.

AUTHOR CONTRIBUTIONS

Conceptualization, Methodology, Investigation, Supervision, Original draft: WJR

Data curation, Methodology, and writing—review & editing: LMS

DATA AVAILABILITY STATEMENT

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

COMPETING INTERESTS

All other authors declare they have no competing interests.

STATEMENT

All experiments were performed in accordance with relevant guidelines and regulations.

References


**Figures**
Figure 1

Spatial distribution of schools in the FD, stratified by rural and urban locations and the neighborhood characteristics (roads, wildfires, among of vegetation, and NDVI).
Figure 2

Effects of neighborhood characteristics (NDVI, amount of vegetation, roads, and wildfire) on students’ performance (primary analysis) stratified by buffer and location of the exposure (home or school).

Note 1: "VEG" is the area (m$^2$) of green spaces within the buffers, "Roads is the length of roads (m) within the buffers, and "Wildfire" is the number of wildfire records within the buffers.

Note 2: Gray color represents the insignificant coefficients (which includes the value 0), the red color represents the significant positive associations, and the green color represents the significant negative associations.
Figure 3

Effects of neighborhood characteristics (NDVI, amount of vegetation, roads, and wildfire) on students’ performance (stratified by type of school – urban and rural) stratified by buffer and location of the exposure (home or school). Note 1: “VEG” is the area (m$^2$) of green spaces within the buffers, “Roads is the length of roads (m) within the buffers, and “Wildfire” is the number of wildfire records within the buffers. Note 2: Gray color represents the insignificant coefficients (which includes the value 0), the red color represents the significant positive associations, and the green color represents the significant negative associations.
Figure 4

Effects of neighborhood characteristics (NDVI, amount of vegetation, roads, and wildfire) on students’ performance (stratified by school level – middle and high school) stratified by buffer and location of the exposure (home or school). Note 1: “VEG” is the area (m²) of green spaces within the buffers, “Roads is the length of roads (m) within the buffers, and “Wildfire” is the number of wildfire records within the buffers. Note 2: Gray color represents the insignificant coefficients (which includes the value 0), the red color represents the significant positive associations, and the green color represents the significant negative associations.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterial.docx