Policy Research Working Paper 10993

Local Economic Shocks and Human Capital Accumulation

Evidence from Rwandan Coffee Mills

Andrew Dabalen
Justice TeiMensah
Aimable Nsabimana

WORLD BANK GROUP
Africa Region
Office of the Chief Economist
December 2024

Policy Research Working Paper 10993

Abstract

This paper examines the medium-term effects of policy-driven income shocks on human capital accumulation in low-income environments. Using administrative data on test scores of the universe of primary school students in Rwanda and the staggered rollout of coffee mills in the country, it shows a positive spillover effect of the coffee mills on students' performance. Early life exposure to coffee mills is associated with a 0.09 standard deviation (4 percent) increase in student test scores. Improvements in household welfare, child health, and school attendance are likely operative channels of impact.

This paper is a product of the Office of the Chief Economist, Africa Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at jmensah2@worldbank.org; adabalen@worldbank.org; and aimable.nsabimana@wider.unu.edu.

Local Economic Shocks and Human Capital Accumulation: Evidence from Rwandan Coffee Mills*

Andrew Dabalen, Justice Tei Mensah, Aimable Nsabimana

JEL classification: I28, Q13, J24
Keywords: Local economic shocks, Coffee Mills, Human capital accumulation, Rwanda

*Dabalen: Office of the Chief Economist, Africa Region, The World Bank. Email: adabalen@worldbank.org; Mensah: Office of the Chief Economist, Africa Region, The World Bank. Email: jmensah2@worldbank.org; and Nsabimana: UNU-WIDER, Pretoria, South Africa, Email: aimable.nsabimana@wider.unu.edu; We wish to express our profound gratitude to the Rwandan National Examination and School Inspection Authority (NESA), Rwanda Basic Education Board (REB) for granting access to standardized students' test score datasets, and to the Rwanda Ministry of Education (MINEDUC) for sharing official documents related to the education sector in the country. We are grateful to Kathleen Beegle, Heidi Kaila, Dean Karlan, Konrad Burchardi, Moussa Blimpo, and participants at the STIAS 2024 Nobel Symposium in Economic Sciences, and the CSAE 2024 Conference for their contributions and comments. We also thank Marijke Verpoorten for sharing the data on the Rwanda coffee census with us. All errors and views expressed in this paper are those of the authors. They do not necessarily represent the views of the World Bank or UNU-WIDER

1 Introduction

Value-addition to raw commodity exports is an important (aspirational) industrial policy strategy for commodity exporters in the Global South. This is because, aside from the increase in export revenues, adding value to commodities prior to export, as opposed to exporting them in their raw form, can stimulate job creation and growth of the local economy (Wang, 2013; Rodrik, 2022). The local economic effects of industrial (trade) policy measures have thus gained significant attention in the literature (Busso et al., 2013; Wang, 2013; Goldberg and Pavcnik, 2016; Grogan, 2023). However, many of these studies focus on the "first-order" impacts of industrial (trade) policies, such as employment (Glick and Roubaud, 2006), wages (Cling et al., 2005; Glick and Roubaud, 2006) and income (Picarelli, 2016). Little is known about the "second-order" effects on outcomes such as health and education (Atkin, 2016). Meanwhile, these "second-order" effects play a key role in determining the long-run impact of these policies.

In this paper, we explore how agro-industrialization policies affect human capital accumulation in low-income environments using administrative data from Rwanda. Theoretically, the effect of agro-industries or any local economic shock on human capital accumulation is ambiguous. On the one hand, local economic shocks boost household income, which enables them to invest in the schooling and healthcare of their children. On the other hand, standard economic models suggest that an increase in wages and employment opportunities for low-skilled workers, resulting from a boom in the local economy, increases the opportunity cost of schooling and reduces the marginal returns to higher education, thereby reducing educational attainment, particularly in environments where regulations on compulsory schooling for children are loosely enforced (Atkin, 2016; Mensah et al., 2019; Carrillo, 2020). Thus, the net effect of these competing factors determines the extent to which local economic shocks influence human capital accumulation in host communities.

We leverage the rapid policy-induced expansion of coffee mills in Rwanda in the 2000s to estimate the effects of early life exposure to local economic shocks on students' performance in the medium term. In an attempt to increase the value of Rwanda's coffee exports, the government implemented a National Coffee Strategy, which sought, among other things, to liberalize the coffee industry to allow competition and also increase the adoption of modern technology in coffee processing through the rapid establishment of coffee mills across the country (Macchiavello and Morjaria, 2021). Between 2000 and 2013, the number of coffee mills in the country increased from 3 to 226 (Figure 1). Prior to this, Rwandan coffee was primarily processed by farmers using traditional dry methods whereby farmers pulp the cherries at home using rocks before drying them on mats, resulting in low-quality beans (Macchiavello and Morjaria, 2021). However, the coffee mills, otherwise referred to as coffee washing stations, use the so-called wet processing approach. This approach relies on the use of water and machinery such as cherry hoppers, which sort the coffee cherries based on quality, after which the sorted cherries are pulped using coffee pulping machines. After pulping, the cherries are further sorted by immersion in water for several hours (Macchiavello and Morjaria, 2021).

Coffee processed at these stations attracts a higher market premium than that processed through traditional methods by farmers. Estimates suggest that the prices of wet-processed coffee via the mills are 40% higher than that of dry-processed coffee (Macchiavello and Morjaria, 2015). This suggests that the introduction of coffee mills induced significant income shocks to coffee farmers in the country. In addition, the introduction of the mills also led to an increase in employment opportunities through wage employment, as labor[1] was hired to work in these mills (Sanin, 2021).

Thus, the extent to which the local economic shocks associated with the introduction of these coffee mills affect the cognitive development of children exposed to these shocks in early life is the focus of this paper. To this end, we combine unique administrative data on test scores of the universe of primary school students in Rwanda with data on the staggered rollout of the coffee mills to estimate the effects of early life exposure to the coffee mills on students' performance in a high-stakes national exam.

To causally identify the impact of these shocks on students' performance in the medium term, we leverage the staggered variations in the rollout of the coffee mills across various locations relative to the timing of birth. Using a 2.5km radius from a coffee mill as the catchment area for coffee mills, we measure early life exposure to coffee mills as children born in a catchment area after the opening of the coffee mill. Essentially, our identification strategy is a difference-in-difference design whereby we compare the educational outcomes of children in a catchment area born after the opening of a mill with those in the catchment area but born before the opening of a mill and those born in communities outside the catchment area. A key identifying assumption advanced here is that in the absence of these coffee mills, the educational outcomes of treated and control students would have evolved along a similar pattern. We present results from an event-study analysis to support this assumption, using the Borusyak et al. (2024) estimator which accounts for heterogenous treatment effects in analysis with staggered treatment. Another potential challenge to our identification strategy relates to the fact that children born just before the arrival of the coffee mills could also experience positive spillover effects, despite being classified as control as per our baseline approach. We address this by presenting additional analyses that estimate the cumulative years of exposure to coffee mills within the first five years of life on educational outcomes and find similar results.

[1]Mostly women

We also provide several tests of robustness to rule out additional threats to our identification strategy, such as the role of selective migration, spatial spillovers, contemporaneous expansion in school infrastructure, and the definition of catchment areas of coffee mills.

Three main findings emerge from the paper. First, early life exposure to the coffee mills is associated with improvements in student performance in the medium term. Specifically, the aggregate test scores of students born in the catchment area of the coffee mills after the establishment of the mills are 0.09 standard deviation (4% relative to the mean) higher than comparable students without exposure to the coffee mills in early life. The effects persist across the various subject groups: STEM (math and science), and Non-STEM (social studies, English, and Kinyarwanda). Second, the positive effects of early life exposure to coffee mills pertain largely to students in rural communities. This provides suggestive evidence of the importance of income shocks in improving human capital accumulation in low-income environments where liquidity constraints are strictly binding. Third, we find suggestive evidence of improvements in household welfare (income), child health, and school attendance as plausible channels.

The effects are robust to several robustness checks, including measurement of early exposure, selective migration, spatial spillovers in exposure, parallel trends assumption, expansion in school infrastructure, and staggered treatment with heterogeneous treatment effects. For instance, evidence from an event study analysis using the Borusyak et al. (2024) estimator, provides support to the so-called "parallel trends" assumption, by showing similar trends in the evolution of test scores of students in treated and control villages before the establishment of coffee mills. However, following the arrival of the mills, the performance of students born in the treated villages begins to diverge from their counterparts in control villages.

This paper is connected to three strands of the literature. First, it builds on the literature that studies long-and medium-term effects of early life conditions on human capital in developing countries (Duflo, 2001; Edmonds et al., 2010; Adhvaryu et al., 2019; Carrillo, 2020; Adhvaryu et al., 2023). Adhvaryu et al. (2019) for instance show that early life exposure to income shocks via commodity price boom is associated with a lower likelihood of experiencing severe mental distress in adulthood. Similarly, Adhvaryu et al. (2023) find that adverse rainfall shocks at the time of birth are associated with lower educational attainment in the long run. However, the authors show that access to social protection programs such as cash transfers plays a key role in minimizing these adverse impacts: the effects on children from beneficiary households of the cash transfer program, PROGRESA, were much lower than those on children from non-beneficiary households. Results from our paper resonate with the prior results by demonstrating the importance of early life shocks on the cognitive development of children in the medium-to-long term.

Second, and related to the above, it contributes to the extensive literature on the long-term effects of early childhood development programs (Heckman et al., 2013; Gertler et al., 2014; Kline and Walters, 2016; Baker et al., 2019). The overwhelming majority of these studies show long-lasting impacts of health and educational investments in early life. Gertler et al. (2014) for instance, show that a program that provided psychosocial stimulation to growth-stunted children in Jamaica led to an increase in future earnings. While our paper speaks to the medium-term impacts of early-life exposure to local economic shocks, our findings reinforce the need for greater attention to early-life conditions as they can influence socioeconomic outcomes in the medium-to-long term.

Finally, this paper contributes to the broad literature on local impacts of industrial shocks (Picarelli, 2016; Blattman and Dercon, 2018; Benshaul-Tolonen, 2019; Von der Goltz and Barnwal, 2019; Bazillier and Girard, 2020; Sanin, 2021; Grogan, 2023). Benshaul-Tolonen (2019), for instance, finds that the opening of industrial mines in Africa is associated with a lower likelihood of infant mortality with improvements in household welfare via employment of mothers as a likely operative channel. Our findings of a positive educational response to the establishment of coffee mills in Rwanda present additional evidence on the impacts of local shocks on livelihoods of host communities.

The paper proceeds as follows. The next section presents background information about the coffee sector, coffee mills, and the education sector in Rwanda. Section 3 describes the various datasets used in the study. The empirical strategy is presented in Section 4, while Sections 5 and 6 present the main results and the underlying mechanisms, respectively. We conclude the paper in Section 7.

2 Background

2.1 The Coffee Sector in Rwanda

Coffee was first introduced to Rwanda in the early 1900s by German missionaries. It has since become a major crop cultivated by farmers in the country during the colonial and post-colonial regimes.[2] At the time of the country's independence in 1962, the sector accounted for 55% of Rwanda's exports. The sector, however, severely dwindled following the collapse of the international coffee market in the 1980s, and the subsequent decline in global coffee prices. The sector

[2] It is important to note that during the pre-independence era, the colonial government made coffee farming compulsory and enacted laws that prohibited the uprooting of coffee trees. Fertilizers and coffee seeds were also distributed for free by the colonial administration to increase productivity (Guariso et al., 2011).

was steadily revived after the 1994 genocide when the new government introduced policy reforms in coffee farming.

Today, coffee remains a major cash crop and accounts for a significant share of Rwanda's export earnings. In 2017, for instance, coffee contributed more than 23 percent of the country's agricultural export values (Macchiavello and Morjaria, 2022). Aside from exports, the sector is a major source of employment in the country, employing thousands of people along the crop value chain.[3] In 2018, for instance, there were more than 400,000 smallholder farmers actively engaged in coffee farming with approximately 90 million coffee trees grown in more than 88 municipalities[4] in the country. The main variety of coffee grown in Rwanda is *Coffea Arabica*, which accounts for about 98% of total coffee grown in the country (Guariso et al., 2011). Figure 2 shows the spatial distribution of coffee intensity at the subnational level in the country.

2.2 Coffee Mills

Between 1999-2000, Rwanda launched the Vision 2020 program and outlined the goals that the country envisioned to achieve by 2020. One of the key goals was to reshape agriculture to become a competitive and market-oriented sector (Boudreaux, 2011). In this respect, Rwanda designed a four-year Coffee Strategy and Action Plan (1999-2003) to achieve high-quality and wet-processed coffee production to participate in the international specialty coffee market. The strategy aimed to increase coffee production, improve coffee quality, and promote equity in value distribution through farmer participation in coffee marketing. To make this possible, the country liberalized the coffee industry by removing all trade barriers and encouraged competition and technology adoption in the sector (Sanin, 2021; Boudreaux, 2011). In light of the coffee policy reform, the Government of Rwanda prioritized shifting production techniques from coffee dry processing to a wet processing approach. [5] This set the foundation for the massive construction of coffee mills across the country to improve the quality of Rwandan coffee and increase forex earnings. As shown in Figure 1, the number of coffee mills increased from 5 in 2002, to 45 in 2005, 186 in 2010, and 312 in 2018.

[3]These include on-farm labor for seedlings, cultivation, harvest, processing in post-harvest, packaging, and trading.

[4]Also referred to as sectors

[5]The dry processing method is the traditional way of coffee processing where all tasks are done by farmers at their homes without any machinery use, while in the wet processing approach, farmers sell their coffee cherries to coffee mills, locally known as Coffee Washing Stations (CWS). In the mills, de-pulping coffee cherries and cleaning are done with specific machines using plenty of water; this is why the method is named wet processing.

2.3 The Education Sector in Rwanda

Rwanda uses the 6-3-3 pre-tertiary education system, which comprises 6 years of primary education; 3 years of middle school (Ordinary level); and 3 years of senior secondary school (advanced level). At the end of primary school (Grade 6), students sit for a mandatory national exam and obtain the primary leaving certificate, which allows them to move on to junior secondary education. This primary school exit examination is the focus of this study.

The exam consists of five subjects (mathematics, science, social studies, Kinyarwanda, and English). Performance in this examination is the key metric for progression into middle school (O-Level). Usually, the best-performing students in the national exams are automatically awarded places in boarding schools. Although students with low performance in the national exam are promoted to the middle school level, they usually remain in the same schools for what is commonly known as 12 years of basic education (12YBE). The average age of starting primary school education in Rwanda is 7, although most of the students complete primary school at age 14 as shown in Table 1.[6]

3 Data

3.1 Test Scores

Our primary dataset is standardized test scores of the universe of primary school students in the primary school national exams between 2012 and 2019, obtained from the National Examination and School Inspection Authority (NESA). It contains the test scores of nearly 1.5 million students who sat for the exam during the period. The dataset contains the standardized test scores in each of the five subjects (mathematics, science, social studies, Kinyarwanda, and English); grade points in each subject; age and gender of the student; and unique school ID. We match this data with a dataset on the geolocation of each school using the school IDs. Figure 3 shows the geographic footprint of primary schools and coffee mills in Rwanda.[7]

NESA uses a nine-point grading system in order from best to worst- 1 (Excellent) to 9 (Poor). Since we are interested in students' aggregate performance in all subjects as well as individual subjects, we use the grades in the respective subjects to generate a standardized grade point. Specifically, following Denteh et al. (2022), we reverse the scales so that a high grade means better performance. Essentially, we re-order the ranking of the grades such that 1 means Poor

[6]It is noteworthy to emphasize that Rwanda introduced 6-year compulsory primary education. However, in 2009 the country extended the program to a 12-year free and compulsory education system.

[7]Our data on the location of coffee mills was obtained from Rwanda Geoportal

and 9 means Excellent. We then aggregate these individual (rescaled) grades to obtain an aggregate grade point ranging from 5 to 45, with grades 5 and 45 representing the lowest and highest grades respectively. We also compute the total grade points for STEM (mathematics and science) and non-STEM (social studies, Kinyarwanda, and English) subjects. It is important to emphasize that the grading system was unchanged during the study period, thus ensuring comparability of the grades across years. Nonetheless, for ease of interpretation and to ensure consistency with the test scores of the individual subjects, we standardized the grade points for the aggregate, STEM, and non-STEM subjects. Finally, using the geo-location of all primary schools in the country and data on the location and year of establishment of coffee mills, we spatially match each school to the nearest coffee mill.

3.2 DHS Data

To explore potential mechanisms, we use data from the Demographic and Health Surveys (DHS) to understand the potential mechanisms. The DHS data is a nationally representative survey and covers a wide range of issues including, but not limited to, health, education, and household assets. For the purposes of this study, we focused on anthropometric measures (wasting and underweight) of children under five years, school attendance, and household wealth. The main advantage of the DHS over the other surveys such as the Integrated Household Living Conditions Survey (EICV) is that it is geocoded, allowing us to spatially match survey locations with the coffee mills datasets to determine treated and control households.

3.3 Complementary Data

We complement our analysis with extensive data on coffee intensity at baseline from Guariso and Verpoorten (2018). Coffee intensity is measured as the number of coffee trees per sq km in 1999, compiled from the Rwanda coffee census. We also include data on road density at the municipality level, average annual temperature, total precipitation, and coffee suitability.

4 Empirical Strategy

Our baseline specification is a difference-in-difference (DID) design, where we compare the differences in test scores of students who were exposed to coffee mills in early life to those born earlier in the community and hence not exposed to the mills in early life, with counterparts born in communities without a coffee mill. Essentially, our estimation strategy leverages plausibly exogenous variations in the establishments of the coffee mills across space and time (see Figure

1). In the early 2000s, Rwanda implemented its Coffee Strategy and Action Plan (1999-2003) with the goal of improving the value of Rwandan coffee exports and increasing earnings. Central to this policy was the establishment of coffee washing stations, otherwise referred to as coffee mills, to provide efficient and cost-effective processing of coffee, thus moving coffee farmers away from traditional methods of coffee washing which often led to low-quality coffee. This led to a massive increase in the number of coffee mills in the country, increasing from 5 in 2002 to 186 in 2010, with an additional 126 mills added by the end of 2018.

To causally estimate the effects of these coffee mills on outcomes, we focus on the massive expansion in coffee mills between 2002 and 2010 driven by the policy shock and compare the outcomes of children exposed to these coffee mills to their counterparts without exposure to the mills in a difference-in-difference framework. To this end, we combine data on the staggered rollout of the coffee mills (i.e., location and date of establishment) with the timing of birth to measure early life exposure to coffee mills. Consequently, our baseline equation can be specified as follows:

$$S_{ijty} = \phi \cdot CatchmentArea_j \times Post_{it} + \mathbf{X}_{\mathbf{ijt}}^{'}\alpha + \theta_j + \delta_t + \lambda_y + \epsilon_{isty} \tag{1}$$

where the outcome variable, S_{ijty}, is the test score of student i in school (community) j, born in year t, and sitting for the exam in year y. $CatchmentArea_j$ is a dummy variable set equal to 1 if the student lives in a community that is within the catchment area of a coffee mill, and 0 otherwise. $Post_{it}$ is a dummy variable set equal to 1 if the student was born after the opening of the nearest coffee mill and 0 otherwise. Hence, $CatchmentArea_j \times Post_{it}$ is an indicator variable for students living within the catchment area of a coffee mill and born after the opening of the mill. $\mathbf{X}_{\mathbf{ijt}}^{'}$ is a vector of controls, including, gender of the student, average annual temperature, and total precipitation in the community during the exam-year, as well as baseline characteristics (coffee suitability and road density) interacted with time trends. Since the coffee mills were rolled out at different times, we compare the outcomes of students with and without exposure to the mills at birth by virtue of their community and date of birth. Therefore, we control for school (community) fixed effect, θ_j, birth-year (cohort) fixed effect, δ_t, and exam-year,λ_y, fixed effect. Standard errors are clustered at the school (community) level. Also, since we run our regressions on multiple outcomes (test scores by subjects), we report the False Discovery Rate (FDR) adjusted p-values by Anderson (2008) to ascertain the robustness of our estimates to multiple hypothesis testing.

Despite the granular data on the temporal and spatial rollout of the coffee mills, information on the exact demarcation of the catchment area of these mills is scant. Surveys however suggest that in some cases, the size of the catchment areas could extend to about 5km, even though they are largely localized serving farmers within the immediate vicinity of the mills (Macchiavello and Morjaria, 2015). This is largely because raw coffee is bulky and the travel cost associated with hauling coffee from farms to the mills, particularly given the topography of Rwanda, increases with distance. Thus, after a given distance threshold, the net benefits from hauling coffee to these mills decrease with distance. Therefore, for our baseline analysis, we choose a conservative distance threshold of 2.5km,[8] as the optimal size of the catchment area. To put this into perspective, we classify students living within a 2.5km radius of a coffee mill as treated while students living beyond the 2.5km buffer as control. We also present alternative estimations and show that our results are fairly robust even if we extend the size of the catchment area up to 5km from a mill. Figure 4 provides a graphical representation of our treatment and control locations. Further, in Figure 8 we estimate the effects of the coffee mills across distance splines using a spatial lag model to validate the choice of our baseline distance for defining treatment and control areas.

It is noteworthy to emphasize that we proxy a student's place of residence (residence) by the location of the school, due to the lack of data on students' residency and place of birth. Admittedly, this is a strong assumption, but given the wide geographic footprints of primary schools in Rwanda, as shown in Figure 3, school location is a reasonable approximation of students' residence.[9] As a result, $\hat{\phi}$, is essentially an intent-to-treat (ITT) estimate of the medium-term effect of exposure to the coffee mills in early life on student test scores.

Further, equation 1 implicitly assumes that the impact of the opening of coffee mills on the cognitive development of children is largely present during the time of birth. Hence, only children born during or after the opening of the coffee mills are likely to be beneficiaries of these positive impacts. However, when a mill starts operating in a community, it continues and does not shut down. Thus, children born say 1-2 years before the opening may potentially benefit from the income shocks associated with the opening of the coffee mills in the community, thus improving their cognitive development and helping them perform well in school. To this end, in Section 5.3.1, we present additional results by replacing our 0/1 treatment variable with a continuous treatment measure, measuring the share of years between ages 0-5 that a student was exposed to coffee mills. Reassuringly, both measures yield similar conclusions: early-life exposure to coffee mills has a positive impact on student performance.

Event Study

Finally, the validity of our research design rests on the assumption of (plausibly) exogenous

[8] plausibly the median size of the catchment areas

[9] Most primary school students in Rwanda live within a 2km radius of their school. See https://blogs.worldbank.org/nasikiliza/accelerating-afe-rwandas-development-midst-overlapping-crises

rollout of the coffee mills across locations. We address this concern in two ways. First, we control for trends in observable determinants of coffee mills rollout by including baseline coffee intensity interacted with linear time trends. This is expected to purge our estimates of the direct effects of coffee cultivation on our outcome. In addition, since we focus on students born during the period of the massive expansion of coffee mills, we control for possible confounders such as access to transport infrastructure, with the inclusion of road density interacted with time trends (Hoynes and Schanzenbach, 2009; Hoynes et al., 2016).

Second, we present event-study results to show that absent the coffee mills, the performance of students in treated and control locations would evolve along a similar trend by estimating the following specification:

$$S_{ijty} = \sum_{\tau=-5;\tau\neq-1}^{5} \phi_\tau \cdot CatchmentArea_j \times Post_{it}^{\tau} + \mathbf{X}_{\mathbf{ijt}}^{'}\alpha + \theta_j + \delta_t + \lambda_y + \epsilon_{isty} \tag{2}$$

However, the recent wave of literature on difference-in-difference with staggered treatment, as in the case of our study, suggests that using the conventional two-way fixed effects (TWFE) estimators for models with staggered treatment may lead to biased estimates due to the presence of negative weighting and heterogeneity in the treatment effects (De Chaisemartin and D'Haultfoeuille, 2018, 2020a,b; Goodman-Bacon, 2021; Callaway and Sant'Anna, 2021; Borusyak et al., 2024). To address these challenges and ensure the robustness of our estimates, we use the "imputation" estimator by Borusyak et al. (2024) to estimate our event study. This estimator ensures consistency in the presence of the TWFE and also deals with issues of negative weighting (Acemoglu et al., 2022; De Chaisemartin and D'Haultfoeuille, 2020b).

5 Results

We begin by presenting descriptive evidence on the trends in test scores between students in treated and control locations born before and after the opening of the (closest) coffee mill. This allows us to check the trends in test scores between treated and control locations before and after the opening of coffee mills, thus providing a validity check to our identifying assumptions. As shown in Figure 5, we observe a similar evolution in test scores between control and treated locations before the arrival of coffee mills. However, the test scores of students born in treated locations after the opening of coffee mills begin to outpace their counterparts in control locations, suggesting a potentially positive impact of coffee mills on students' test scores. Results from the event study analysis presented in the subsequent sections provide additional validation.

5.1 Main Results

Table 2 presents results on the effect of early-life exposure to coffee mills on aggregate (all subjects) test scores as well as STEM and non-STEM courses. For each outcome, we estimate three variant specifications. Our preferred results are in columns 3, 6, and 9. Also, we present two sets of results: (i) panel A where our dependent variable is the standardized grade points; and (ii) panel B, where the dependent variable is the raw (unstandardized) grade points.

Starting with panel A, we find that exposure to coffee mills in early life is associated with a 0.09 standard deviation (sd) increase in overall test scores in the primary school exit exams (column 3). Disaggregating the overall effects into performance in STEM and non-STEM courses, we find relatively large effects on non-STEM relative to STEM courses. Early life exposure to coffee mills increases test scores in STEM and non-STEM subjects by 0.06 and 0.1 sd respectively.

In panel B, we find similar results. Exposure to students exposed to coffee mills in early life experienced a 0.68 increase in grade points relative to cohorts without exposure to the coffee mills (column 3). Relative to the mean, this represents a 4.4% increase in aggregate test scores.[10] The corresponding increase in test scores in STEM and non-STEM courses are 3.7% and 4.7% respectively.

To further understand the impact of exposure to coffee mills on students' performance, Table 3 presents disaggregated analysis by looking at the effects on the five respective subjects (Math, Science, Social, English, & Kinyarwanda). Exposure to coffee mills increases Math, and Science test scores by about 0.08 sd. Meanwhile, the effect on students' test scores in Social Studies, English, and Kinyarwanda is 0.08, 0.09, and 0.06 sd respectively. Overall, the above results suggest that exposure to coffee mills at the time of birth improves the cognitive development of children leading to increased academic performance.

5.1.1 Event Study

To further understand the effects of early-life exposure to coffee mills, we present results from an event study analysis using the Borusyak et al. (2024) "imputation" estimator, as shown in Figure 6, which corrects for biases arising from heterogeneous treatment effects that are akin to staggered treatments.[11] The event study analysis allows us to: (i) examine the validity of our parallel trends assumption, and (ii) explore how the effects of the coffee mills vary with time.

Starting with panel (a), the results show relatively similar trends in test scores of students

[10] 0.6791/15.4015=0.044

[11] It is important to note that, unlike the conventional event study estimators, the reference group for pre-trends test in the Borusyak et al. (2024) "imputation" estimator is all k periods prior to the event date and all never-treated observations.

in treatment and control communities born before the opening of the nearest coffee mill, as the pre-treatment estimates are close to zero and statistically insignificant at the 95% confidence interval. However, we observe positive and (mostly) statistically significant estimates in the post-treatment period, suggesting an improvement in test scores for students in treatment villages born after the opening of a coffee mill.[12] Similar results are obtained for STEM and non-STEM test scores in panels (b) and (c) respectively. Additional event study analysis using the TWFE estimator (Figure A1 in the online appendix) yields similar conclusions. Overall, our event study estimates provide suggestive evidence in support of our identifying assumptions.

5.2 Effects by Gender

Next, we explore the effects of early-life exposure to coffee mills on test scores across gender as shown in Table 4. Across the table, we observe a positive and statistically significant effect on test scores of both male and female students, although the effect size appears marginally higher for male than female students. In the top panel, the results show that early-life exposure to coffee mills is associated with a 0.09, 0.06, and 0.1 sd increase in test scores in all (aggregate), STEM, and non-STEM subjects respectively for male students. Meanwhile, the effect on test scores for female students in the respective subjects are 0.08, 0.05, and 0.09 sd. Thus the results suggest that exposure to coffee mills is beneficial to students regardless of their gender.

5.3 Robustness Checks

Having established the positive effects of exposure to coffee mills on test scores, we explore in this section, the robustness of our baseline results to alternative hypothesis.

5.3.1 Alternative Measure of Exposure

In this section, we present an alternate measure of exposure to coffee mills using the share of years between age 0 and 5 that a student was exposed to coffee mills.[13] In contrast to our baseline binary measure of exposure, this measure accounts for the cumulative exposure to coffee mills. It also allows us to capture the spillover effects on children born just before the opening of the coffee mill but still benefited from the economic impacts of the mills within

[12]The only exceptions are the estimates for the third and fourth events after treatment, which, despite being positive, are imprecisely (statistically) estimated.

[13]This is measured by using the same 2.5km distance buffer and then computing the number of years within the first 5 years of the students for which the nearest coffee mill was operational.

their first 5 years which experts argue is the most critical stage in the cognitive development of children. The results in Table 6 are largely in line with our baseline results of a positive impact of exposure to coffee mills on students' test scores. Across the various subjects, we find a positive association between exposure to coffee mills and student performance.

5.3.2 Rural-Urban

Given the fact that agriculture is the mainstay of most rural households, income shocks to agricultural production such as the value addition associated with coffee mills are expected to be more impactful on livelihoods in rural areas relative to urban areas. Thus, a priori, we do not expect early-life exposure to these coffee mills to affect students' test scores in urban areas as would in rural areas.

The results in Table 5 confirm this assertion. While exposure to coffee mills exerts a positive and statistically significant effect on the test scores of students in rural (top panel), we find no evidence that exposure to coffee mills has an impact on the test scores of students in urban communities (lower panel). These results provide further support to our identification strategy: it suggests that our treatment variables are not picking up confounders such as changes in local economic conditions not associated with the introduction of coffee mills.

We also explore the extent to which our results may be driven by the inclusion of Kigali Province, which happens to be the most developed geographic area in the country[14] and has historically also been a coffee-producing region. In Table A1 in the (online) Appendix, we re-estimate our baseline regression while excluding the sample of students within the Kigali Province. Our results remain qualitatively and quantitatively similar to the baseline: a further assurance of the robustness of our baseline estimates.

5.3.3 Selective Migration

One concern is that given the importance of coffee mills in improving the value of coffee and associated effects on local economic growth, it may induce migration to areas within the catchment area of the mills. This becomes an issue if, for instance, well-endowed households strategically migrate to places with either a high potential for a coffee mill or where a coffee mill has been established. And if this is the case, then the estimated positive effect may be driven by the "selective" migration of (talented) children from these households to treated villages and not necessarily due to the impact of income shocks associated with the mills in improving cognitive development of children in treated villages.

[14] despite having rural communities therein

While this is difficult to test, empirically, we provide some checks to assuage this concern. We use census data on migration rates at the municipality level and sequentially trim our sample by dropping observations in municipalities with high migration rates.[15] Specifically, we estimate the baseline specifications under the three sets of fixed effects in Table 2 while systematically dropping observations in municipalities where the migration rates exceed the following thresholds: 60%, 50%, 40%, 30%, 20%, 10%, 6% and 3% respectively. The results are shown in Figure 7. Across the various specifications, the results for the respective subject groupings remain positive and statistically significant, regardless of the migration threshold that is applied. Thus, we can rule out the potential of selective migration confounding our results.

5.3.4 Alternative Measure of Catchment Area

So far, we have used a very conservative measure of treated locations using a 2.5km radius around the coffee mills. However, as shown in Macchiavello and Morjaria (2021), the size of the catchment area of coffee mills could extend to about 5km. Thus, if indeed the true benefits of coffee mills extend beyond 2.5km, then our baseline estimates may underestimate the true effect of the coffee mills on students' outcomes.

In this section, we redefine the catchment area to include communities within a 5km radius of a coffee mill and replicate the analysis forgoing analysis. The corresponding results are shown in Tables A3 - A7 (Section B.2) in the appendix. Reassuringly, the results align with the baseline results, albeit the effect sizes are relatively low compared to the latter. Overall, both sets of results point to a positive impact of exposure to coffee mills in early life on educational performance in the medium term.

5.3.5 Spillovers and Geographic Distribution of Impacts

Next, we address two potential threats to our assignment of treatment and control locations: I. There may be concerns that our estimated increase in student performance could be driven by a deterioration in student performance in control communities driven by say a negative spillover effect of the coffee mills on nearby communities. For instance, if the coffee mills induce a displacement of (low-income) non-coffee farming households from treatment to control communities, then the treatment effects could in principle be non-linearly distributed across space. II. Contrary to (I), there is also the possibility that given our "arbitrary" and fuzzy definition of catchment areas, some farmers (students) in nearby communities (just outside our treatment threshold) could in principle benefit from the coffee mills (i.e. treated) despite being assigned

[15] Municipality here is the second administrative region.

as control per our design. In that case, our treatment effects could be underestimated. We address these concerns using two approaches.

First, in relation to the geographic distribution of the effects, we estimate a spatial lag model and explore the effects of coffee mills on student outcomes across distance splines. Specifically, we relax our baseline assumption of using a 2.5km buffer as treatment and instead allow the effects to vary across distance bins (0-2.5, 2.5-5, 5-10, 10-20, 20-30).[16] The results, as shown in Figure 8, indicate an increase in test scores of students born within 2.5km of an operational coffee mill. The effects of those within the 2.5-5km distance bin are approximately 0, which could explain why the effect size when using a 5km radius as a measure of the catchment area as shown in Sections 5.3.4 and 4 are relatively lower than the baseline. Interestingly, the results for students within the 5-10km distance bin are negative albeit statistically insignificant. Beyond that, the estimates are also statistically insignificant. Overall, the results from the spatial lag estimation provide additional support to the choice of our baseline measure of a catchment area, as the main effects are concentrated within the 2.5km radius of the coffee mill.

To address the concern of "contamination" effects arising from the possibility of having beneficiary households in the control, we use the approach outlined in Section 5.3.3 to estimate our baseline specification, while systematically restricting our control group to students within a given distance threshold from the coffee mills, and maintaining our 2.5km buffer around the mills as our treatment. For example, we explore a scenario where the control group consists of students living beyond 5km from a mill, hence effectively excluding those living within 2.5-5km from the analysis. In other words, we systematically trim our control group by excluding students around the fringes of the buffer who could be potential beneficiaries. Results are shown in Figure 9 .

Once again, across the various specifications, the results for the respective subject groupings remain positive and statistically significant, regardless of the trimming applied to the control group – thereby ruling out concerns about the potential effects of spillovers contaminating our baseline estimates.

5.3.6 Access to Schooling

Finally, another potential threat to our identification strategy is changes in school access that coincide with the rollout of coffee mills. For instance, if the rollout of coffee mills coincided with say a massive expansion in school access such that treated locations disproportionately benefited from school construction relative to control locations, then one can argue that the estimated treatment effect may be driven by the changes in access to schooling opportunities

[16] The reason for the uneven sizes of the distance bins is due to sample distribution.

rather than the income effects associated with the coffee mills. This is important because mass construction of schools could, for instance, not only expand access to schooling opportunities but also have a positive impact on the quality of teaching and learning by reducing overcrowding.

To address this issue, we combine data on primary school construction in Rwanda with the data on coffee mills and explore the association between the rollout of coffee mills and changes in school access. Specifically, using data geolocation and the year of construction of the universe of primary schools in the country, we compute the number of (new) primary schools as well as the cumulative number of primary schools operational per year and correlate it with the rollout of coffee mills at the municipality level.[17] The results as shown in Table A2 in the (online) Appendix, reveal no association between the rollout of the coffee mills and expansion in primary schools. Thus, we can rule out the possibility of a contemporaneous expansion in school access contaminating our results.

6 Mechanisms

In this section, we explore the potential pathways through which exposure to coffee mills influences student performance. We hypothesize and test three key channels: income, health, and school attendance. To test these channels, we combine data from the Rwanda DHS with data on the location and operational status of the coffee mills. However, unlike the main analysis where we define the catchment area using a 2.5km radius from a coffee mill, we are unable to use the same distance threshold with the DHS data, due to the inherent displacements of the coordinates in the DHS survey. Due to privacy concerns, the geo-coordinates of the DHS are displaced by up to 2km in urban centers and 5km in rural areas.[18] As a result, defining treatment and control locations based on distance thresholds below 5km, especially in the context of our study where 88% of the sample are rural, is likely to induce significant measurement errors. Therefore, for the analysis on the mechanism, we define the catchment using the alternative measure, which is the 5km radius from the coffee mill. Given that the results in Section 5.3.4 confirm our baseline results, we argue that the benefits of using the 5km threshold for the analysis based on the DHS data far outweigh the potential measurement errors associated with defining treatment and control using the more restrictive distance threshold, i.e., 2.5km.

Consequently, we estimate the following specifications:

[17] Municipality here refers to the third administrative region, otherwise referred to as "sectors".

[18] https://dhsprogram.com/pubs/pdf/SAR7/SAR7.pdf

$$Y_{ijt} = \phi \cdot CatchmentArea_j \times Post_t + \beta \cdot Mills_j + \mathbf{X_{ijt}^{'}} \alpha + \lambda_t + \epsilon_{ijt} \tag{3}$$

where Y_{ijt} is the outcome if individual (household) i, in locality j, surveyed in year t. $CatchmentArea_j$ is a dummy variable set equal to 1 if the community is located within 0-5km from a coffee mill and 0 otherwise. $Post_t$ is a dummy defined equal to 1 if the coffee mill is operational at the time of the survey and 0 otherwise. $CatchmentArea_j \times Post_t$ is essentially a dummy set equal to 1 for those living within 5km from a coffee mill that is operational during the survey year and 0 otherwise. λ_t represent location and survey year fixed effects respectively. All other variables remain as previously defined.

6.1 Income

The underlying premise of this paper is that the opening of a coffee mill leads to an income shock to residents within the catchment area of the mill, due to the value-addition of coffee washed in these mills relative to those washed using traditional methods. While the direct effect of the income shock will be borne by coffee farmers, the indirect effect on the local economy cannot be overemphasized.[19]

To provide evidence on this channel, we leverage the wealth and asset poverty measures from the DHS dataset, due to the absence of income measures from the DHS dataset.[20] Table 7 presents results on the association between coffee mills and household income/welfare using proximate measures of household poverty status and wealth score from the DHS dataset.[21] The results show a positive income effect of coffee mills on treated households. Households living within the catchment area of coffee mills experience a 0.05 std increase in their wealth score (column 1), and a 4 percentage points (pp) decrease in the probability of being (asset) poor (column 3).

6.2 Child Health

Conditional on the positive income effect of coffee mills in the catchment area, a proximate outcome is an improvement in the health of children living within the catchment areas. We test

[19] However, disentangling these effects is challenging due to data constraints.

[20] Rwanda's Integrated Household Living Conditions Survey (EICV) has direct measures of household income, however, the dataset is not geocoded hence limiting its usage in the context of this study, especially as it pertains to identifying households in treated and control communities. In fact, the only location identifier in the (publicly available version) of the dataset is at the district level.

[21] The DHS measure of poverty is a measure of the relative (asset) wealth of a household. A household is classified as poor if its wealth score is within the two lowest quintiles of the wealth distribution in the country.

this assumption by looking at the association between coffee mills and two key health outcomes that have a significant bearing on the cognitive development of children: high-for-age (HAZ) scores and stunting (Grantham-McGregor et al., 1996; Mendez and Adair, 1999).

Our results in Table 8 provide evidence that coffee mills are associated with an improvement in child health. Children living within the catchment area of coffee mills are 5 pp less likely to be stunted (column 3) and experience 0.1 std increase in their HAZ scores (column 1).

6.3 School Attendance

We conclude our section on mechanisms by evaluating the effects on school attendance. Again, our working hypothesis is that the positive income shocks associated with coffee mills enable parents to increase educational investments in their children thereby allowing (incentivizing) the children to attend school and learn.

Ideally, testing this assumption will require granular data on student attendance at the school level before and after the opening of coffee mills. However, this data is unavailable in the study context. The available data on attendance is aggregated at the district level, thereby limiting its use, at least for this study. To this end, we rely on proximate measures of attendance from the DHS dataset, which asks if a "person attended school during the current school year", and restrict the sample to children within the primary school-going age (7 and 16) in Rwanda. Admittedly, this measure is imperfect as it may reflect enrolment rather than (regular) attendance, but we use it as a proxy for attendance.

We, therefore, combine the DHS data and information on the location and opening of the coffee mills to estimate the effects of the coffee mills on our proxy measure of "school attendance". The results as shown in Table 9 suggest a positive effect of the mills opening on school attendance rates. Starting in column 2, we find that children living within the catchment area of an operational coffee mill are about 2 pp more likely to attend school than their counterparts in control communities. Splitting the sample into rural and urban sub-samples, we find that the effects of the coffee mills on schooling are largely concentrated in rural areas (column 4). There is no statistically significant effect on school attendance on children in urban areas. This result concurs with our main results in Table 5[22] which also shows that exposure to coffee mills increases the test scores of children in rural areas with no (statistically) discernible effects on urban students. Overall, the results herein provide suggestive evidence that schooling and learning are potential channels through which the income shocks associated with coffee mills affect students' performance. This result is in line with prior literature which establishes a

[22] As well as Table A6 in the Appendix.

strong negative relationship between absenteeism (dropouts) and students' performance (Durden and Ellis, 1995; Bedi and Marshall, 1999; Chen and Lin, 2008).

7 Concluding Remarks

This paper examines the effects of agro-industries on human capital accumulation by investigating the impact of coffee mills in Rwanda on the performance of primary school children who were exposed to these mills in early life. We also explore the potential mechanisms that explain our results.

The study uses administrative data on test scores of the universe of primary school students sitting for a high-stakes national exam between 2012 and 2019 together with geocoded data on the location and timing of the establishment of coffee mills in Rwanda. Our identification strategy essentially exploits the staggered rollout of the coffee mills along with the timing of birth of the students to measure the medium-term effect of early life exposure to the coffee mills on test scores in a difference-in-difference design.

We find evidence of a positive impact of exposure to the coffee mills on student performance. On average, test scores of students born within the catchment area of coffee mills are about 0.09 standard deviation (4%) higher relative to their counterparts in control villages. Our results suggest improvements in household welfare (income), employment, child health, and school attendance as plausible channels.

Overall, the findings of the paper underscore the spillover effects of agro-industrialization, such as the establishment of coffee mills in local communities. More importantly, the paper highlights the need for developing countries that rely on raw commodity exports to invest in value-addition; this not only increases the value of their exports, but also has spillover effects on local economies.

References

Acemoglu, D., He, A., and le Maire, D. (2022). Eclipse of rent-sharing: The effects of managers' business education on wages and the labor share in the us and denmark. Technical report, National Bureau of Economic Research.

Adhvaryu, A., Fenske, J., and Nyshadham, A. (2019). Early life circumstance and adult mental health. *Journal of Political Economy*, 127(4):1516–1549.

Adhvaryu, A., Molina, T., Nyshadham, A., and Tamayo, J. (2023). Helping Children Catch UP: Early Life Shocks and the Progresa Experiment*. *The Economic Journal*, page uead067.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association*, 103(484):1481–1495.

Atkin, D. (2016). Endogenous skill acquisition and export manufacturing in mexico. *American Economic Review*, 106(8):2046–2085.

Baker, M., Gruber, J., and Milligan, K. (2019). The long-run impacts of a universal child care program. *American Economic Journal: Economic Policy*, 11(3):1–26.

Bazillier, R. and Girard, V. (2020). The gold digger and the machine. evidence on the distributive effect of the artisanal and industrial gold rushes in burkina faso. *Journal of Development Economics*, 143:102411.

Bedi, A. S. and Marshall, J. H. (1999). School attendance and student achievement: Evidence from rural honduras. *Economic development and cultural change*, 47(3):657–682.

Benshaul-Tolonen, A. (2019). Local industrial shocks and infant mortality. *The Economic Journal*, 129(620):1561–1592.

Blattman, C. and Dercon, S. (2018). The impacts of industrial and entrepreneurial work on income and health: Experimental evidence from ethiopia. *American Economic Journal: Applied Economics*, 10(3):1–38.

Borusyak, K., Jaravel, X., and Spiess, J. (2024). Revisiting event-study designs: robust and efficient estimation. *Review of Economic Studies*, page rdae007.

Boudreaux, K. C. (2011). Economic liberalization in rwanda's coffee sector: A better brew for success. *Yes Africa Can*, page 185.

Busso, M., Gregory, J., and Kline, P. (2013). Assessing the incidence and efficiency of a prominent place based policy. *American Economic Review*, 103(2):897–947.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, 225(2):200–230.

Carrillo, B. (2020). Present bias and underinvestment in education? long-run effects of childhood exposure to booms in colombia. *Journal of Labor Economics*, 38(4):1127–1265.

Chen, J. and Lin, T.-F. (2008). Class attendance and exam performance: A randomized experiment. *The Journal of Economic Education*, 39(3):213–227.

Cling, J.-P., Razafindrakoto, M., and Roubaud, F. (2005). Export processing zones in madagascar: a success story under threat? *World development*, 33(5):785–803.

De Chaisemartin, C. and D'Haultfoeuille, X. (2018). Fuzzy differences-in-differences. *The Review of Economic Studies*, 85(2):999–1028.

De Chaisemartin, C. and D'Haultfoeuille, X. (2020a). Difference-in-differences estimators of intertemporal treatment effects. Technical report.

De Chaisemartin, C. and D'Haultfoeuille, X. (2020b). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.

Denteh, A., Asare, S., and Senadza, B. (2022). Is four better than three? the effect of the 4-year high school policy on academic performance in ghana. *The Effect of The*.

Duflo, E. (2001). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American economic review*, 91(4):795–813.

Durden, G. C. and Ellis, L. V. (1995). The effects of attendance on student learning in principles of economics. *The American Economic Review*, 85(2):343–346.

Edmonds, E. V., Pavcnik, N., and Topalova, P. (2010). Trade adjustment and human capital investments: Evidence from indian tariff reform. *American Economic Journal: Applied Economics*, 2(4):42–75.

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., Chang, S. M., and Grantham-McGregor, S. (2014). Labor market returns to an early childhood stimulation intervention in jamaica. *Science*, 344(6187):998–1001.

Glick, P. and Roubaud, F. (2006). Export processing zone expansion in madagascar: What are the labour market and gender impacts? *Journal of African Economies*, 15(4):722–756.

Goldberg, P. K. and Pavcnik, N. (2016). The effects of trade policy. In *Handbook of commercial policy*, volume 1, pages 161–206. Elsevier.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

Grantham-McGregor, S. M., Walker, S. P., Himes, J. H., and Powell, C. A. (1996). Stunting and mental development in children. *Nutrition Research*, 16(11-12):1821–1828.

Grogan, L. (2023). Manufacturing employment and women's agency: Evidence from lesotho 2004–2014. *Journal of Development Economics*, 160:102951.

Guariso, A., Ngabitsinze, C., and Verpoorten, M. (2011). The Rwandan coffee sector: Out of the ordinary. *L'Afrique des grands lacs: Annuaire*, 2012:397–426.

Guariso, A. and Verpoorten, M. (2018). Aid, trade and the post-war recovery of the Rwandan coffee sector. *Journal of eastern African studies*, 12(3):552–574.

Heckman, J., Pinto, R., and Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*, 103(6):2052–2086.

Hoynes, H., Schanzenbach, D. W., and Almond, D. (2016). Long-run impacts of childhood access to the safety net. *American Economic Review*, 106(4):903–934.

Hoynes, H. W. and Schanzenbach, D. W. (2009). Consumption responses to in-kind transfers: Evidence from the introduction of the food stamp program. *American Economic Journal: Applied Economics*, 1(4):109–139.

Kline, P. and Walters, C. R. (2016). Evaluating public programs with close substitutes: The case of head start. *The Quarterly Journal of Economics*, 131(4):1795–1848.

Macchiavello, R. and Morjaria, A. (2015). Fully washed coffee exports in rwanda: Market structure and policy implications. *International Growth Centre, Policy Brief.*

Macchiavello, R. and Morjaria, A. (2021). Competition and relational contracts in the rwanda coffee chain. *The Quarterly Journal of Economics*, 136(2):1089–1143.

Macchiavello, R. and Morjaria, A. (2022). Acquisitions, management, and efficiency in rwanda's coffee industry. Technical report, National Bureau of Economic Research.

Mendez, M. A. and Adair, L. S. (1999). Severity and timing of stunting in the first two years of life affect performance on cognitive tests in late childhood. *The Journal of nutrition*, 129(8):1555–1562.

Mensah, J. T., Adu, G., Amuakwa-Mensah, F., and Marbuah, G. (2019). Mines and minds: Mining boom and human capital accumulation. Technical report.

Picarelli, N. (2016). Who really benefits from export processing zones? evidence from nicaraguan municipalities. *Labour Economics*, 41:318–332.

Rodrik, D. (2022). An industrial policy for good jobs. *Hamilton Project, Brookings Institution.*

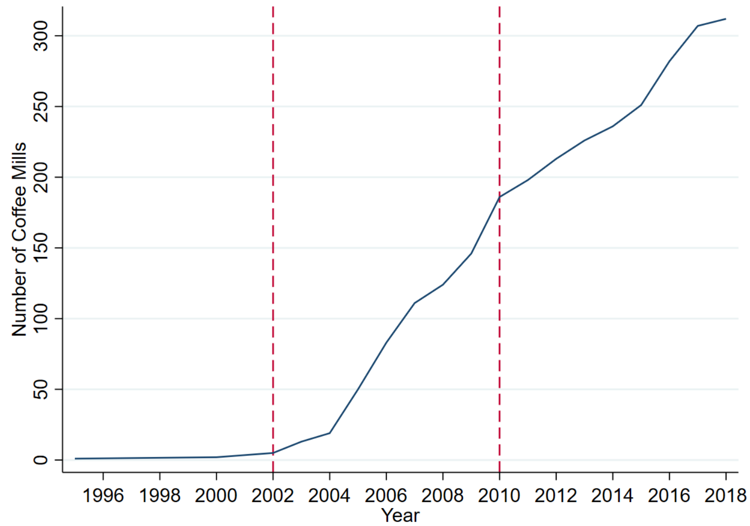
Sanin, D. (2021). Paid work for women and domestic violence: Evidence from the rwandan coffee mills. Technical report, Technical report, Working Paper.

Von der Goltz, J. and Barnwal, P. (2019). Mines: The local wealth and health effects of mineral mining in developing countries. *Journal of Development Economics*, 139:1–16.

Wang, J. (2013). The economic impact of special economic zones: Evidence from chinese municipalities. *Journal of development economics*, 101:133–147.

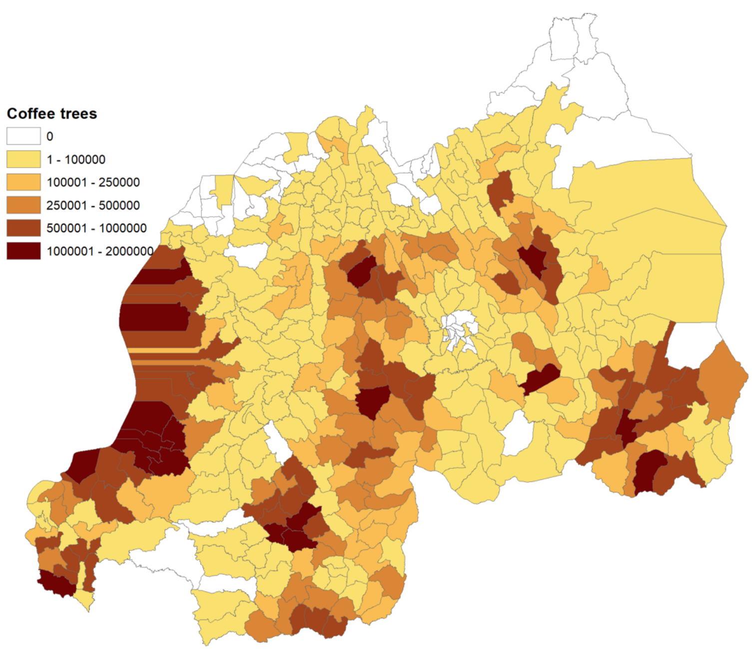
Figures

Figure 1: Total Number of Coffee Mills



Note: The above figure shows the evolution of Coffee Mills between 1999 to 2018 in Rwanda.

Figure 2: Distribution of coffee trees in Rwanda



Note: The figure shows the distribution of coffee trees across municipalities in Rwanda

Figure 3: Spatial distribution of primary schools

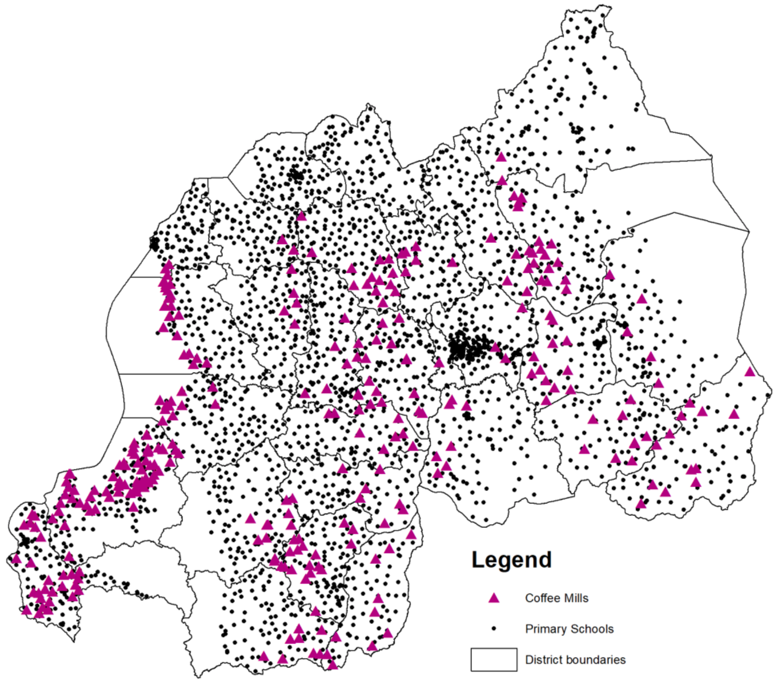
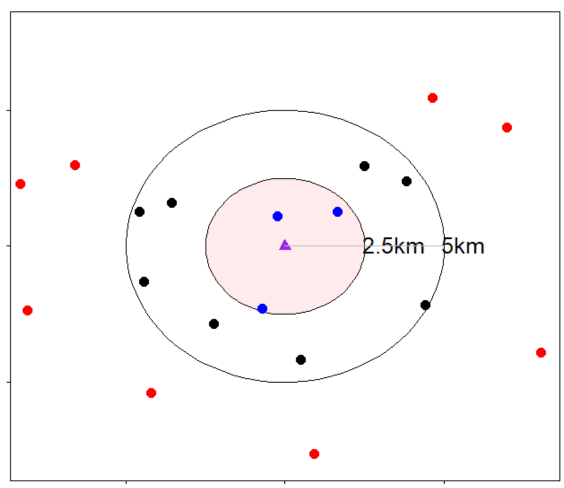
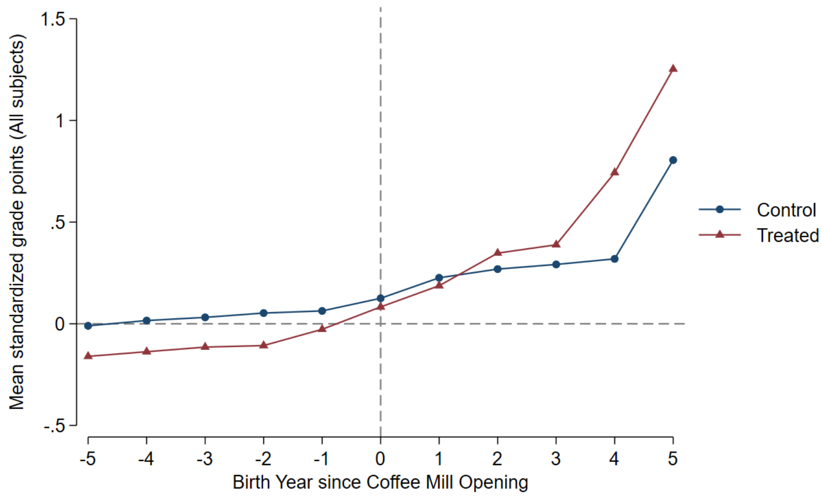


Figure 4: Graphical Representation of Catchment Area of Coffee Mills



This figure is a graphical depiction of our definition of the catchment area of the coffee mill. The inner circle depicts our baseline measure of the catchment area using a 2.5km radius around a coffee mill. The outer circle depicts the alternative measure of the catchment area using a 5km radius around a coffee mill. The dots (blue, black, red) depict school locations while the triangle depicts a coffee mill. Hence using the baseline measure of catchment area, the blue dots represent treated locations while the black and red dots represent the control locations.

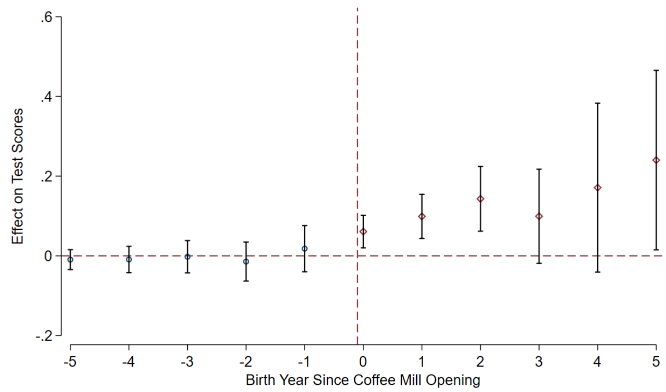
Figure 5: Trends in Students' Test Scores



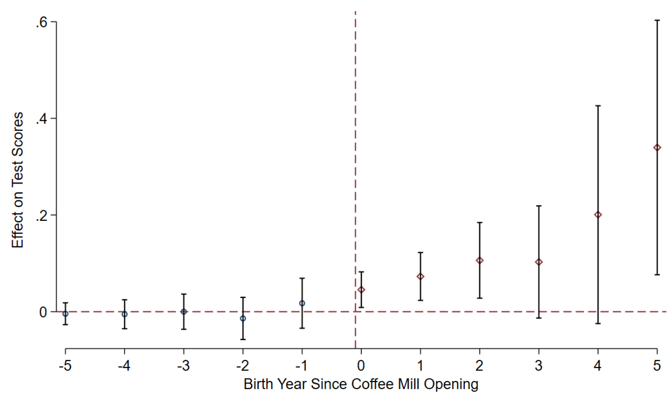
Note: This figure illustrates the trends in test scores of students in control and treated villages born before and after the opening of the nearest coffee mill.

Figure 6: Event Study: Exposure to Coffee Mills and Test Scores

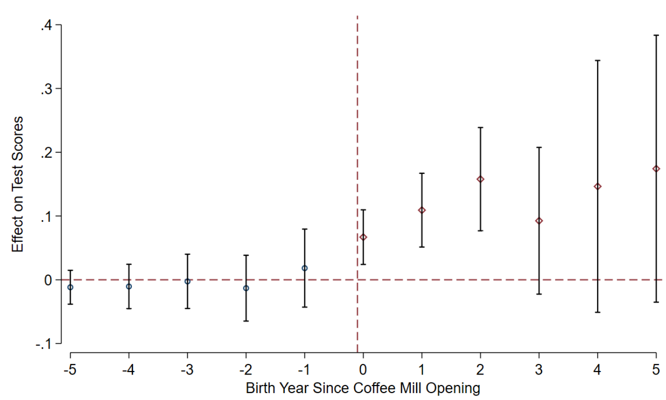
(a) All Subjects



(b) STEM



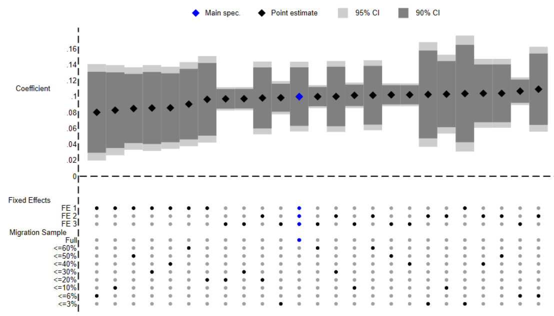
(c) Non-STEM



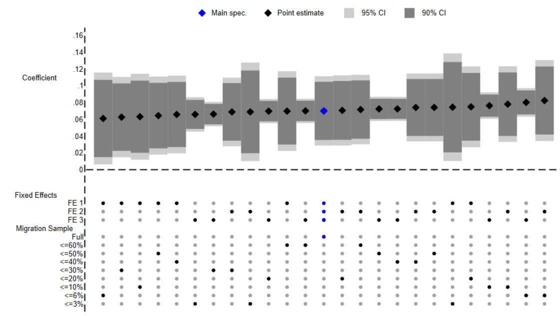
Notes: The figure shows point estimates and 95% confidence interval of the effect of early-life exposure to coffee mills and test scores using the Borusyak et al. (2024) imputation estimator.

Figure 7: Coffee Mills and Students Performance: Selective Migration

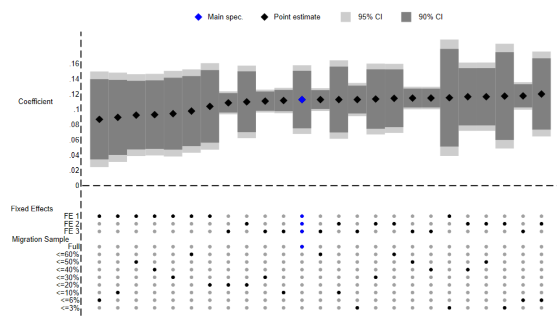
(a) All Subjects



(b) STEM



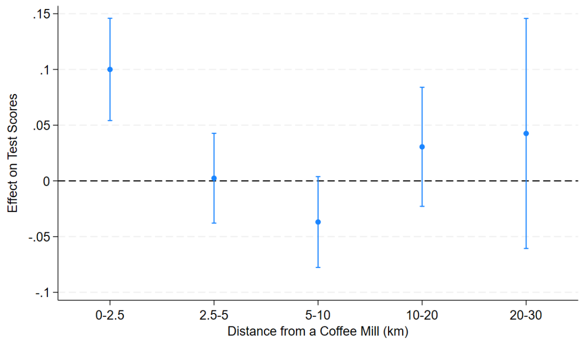
(c) Non-STEM



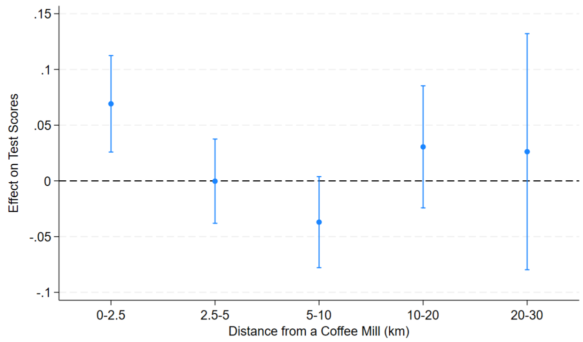
This figure presents estimates of the effect of coffee mills on student test scores across the various specifications outlined in Table 2 while trimming the sample according to the baseline migration rates at the municipality level. This is done to assess the robustness of the estimates to selective migration of people induced by the opening of coffee mills. FE1, FE2, and FE3 correspond to the fixed effects used in columns 1, 2, and 3 of Table 2 respectively. The estimates marked in blue are from the baseline while the remaining are from the respective sample trim and fixed effect specifications.

Figure 8: Spatial Lag Model

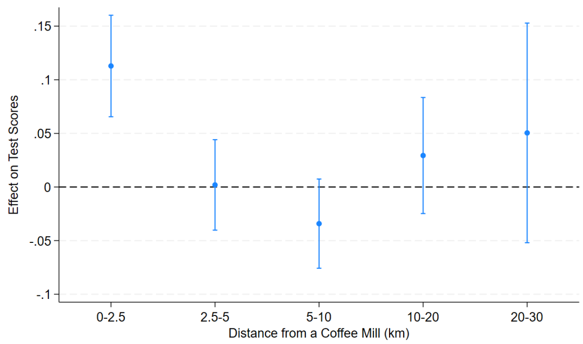
(a) All Subjects



(b) STEM



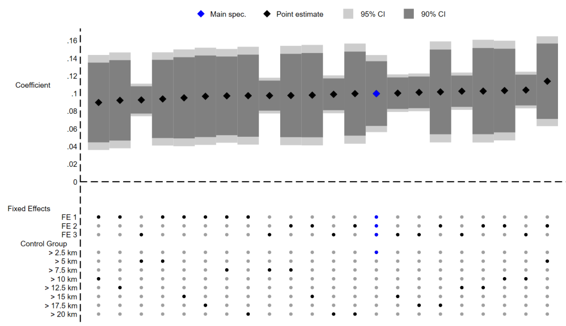
(c) Non-STEM



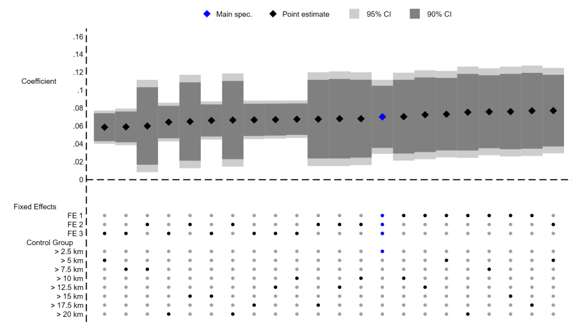
This figure plots point estimates and 95% confidence intervals along the respective distance splines. The control group is students living beyond 30km from a coffee mill.

Figure 9: Coffee Mills and Students Performance: Addressing Spillovers

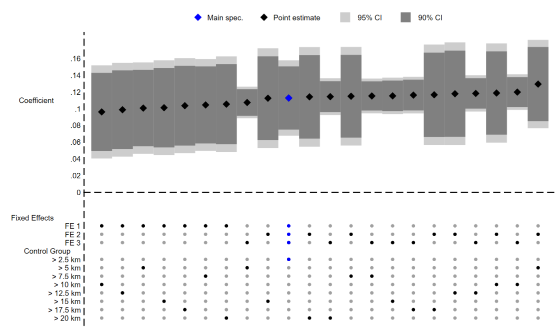
(a) All Subjects



(b) STEM



(c) Non-STEM



This figure presents estimates of the effect of coffee mills on student test scores across the various specifications outlined in Table 2 while trimming the control group using the distance thresholds specified in the respective sections. This is done to assess the robustness of the estimates to the selective migration of people induced by the opening of coffee mills. FE1, FE2, and FE3 correspond to the fixed effects used in columns 1, 2, and 3 of Table 2 respectively. The estimates marked in blue are from the baseline while the remaining are from the respective sample trim and fixed effect specifications.

Tables: Main Results

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Female	0.55	0.497	0	1	1492519
Age	14.256	1.665	9	21	1492520
Rural	0.881	0.324	0	1	1491692
CatchmentArea	0.187	0.39	0	1	1492520
CatchmentArea X Post	0.016	0.124	0	1	1492520
Standardized Grade Points in:					
All subjects	-0.038	0.964	-1.459	5.367	1492516
STEM subjects	-0.037	0.959	-1.05	6.503	1492519
Non-STEM subjects	-0.037	0.969	-1.678	4.419	1492515
Unstandardized Grade Points in:					
All subjects	15.421	7.675	5	45	1492516
STEM subjects	4.867	3.144	2	18	1492519
Non-STEM subjects	10.554	4.818	3	27	1492515
Standardized Test Scores in:					
Math	-0.035	0.959	-1.295	8.511	1492520
Science	-0.036	0.963	-1.542	7.121	1492520
Social studies	-0.037	0.966	-1.436	5.102	1492520
English	-0.044	0.948	-2.226	6.826	1492520
Kinyarwanda	-0.022	0.995	-3.406	3.801	1492520
Baseline attributes:					
Road density (m/km2)	12.67	9.500	0	34.835	1492520
Coffee density in 1999 (trees/km2)	3338.4	5034.6	0	33542.6	1492520
Distance to closest main city	24.119	17.328	0.544	91.796	1487362
Health Data:					
Height-for-age (HAZ)	-1.71	1.43	-5.96	5.99	22794
Weight-for-Height (WHZ)	0.34	1.17	-4.91	4.97	22794
Weight-for-age (WAZ)	-0.74	1.11	-5.69	4.06	22796
Stunting	0.42	0.49	0	1	22794
Wasting	0.03	0.17	0	1	22794
Underweight	0.12	0.32	0	1	22796

Table 2: Early Life Exposure to Coffee Mills and Test Scores

	All Subjects			STEM			Non STEM		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	A. Standardized Grade Points								
CatchmentArea X Post	0.0905***	0.1015***	0.0882***	0.0699***	0.0717***	0.0595***	0.0976***	0.1143***	0.1009***
	(0.0269)	(0.0224)	(0.0214)	(0.0242)	(0.0211)	(0.0206)	(0.0278)	(0.0230)	(0.0219)
	[0.002]	[0.001]	[0.001]	[0.006]	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]
R-squared	0.3866	0.4260	0.4261	0.3796	0.4152	0.4153	0.3610	0.4030	0.4031
	B. Grade Points (unstandardized)								
CatchmentArea X Post	0.7162***	0.7879***	0.6791***	0.2449***	0.2219***	0.1824***	0.4713***	0.5660***	0.4967***
	(0.2058)	(0.1702)	(0.1633)	(0.0728)	(0.0623)	(0.0614)	(0.1373)	(0.1134)	(0.1080)
	[0.002]	[0.001]	[0.001]	[0.002]	[0.001]	[0.002]	[0.002]	[0.001]	[0.001]
Mean dep var	15.4019	15.4015	15.4015	4.8608	4.8607	4.8607	10.5410	10.5408	10.5408
R-squared	0.4085	0.4478	0.4478	0.4092	0.4451	0.4452	0.3768	0.4189	0.4190
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exam Year FE	Yes	No	No	Yes	No	No	Yes	No	No
YOB FE	Yes	No	No	Yes	No	No	Yes	No	No
Municipality X Exam Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Municipality X YOB FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
CatchmentArea X Exam Year FE	No	No	Yes	No	No	Yes	No	No	Yes
R-squared	0.3866	0.4260	0.4260	0.3796	0.4152	0.4152	0.3609	0.4030	0.4030
Observations	1487355	1487036	1487036	1487358	1487039	1487039	1487354	1487035	1487035

Notes: In panel A, the dependent variable is the standardized aggregate grade points in all subjects in the Rwanda Primary School National exam. The dependent variable in panel B is the (unstandardized) aggregate grade points in all subjects. STEM subjects include Math and Science while Non STEM courses include English, Social Studies, and Kinyarwanda. CatchmentArea X Post is an indicator variable equal to 1 if a student was born within the vicinity (≤ 2.5km) of an operational coffee mill and 0 if otherwise. Controls include gender of the student, average annual temperature and precipitation in the exam year, as well as baseline controls such as coffee suitability and road density interacted with time trend. Robust standard errors clustered at school level in parenthesis. False discovery rate (FDR) adjusted p-values for CatchmentArea X Post are reported in square brackets (Anderson 2008). The outcome variables accounted for in the p-values adjustments include the standardized test scores of: All subjects (aggregate), STEM, and Non-STEM subjects, as well as the subject-specific test scores i.e., Maths, Science, Social, English, and Kinyarwanda.
* Significant at 10 percent level ** Significant at 5 percent level *** Significant at 1 percent level

Table 3: Early Life Exposure to Coffee Mills and Test Scores: Effects by Subjects

	All Subjects			STEM					
	Aggregate			Maths			Science		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
CatchmentArea X Post	0.0905***	0.1015***	0.0882***	0.0700***	0.0566**	0.0472**	0.0724***	0.0904***	0.0789***
	(0.0269)	(0.0224)	(0.0214)	(0.0231)	(0.0225)	(0.0221)	(0.0268)	(0.0222)	(0.0213)
	[0.002]	[0.001]	[0.001]	[0.004]	[0.015]	[0.016]	[0.009]	[0.001]	[0.001]
R-squared	0.3866	0.4260	0.4261	0.3459	0.3764	0.3765	0.3783	0.4191	0.4191
Observations	1487355	1487036	1487036	1487359	1487040	1487040	1487359	1487040	1487040
	Non STEM								
	Social			English			Kinyarwanda		
CatchmentArea X Post	0.0728**	0.0848***	0.0760***	0.0935***	0.1054***	0.0914***	0.0934***	0.1115***	0.1006***
	(0.0287)	(0.0230)	(0.0223)	(0.0260)	(0.0206)	(0.0195)	(0.0254)	(0.0232)	(0.0225)
	[0.014]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
R-squared	0.3842	0.4227	0.4227	0.4057	0.4469	0.4470	0.2317	0.2767	0.2768
Observations	1487359	1487040	1487040	1487359	1487040	1487040	1487359	1487040	1487040
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exam Year FE	Yes	No	No	Yes	No	No	Yes	No	No
YOB FE	Yes	No	No	Yes	No	No	Yes	No	No
Municipality X Exam Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Municipality X YOB FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
CatchmentArea X Exam Year FE	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Dependent variable is the standardized test scores in the respective subject groups in the Rwanda Primary School National exam. STEM subjects include Math and Science while Non STEM courses include English, Social Studies, and Kinyarwanda. CatchmentArea X Post is an indicator variable equal to 1 if a student was born within the vicinity (≤ 2.5km) of an operational coffee mill and 0 if otherwise. Controls include gender of the student, average annual temperature and precipitation in the exam year, as well as baseline controls such as coffee suitability and road density interacted with time trend. Robust standard errors clustered at school level in parenthesis. False discovery rate (FDR) adjusted p-values for CatchmentArea X Post are reported in square brackets (Anderson 2008). The outcome variables accounted for in the p-values adjustments include the standardized test scores of: All subjects (aggregate), STEM, and Non-STEM subjects, as well as the subject-specific test scores i.e., Maths, Science, Social, English, and Kinyarwanda.
* Significant at 10 percent level ** Significant at 5 percent level *** Significant at 1 percent level

Table 4: Early Life Exposure to Coffee Mills and Test Scores: Effects by Gender

	All Subjects			STEM			Non STEM		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Male				
CatchmentArea X Post	0.0951***	0.1022***	0.0932***	0.0784***	0.0739***	0.0645**	0.0998***	0.1144***	0.1059***
	(0.0317)	(0.0280)	(0.0272)	(0.0303)	(0.0271)	(0.0267)	(0.0317)	(0.0282)	(0.0274)
	[0.004]	[0.001]	[0.001]	[0.012]	[0.009]	[0.009]	[0.003]	[0.001]	[0.001]
R-squared	0.3793	0.4222	0.4222	0.3699	0.4099	0.4100	0.3574	0.4026	0.4027
Observations	668639	668285	668285	668639	668285	668285	668638	668284	668284
					Female				
CatchmentArea X Post	0.0837***	0.0973***	0.0801***	0.0604***	0.0662***	0.0513***	0.0929***	0.1108***	0.0932***
	(0.0257)	(0.0209)	(0.0202)	(0.0222)	(0.0196)	(0.0193)	(0.0273)	(0.0219)	(0.0210)
	[0.002]	[0.001]	[0.001]	[0.009]	[0.002]	[0.002]	[0.002]	[0.001]	[0.001]
R-squared	0.3984	0.4432	0.4433	0.3889	0.4300	0.4300	0.3713	0.4188	0.4189
Observations	818714	818396	818396	818717	818399	818399	818714	818396	818396
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exam Year FE	Yes	No	No	Yes	No	No	Yes	No	No
YOB FE	Yes	No	No	Yes	No	No	Yes	No	No
Municipality X Exam Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Municipality X YOB FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
CatchmentArea X Exam Year FE	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Dependent variable is the standardized aggregate test scores in the respective subject groups in the Rwanda Primary School National exam. STEM subjects include Math and Science while Non STEM courses include English, Social Studies, and Kinyarwanda. CatchmentArea X Post is an indicator variable equal to 1 if a student was born within the vicinity (≤ 2.5km) of an operational coffee mill and 0 if otherwise. Controls include gender of the student, average annual temperature, and precipitation in the exam year, as well as baseline controls such as coffee suitability and road density interacted with time trend. Robust standard errors clustered at school level in parenthesis. False discovery rate (FDR) adjusted p-values for CatchmentArea X Post are reported in square brackets (Anderson 2008). The outcome variables accounted for in the p-values adjustments include the standardized test scores of: All subjects (aggregate), STEM, and Non-STEM subjects, as well as the subject-specific test scores i.e., Maths, Science, Social, English, and Kinyarwanda.
* Significant at 10 percent level ** Significant at 5 percent level *** Significant at 1 percent level

Table 5: Early Life Exposure to Coffee Mills and Test Scores: Effects by Location

	All Subjects			STEM			Non STEM		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Rural				
CatchmentArea X Post	0.0799***	0.0890***	0.0730***	0.0604**	0.0585***	0.0450**	0.0867***	0.1027***	0.0860***
	(0.0273)	(0.0229)	(0.0217)	(0.0241)	(0.0211)	(0.0203)	(0.0286)	(0.0239)	(0.0225)
	[0.005]	[0.001]	[0.001]	[0.015]	[0.008]	[0.009]	[0.004]	[0.001]	[0.001]
R-squared	0.3112	0.3588	0.3588	0.3017	0.3450	0.3451	0.2925	0.3423	0.3424
Observations	1314120	1313807	1313807	1314123	1313810	1313810	1314119	1313806	1313806
					Urban				
CatchmentArea X Post	0.1524	0.1509	0.1525	0.1562	0.1357	0.1339	0.1458	0.1584	0.1615
	(0.1363)	(0.1245)	(0.1295)	(0.1335)	(0.1331)	(0.1367)	(0.1328)	(0.1174)	(0.1225)
	[0.275]	[0.242]	[0.241]	[0.256]	[0.309]	[0.309]	[0.28]	[0.196]	[0.194]
R-squared	0.5513	0.5717	0.5718	0.5355	0.5554	0.5556	0.5254	0.5474	0.5475
Observations	172407	172371	172371	172407	172371	172371	172407	172371	172371
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exam Year FE	Yes	No	No	Yes	No	No	Yes	No	No
YOB FE	Yes	No	No	Yes	No	No	Yes	No	No
Municipality X Exam Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Municipality X YOB FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
CatchmentArea X Exam Year FE	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Dependent variable is the standardized aggregate test scores in the respective subject groups in the Rwanda Primary School National exam. STEM subjects include Math and Science while Non STEM courses include English, Social Studies, and Kinyarwanda. CatchmentArea X Post is an indicator variable equal to 1 if a student was born within the vicinity (≤ 2.5km) of an operational coffee mill and 0 if otherwise. Controls include gender of the student, average annual temperature and precipitation in the exam year, as well as baseline controls such as coffee suitability and road density interacted with time trend. Robust standard errors clustered at school level in parenthesis. False discovery rate (FDR) adjusted p-values for CatchmentArea X Post are reported in square brackets (Anderson 2008). The outcome variables accounted for in the p-values adjustments include the standardized test scores of: All subjects (aggregate), STEM, and Non-STEM subjects, as well as the subject-specific test scores i.e., Maths, Science, Social, English, and Kinyarwanda.
* Significant at 10 percent level ** Significant at 5 percent level *** Significant at 1 percent level

Table 6: Early Life Exposure to Coffee Mills and Test Scores: Alternative Measure of Exposure

	All Subjects			STEM			Non STEM		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Exposure (% of age 0-5)	0.0336**	0.0136	0.0336***	0.0242*	0.0073	0.0215*	0.0370**	0.0166	0.0393***
	(0.0152)	(0.0135)	(0.0129)	(0.0137)	(0.0125)	(0.0122)	(0.0159)	(0.0142)	(0.0133)
	[0.173]	[0.433]	[0.433]	[0.233]	[0.59]	[0.581]	[0.173]	[0.433]	[0.433]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exam Year FE	Yes	No	No	Yes	No	No	Yes	No	No
YOB FE	Yes	No	No	Yes	No	No	Yes	No	No
Municipality X Exam Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Municipality X YOB FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
CatchmentArea X Exam Year FE	No	No	Yes	No	No	Yes	No	No	Yes
R-squared	0.3866	0.4260	0.4260	0.3796	0.4152	0.4152	0.3609	0.4029	0.4030
Observations	1487355	1487036	1487036	1487358	1487039	1487039	1487354	1487035	1487035

Notes: Dependent variable is the standardized aggregate test scores in the respective subject groups in the Rwanda Primary School National exam. STEM subjects include Math and Science while Non STEM courses include English, Social Studies, and Kinyarwanda. Exposure (% of age 0-5) measures the share of a years between conception and age 5 that a student was exposed to an operational coffee mill (within ($\leq 5km$) of locality). Controls include gender of the student, average annual temperature, and precipitation in the exam year, as well as baseline controls such as coffee suitability and road density interacted with time trend. Robust standard errors clustered at the school level in parenthesis. False discovery rate (FDR) adjusted p-values for Exposure (% of age 0-5) are reported in square brackets (Anderson 2008). The outcome variables accounted for in the p-values adjustments include the standardized test scores of: All subjects (aggregate), STEM, and Non-STEM subjects, as well as the subject-specific test scores i.e., Maths, Science, Social, English, and Kinyarwanda.

* Significant at 10 percent level ** Significant at 5 percent level *** Significant at 1 percent level

Tables: Mechanisms

Table 7: Exposure to Coffee Mills and Household Wealth

	Wealth Score		Poor (0/1)	
	(1)	(2)	(3)	(4)
CatchmentArea X Post	0.0547**	0.0252	-0.0435***	-0.0293*
	(0.0256)	(0.0269)	(0.0163)	(0.0166)
CatchmentArea	-0.0211	-0.0073	0.0184	0.0116
	(0.0280)	(0.0286)	(0.0186)	(0.0189)
Controls	No	Yes	No	Yes
Municipality FE	Yes	Yes	Yes	Yes
Province X Year FE	Yes	Yes	Yes	Yes
R-squared	0.5625	0.5640	0.2083	0.2095
Observations	23513	23513	23513	23513

Notes: CatchmentArea X Post is an indicator variable equal to 1 for a child living in a community within the vicinity of an operational coffee mill and 0 if otherwise. CatchmentArea is an indicator variable equal to 1 for a child living in a community within the vicinity of a coffee mill. Controls include rural-urban status, education level, and time trends interacted with coffee suitability and road density."
* Significant at 10 percent level ** Significant at 5 percent level *** Significant at 1 percent level

Table 8: Exposure to Coffee Mills and Child Health

	HAZ Score		Stunted (0/1)	
	(1)	(2)	(3)	(4)
CatchmentArea X Post	0.0967*	0.1021**	-0.0464***	-0.0456***
	(0.0511)	(0.0500)	(0.0171)	(0.0173)
CatchmentArea	-0.0108	-0.0634	0.0313	0.0443**
	(0.0590)	(0.0560)	(0.0192)	(0.0187)
Municipality FE	Yes	Yes	Yes	Yes
YOB FE	Yes	Yes	Yes	Yes
Province X Year FE	Yes	Yes	Yes	Yes
R-squared	0.2030	0.2475	0.1428	0.1745
Observations	22794	22794	22794	22794

Notes: CatchmentArea X Post is an indicator variable equal to 1 for a child living in a community within the vicinity of an operational coffee mill and 0 if otherwise. CatchmentArea is an indicator variable equal to 1 for a child living in a community within the vicinity of a coffee mill. Controls include gender of the child, twin status, kid birth order, age of mother, rural/urban, household wealth, and time trends interacted with coffee suitability and road density. Standard errors are clustered at the enumeration area level.
* Significant at 10 percent level ** Significant at 5 percent level *** Significant at 1 percent level

Table 9: Exposure to Coffee Mills and School Attendance

	Dep. Var: Child attended school during the current school year (0/1)					
	All		Rural		Urban	
	(1)	(2)	(3)	(4)	(5)	(6)
CatchmentArea X Post	0.0167**	0.0176**	0.0176**	0.0169**	0.0033	0.0035
	(0.0075)	(0.0076)	(0.0085)	(0.0086)	(0.0203)	(0.0214)
CatchmentArea	0.0036	0.0040	0.0084	0.0099	-0.0259	-0.0265
	(0.0083)	(0.0084)	(0.0094)	(0.0095)	(0.0228)	(0.0236)
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
YOB FE	Yes	Yes	Yes	Yes	Yes	Yes
Province X Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.0847	0.0852	0.0844	0.0847	0.1240	0.1260
Observations	42640	42346	35229	34996	7411	7350

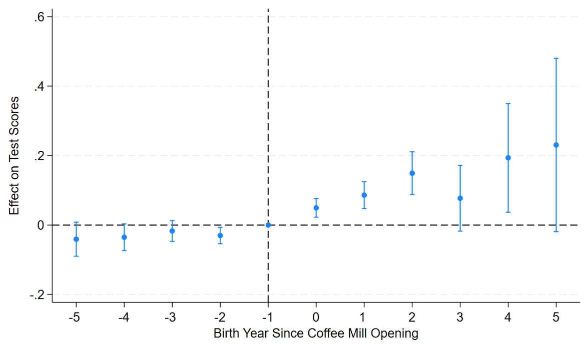
Notes: CatchmentArea X Post is an indicator variable equal to 1 for a child living in a community within the vicinity of an operational coffee mill and 0 if otherwise. CatchmentArea is an indicator variable equal to 1 for a child living in a community within the vicinity of a coffee mill. Controls include gender of the child, rural/urban, and time trends interacted with coffee suitability and road density. Standard errors are clustered at the enumeration area level.
* Significant at 10 percent level ** Significant at 5 percent level *** Significant at 1 percent level

B Online Appendix

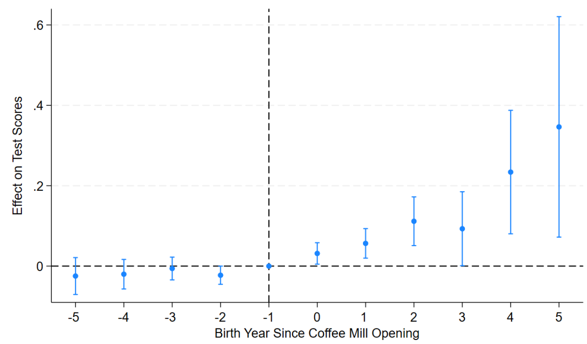
B.1 Additional Figures and Tables

Figure A1: Event Study: Exposure to Coffee Mills and Test Scores

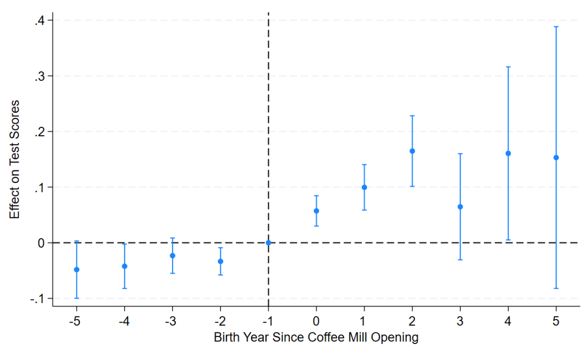
(a) All Subjects



(b) STEM



(c) Non-STEM



Notes: This figure presents events study estimates of the effects of exposure to coffee mills in early life on students' test scores using a two-way fixed effect (TWFE) estimator.

Table A1: Early Life Exposure to Coffee Mills and Test Scores: Sample excluding Kigali

	All Subjects			STEM			Non STEM		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
CatchmentArea X Post	0.0831***	0.0974***	0.0833***	0.0636***	0.0651***	0.0537***	0.0897***	0.1117***	0.0967***
	(0.0268)	(0.0226)	(0.0216)	(0.0239)	(0.0209)	(0.0205)	(0.0279)	(0.0234)	(0.0223)
	[0.003]	[0.001]	[0.001]	[0.01]	[0.003]	[0.004]	[0.003]	[0.001]	[0.001]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exam Year FE	Yes	No	No	Yes	No	No	Yes	No	No
YOB FE	Yes	No	No	Yes	No	No	Yes	No	No
Municipality X Exam Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Municipality X YOB FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
CatchmentArea X Exam Year FE	No	No	Yes	No	No	Yes	No	No	Yes
R-squared	0.3614	0.4033	0.4033	0.3542	0.3918	0.3918	0.3375	0.3820	0.3821
Observations	1410066	1409774	1409774	1410069	1409777	1409777	1410065	1409773	1409773

Notes: Dependent variable is the standardized aggregate test scores in the respective subject groups in the Rwanda Primary School National exam. STEM subjects include Math and Science while Non STEM courses include English, Social Studies, and Kinyarwanda. CatchmentArea X Post is an indicator variable equal to 1 if a student was born within the vicinity ($\leq$ 2.5km) of an operational coffee mill and 0 if otherwise. Controls include gender of the student, average annual temperature and precipitation in the exam year, as well as baseline controls such as coffee suitability and road density interacted with time trend. Robust standard errors clustered at school level in parenthesis. False discovery rate (FDR) adjusted p-values for CatchmentArea X Post are reported in square brackets (Anderson 2008). The outcome variables accounted for in the p-values adjustments include the standardized test scores of: All subjects (aggregate), STEM, and Non-STEM subjects, as well as the subject-specific test scores i.e., Maths, Science, Social, English, and Kinyarwanda.
* Significant at 10 percent level ** Significant at 5 percent level *** Significant at 1 percent level

Table A2: Coffee Mills and School Construction

	# New Schools Constructed				Total Schools			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coffee Mills (0/1)	-0.0140	-0.0206			-0.1159*	-0.0934		
	(0.0119)	(0.0126)			(0.0674)	(0.0669)		
# Coffee Mills Operating			-0.0034	-0.0043			-0.0610*	-0.0471
			(0.0058)	(0.0061)			(0.0349)	(0.0341)
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	No	Yes	No	Yes	No	Yes	No
Province-Year FE	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.0898	0.1022	0.0897	0.1020	0.9455	0.9484	0.9456	0.9485
Observations	7218	7218	7218	7218	7218	7218	7218	7218

Notes:
* Significant at 10 percent level ** Significant at 5 percent level *** Significant at 1 percent level

B.2 Using an Alternative Measure of Catchment Area

So far, our baseline analysis relies on a 2.5km buffer around the location of coffee mills in identifying treated and control communities. In this section, we relax this strict definition of the catchment area using a 5km radius from a coffee mill and replicate the baseline analysis. Results are shown in Tables A3 - A7

Table A3: Early Life Exposure to Coffee Mills and Test Scores

	All Subjects			STEM			Non STEM		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	A. Standardized Grade Points								
CatchmentArea X Post	0.0492***	0.0522***	0.0487***	0.0353**	0.0372**	0.0337**	0.0540***	0.0571***	0.0537***
	(0.0178)	(0.0171)	(0.0169)	(0.0159)	(0.0161)	(0.0159)	(0.0187)	(0.0180)	(0.0176)
	[0.015]	[0.013]	[0.014]	[0.042]	[0.035]	[0.047]	[0.014]	[0.013]	[0.013]
R-squared	0.3866	0.4260	0.4260	0.3796	0.4152	0.4152	0.3609	0.4030	0.4030
	B. Grade Points (unstandardized)								
CatchmentArea X Post	0.4057***	0.3944***	0.3633***	0.1482***	0.1140**	0.1019**	0.2576***	0.2804***	0.2614***
	(0.1358)	(0.1313)	(0.1291)	(0.0475)	(0.0476)	(0.0471)	(0.0920)	(0.0888)	(0.0871)
	[0.013]	[0.013]	[0.014]	[0.013]	[0.031]	[0.045]	[0.014]	[0.013]	[0.013]
Mean dep var	15.4019	15.4015	15.4015	4.8608	4.8607	4.8607	10.5410	10.5408	10.5408
R-squared	0.4085	0.4477	0.4478	0.4092	0.4451	0.4451	0.3768	0.4189	0.4189
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exam Year FE	Yes	No	No	Yes	No	No	Yes	No	No
YOB FE	Yes	No	No	Yes	No	No	Yes	No	No
Municipality X Exam Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Municipality X YOB FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
CatchmentArea X Exam Year FE	No	No	Yes	No	No	Yes	No	No	Yes
R-squared	0.3866	0.4260	0.4260	0.3796	0.4152	0.4152	0.3609	0.4030	0.4030
Observations	1487355	1487036	1487036	1487358	1487039	1487039	1487354	1487035	1487035

Notes: In panel A, the dependent variable is the standardized aggregate grade points in all subjects in the Rwanda Primary School National exam. The dependent variable in panel B is the (unstandardized) aggregate grade points in all subjects. STEM subjects include Math and Science while Non STEM courses include English, Social Studies, and Kinyarwanda. CatchmentArea X Post is an indicator variable equal to 1 if a student was born within the vicinity (≤ 5km) of an operational coffee mill and 0 if otherwise. Controls include gender of the student, average annual temperature and precipitation in the exam year, as well as baseline controls such as coffee suitability and road density interacted with time trend. Robust standard errors clustered at school level in parenthesis. False discovery rate (FDR) adjusted p-values for $CatchmentAreaXPost$ are reported in square brackets (Anderson 2008). The outcome variables accounted for in the p-values adjustments include the standardized test scores of: All subjects (aggregate), STEM, and Non-STEM subjects, as well as the subject-specific test scores i.e., Maths, Science, Social, English, and Kinyarwanda.
* Significant at 10 percent level ** Significant at 5 percent level *** Significant at 1 percent level

Table A4: Early Life Exposure to Coffee Mills and Test Scores: Effects by Subjects

	All Subjects			STEM					
	Aggregate			Maths			Science		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
CatchmentArea X Post	0.0492***	0.0522***	0.0487***	0.0414***	0.0376**	0.0352**	0.0273	0.0366**	0.0341**
	(0.0178)	(0.0171)	(0.0169)	(0.0156)	(0.0170)	(0.0167)	(0.0176)	(0.0170)	(0.0169)
	[0.015]	[0.013]	[0.014]	[0.018]	[0.042]	[0.048]	[0.138]	[0.045]	[0.056]
R-squared	0.3866	0.4260	0.4260	0.3459	0.3764	0.3764	0.3782	0.4190	0.4190
Observations	1487359	1487040	1487040	1487359	1487040	1487040	1487359	1487040	1487040
	Non STEM								
	Social			English			Kinyarwanda		
CatchmentArea X Post	0.0330*	0.0321*	0.0294*	0.0486***	0.0546***	0.0488***	0.0594***	0.0639***	0.0631***
	(0.0189)	(0.0173)	(0.0172)	(0.0172)	(0.0168)	(0.0165)	(0.0184)	(0.0186)	(0.0183)
	[0.098]	[0.08]	[0.103]	[0.014]	[0.013]	[0.013]	[0.013]	[0.013]	[0.013]
R-squared	0.3841	0.4226	0.4226	0.4056	0.4469	0.4469	0.2316	0.2767	0.2767
Observations	1487359	1487040	1487040	1487359	1487040	1487040	1487359	1487040	1487040
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exam Year FE	Yes	No	No	Yes	No	No	Yes	No	No
YOB FE	Yes	No	No	Yes	No	No	Yes	No	No
Municipality X Exam Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Municipality X YOB FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
CatchmentArea X Exam Year FE	No	No	Yes	No	No	Yes	No	No	Yes

Notes: The dependent variable is the standardized test scores in the respective subject groups in the Rwanda Primary School National exam. STEM subjects include Math and Science while Non STEM courses include English, Social Studies, and Kinyarwanda. CatchmentArea X Post is an indicator variable equal to 1 if a student was born within the vicinity (≤ 5km) of an operational coffee mill and 0 if otherwise. Controls include gender of the student, average annual temperature, and precipitation in the exam year, as well as baseline controls such as coffee suitability and road density interacted with time trend. Robust standard errors clustered at the school level in parenthesis. False discovery rate (FDR) adjusted p-values for $CatchmentAreaXPost$ are reported in square brackets (Anderson 2008). The outcome variables accounted for in the p-values adjustments include the standardized test scores of: All subjects (aggregate), STEM, and Non-STEM subjects, as well as the subject-specific test scores i.e., Maths, Science, Social, English, and Kinyarwanda.
* Significant at 10 percent level ** Significant at 5 percent level *** Significant at 1 percent level

Table A5: Early Life Exposure to Coffee Mills and Test Scores: Effects by Gender

	All Subjects			STEM			Non STEM		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Male				
CatchmentArea X Post	0.0597***	0.0622***	0.0607***	0.0460**	0.0492**	0.0466**	0.0636***	0.0656***	0.0648***
	(0.0210)	(0.0217)	(0.0215)	(0.0195)	(0.0210)	(0.0208)	(0.0216)	(0.0222)	(0.0220)
	[0.015]	[0.015]	[0.015]	[0.034]	[0.035]	[0.043]	[0.015]	[0.015]	[0.015]
R-squared	0.3792	0.4221	0.4222	0.3698	0.4099	0.4099	0.3574	0.4026	0.4026
Observations	668639	668285	668285	668639	668285	668285	668638	668284	668284
					Female				
CatchmentArea X Post	0.0379**	0.0409**	0.0345**	0.0243	0.0254*	0.0201	0.0434**	0.0464***	0.0397**
	(0.0171)	(0.0160)	(0.0158)	(0.0151)	(0.0148)	(0.0148)	(0.0181)	(0.0170)	(0.0167)
	[0.042]	[0.023]	[0.043]	[0.124]	[0.103]	[0.196]	[0.031]	[0.015]	[0.032]
R-squared	0.3983	0.4432	0.4432	0.3888	0.4299	0.4300	0.3712	0.4187	0.4187
Observations	818714	818396	818396	818717	818399	818399	818714	818396	818396
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exam Year FE	Yes	No	No	Yes	No	No	Yes	No	No
YOB FE	Yes	No	No	Yes	No	No	Yes	No	No
Municipality X Exam Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Municipality X YOB FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
CatchmentArea X Exam Year FE	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Dependent variable is the standardized aggregate test scores in the respective subject groups in the Rwanda Primary School National exam. STEM subjects include Math and Science while Non STEM courses include English, Social Studies, and Kinyarwanda. Mills X Post is an indicator variable equal to 1 if a student was born within the vicinity ($\leq 5km$) of an operational coffee mill and 0 if otherwise. Controls include gender of the student, average annual temperature and precipitation in the exam year, as well as baseline controls such as coffee suitability and road density interacted with time trend. Robust standard errors clustered at school level in parenthesis. False discovery rate (FDR) adjusted p-values for $MillsXPost$ are reported in square brackets (Anderson 2008). The outcome variables accounted for in the p-values adjustments include the standardized test scores of: All subjects (aggregate), STEM, and Non-STEM subjects, as well as the subject-specific test scores i.e., Maths, Science, Social, English, and Kinyarwanda.

* Significant at 10 percent level ** Significant at 5 percent level *** Significant at 1 percent level

Table A6: Early Life Exposure to Coffee Mills and Test Scores: Effects by Location

	All Subjects			STEM			Non STEM		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Rural				
CatchmentArea X Post	0.0459**	0.0505***	0.0447**	0.0342**	0.0353**	0.0306*	0.0493**	0.0557***	0.0495***
	(0.0187)	(0.0179)	(0.0175)	(0.0164)	(0.0167)	(0.0163)	(0.0197)	(0.0186)	(0.0182)
	[0.029]	[0.014]	[0.023]	[0.049]	[0.048]	[0.077]	[0.026]	[0.013]	[0.015]
R-squared	0.3112	0.3587	0.3588	0.3017	0.3450	0.3450	0.2925	0.3422	0.3423
Observations	1314120	1313807	1313807	1314123	1313810	1313810	1314119	1313806	1313806
					Urban				
CatchmentArea X Post	0.0432	0.0369	0.0415	0.0367	0.0045	0.0085	0.0466	0.0545	0.0588
	(0.0652)	(0.0599)	(0.0603)	(0.0626)	(0.0589)	(0.0594)	(0.0656)	(0.0667)	(0.0666)
	[0.535]	[0.559]	[0.525]	[0.572]	[0.939]	[0.898]	[0.516]	[0.455]	[0.42]
R-squared	0.5513	0.5717	0.5718	0.5355	0.5554	0.5555	0.5253	0.5474	0.5475
Observations	172407	172371	172371	172407	172371	172371	172407	172371	172371
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exam Year FE	Yes	No	No	Yes	No	No	Yes	No	No
YOB FE	Yes	No	No	Yes	No	No	Yes	No	No
Municipality X Exam Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Municipality X YOB FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
CatchmentArea X Exam Year FE	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Dependent variable is the standardized aggregate test scores in the respective subject groups in the Rwanda Primary School National exam. STEM subjects include Math and Science while Non STEM courses include English, Social Studies, and Kinyarwanda. Mills X Post is an indicator variable equal to 1 if a student was born within the vicinity ($\leq 5km$) of an operational coffee mill and 0 if otherwise. Controls include gender of the student, average annual temperature and precipitation in the exam year, as well as baseline controls such as coffee suitability and road density interacted with time trend. Robust standard errors clustered at school level in parenthesis. False discovery rate (FDR) adjusted p-values for $MillsXPost$ are reported in square brackets (Anderson 2008). The outcome variables accounted for in the p-values adjustments include the standardized test scores of: All subjects (aggregate), STEM, and Non-STEM subjects, as well as the subject-specific test scores i.e., Maths, Science, Social, English, and Kinyarwanda.

* Significant at 10 percent level ** Significant at 5 percent level *** Significant at 1 percent level

Table A7: Early Life Exposure to Coffee Mills and Test Scores: Alternative Measure of Exposure

	All Subjects			STEM			Non STEM		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Exposure (% of age 0-5)	0.0336**	0.0136	0.0336***	0.0242*	0.0073	0.0215*	0.0370**	0.0166	0.0393***
	(0.0152)	(0.0135)	(0.0129)	(0.0137)	(0.0125)	(0.0122)	(0.0159)	(0.0142)	(0.0133)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Exam Year FE	Yes	No	No	Yes	No	No	Yes	No	No
YOB FE	Yes	No	No	Yes	No	No	Yes	No	No
Municipality X Exam Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Municipality X YOB FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
CatchmentArea X Exam Year FE	No	No	Yes	No	No	Yes	No	No	Yes
R-squared	0.3866	0.4260	0.4260	0.3796	0.4152	0.4152	0.3609	0.4029	0.4030
Observations	1487355	1487036	1487036	1487358	1487039	1487039	1487354	1487035	1487035

Notes: Dependent variable is the standardized aggregate test scores in the respective subject groups in the Rwanda Primary School National exam. STEM subjects include Math and Science while Non STEM courses include English, Social Studies, and Kinyarwanda. Exposure (% of age 0-5) measures the share of a years between conception and age 5 that a student was exposed to an operational coffee mill (within ($\leq 5km$) of locality). Controls include gender of the student, average annual temperature and precipitation in the exam year, as well as baseline controls such as coffee suitability and road density interacted with time trend. Robust standard errors clustered at school level in parenthesis. False discovery rate (FDR) adjusted p-values for $MillsXPost$ are reported in square brackets (Anderson 2008). The outcome variables accounted for in the p-values adjustments include the standardized test scores of: All subjects (aggregate), STEM, and Non-STEM subjects, as well as the subject-specific test scores i.e., Maths, Science, Social, English, and Kinyarwanda.

* Significant at 10 percent level ** Significant at 5 percent level *** Significant at 1 percent level