Centro de Estudios®
Espinosa Yglesias
PROMOVEMOS LA IGUALDAD
DE OPORTUNIDADES

The potential effects of the COVID-19 pandemic on learning

Autores:

Luis Monroy-Gómez-Franco
City University of New York /
Centro de Estudios Espinosa Yglesias

Roberto Vélez Grajales
Centro de Estudios Espinosa Yglesias

Luis Felipe López-Calva
Programa de las Naciones Unidas para el Desarrollo

Documento de trabajo no.

08 / 2021

Centro auspiciado por: ESRU Fundación Espinosa Rugarcía

The potential effects of the COVID-19 pandemic on learning *

Luis Monroy-Gómez-Franco†, Roberto Vélez-Grajales‡, Luis F. López-Calva§

September, 2021

Abstract

In this paper, we use a new database for Mexico to model the possible long-run effects of the pandemic on learning. First, based on the framework of Neidhöfer et al. (2021), we estimate the loss of schooling due to the transition from in-person to remote learning using data from the National Survey on Social Mobility (ESRU-EMOVI-2017), census data and national statistics of COVID-19 incidence. In this estimation, we account for the attenuation capacity of households by considering the parental educational attainment and the economic resources available to the household in the calculation of the short-run cost. Secondly, we estimate the potential long-run consequences of this shock through a calibrated learning profile for five Mexican regions following Kaffenberger and Pritchett (2021). Assuming the distance learning policy adopted by the Mexican government is entirely effective, our results indicate that a learning loss equivalent to the learning during a third of a school year in the short run translates into a learning loss equivalent to an entire school year further up the educational career of students. On the other hand, if the policy was ineffective, the short-run loss increases to an entire school year and becomes a loss of two years of learning in the long run. Our results suggest substantial variation at the regional level, with the most affected region, the South experiencing a loss thrice as large as that of the least affected region, the Centre region.

JEL-Classification: I24, I25, J62, O15

Keywords: Learning, COVID-19, Mexico, Education

*We deeply appreciate the comments made by Miles Corak, Jonathan Conning, María de los Angeles Monroy Gómez-Franco, Guido Neidhöffer and Nuria Rodríguez-Planas to a previous draft of this document. We also thank the Centro de Estudios Espinosa Yglesias for its funding support. The findings, interpretations, and conclusions in this paper are entirely those of the authors. They do not necessarily represent the views of the UNDP.

†PhD. in Economics Candidate; The Graduate Center, CUNY. lmonroygomezfranco@gradcenter.cuny.edu

‡Director, Centro de Estudios Espinosa Yglesias

§United Nations Development Programme

Introduction

Social distancing policies have taken the forefront in the effort to reduce the spread of COVID-19. As a result, almost all countries affected by the pandemic suspended activities that require the congregation of large groups of people, among which in-person education has a central place. Primarily among the costs of this decision are the short-run learning losses experienced by students and the yet to be seen impact on their school trajectories. In this paper, we provide the first set of estimates of short-run and long-run costs in learning for the Mexican case, paying special attention to how these costs vary across subnational regions. Depending on the assumed effectiveness of the remote learning model, we estimate that, at the national level, the lower bound of the short-run learning cost of the pandemic is equivalent to a loss between a third and more than a whole school year of learning. The lower bound on the long-run costs lies between a gap equivalent to 1.27 school years of learning and 2.11 school years of learning with respect to the expected learning attainment.

However, the large regional variation in the size of the COVID-19 shocks and the household's attenuation capacity crisscross these results. For example, in the south region of the country, where parental educational attainment and economic resources are the lowest, the lower bound on the short-run costs lies between the equivalent to 0.47 and 1.12 school years of learning. In contrast, in the centre of the country, the lower bound of the short-run learning cost lie between 0.26 and 0.96 of a school year. We observe a similar pattern in the long-run gap with respect to the expected learning progression. In the south, this long-run gap between attained learning and expected learning levels is equivalent to 1.91 and 3.02 school years of learning, while it lies just half that level in the centre.

The Mexican case is interesting for several reasons. First, the country as a whole remained in distanced learning from the spring of 2020 to the fall of 2021[1]. This implies that children currently in the school system have experienced more than a year of distanced learning with limited public compensatory measures, thus displacing the major part of the adjustment costs to households. Given the relative importance of Mexico among developing economies, providing an estimation of the possible costs of the displacement in instruction mode can highlight the specificities of the pandemic effects on middle-income economies.

Secondly, Mexico is a country with high levels of inequality of opportunity and low social mobility (Vélez-Grajales et al., 2018; Monroy-Gómez-Franco et al., 2021; Vélez-Grajales and Monroy-Gómez-Franco, 2017) A significant determinant of this structural pattern is the

[1] In-person instruction re-started only through a hybrid mode in August 30 of 2021.

heightened role of private resources of origin determining the economic trajectory of a person, which arises from the low effectiveness of public systems to support the socioeconomic achievement of the majority of the Mexican population. Moreover, Mexico is a country with significant spatial disparities, where there is a substantial degree of regional heterogeneity in terms of both social mobility and inequality of opportunity (Delajara et al., 2021; Monroy-Gómez-Franco, 2021b). Thus, it is likely that the displacement to distance learning increased these inequalities and compounds these already significant regional differences.

These inequalities can be seen in the educational system, even before the pandemic hit. Although the Mexican educational system includes both public and private schools, the majority of the population attends the public sector at all levels of education, ranging from 88% of all students in primary education to 69% of all students in college (Secretaría de Educación Pública, 2020). The main reason for this is that public education is mostly free for all the country's inhabitants, and taxes at the federal level fund it. However, a student's experience in the private sector is very different from that of a student in the public sector. For example, data for 2014 (the most recent data) shows that while 100% of private schools had student dedicated full-bathroom installations, less than 70% of schools in rural and indigenous communities had them. Likewise, 90% of private schools had at least one functioning computer, and less than 60% of public schools had one. In schools in rural and indigenous communities, less than 40% of schools fulfil this criterion. Finally, 80% of private schools had internet access, while less than 40% of public schools had it, and almost no schools in rural and indigenous communities had an internet connection. So far, public programs have been unable to close these gaps (Miranda-López, 2018).

These characteristics of the Mexican educational system show that the abrupt transition to remote instruction put the majority of the population at a disadvantage as the infrastructure to perform the transition was not in place in most schools. The households with the necessary economic resources to acquire this infrastructure did so, but according to the latest data, they constitute less than 30% of the population (Instituto Nacional de Estadística y Geografía, 2021). Thus, the Mexican case is a clear example of how preexisting inequalities can interact with the pandemic shock, leading to more significant and more persistent inequalities in the absence of countervailing public policy interventions.

Our work represents an application of the framework developed in Monroy-Gómez-Franco (2021a) to assess educational disruptions' short and long-run effects on learning. The model translates the days of schooling loss as calculated by Neidhöfer et al. (2021) into the learning dimension space. Specifically, this loss represents the share of learning that a student is losing due to the transition, with respect to what she would have learned on a school year where no disruption occurred. We call this the short-run or immediate learning cost. Additionally, it expands the definition of households' attenuation capacity by including parental education and economic resources of the household of origin, whereas Neidhöfer et al. (2021) only consider the role played by parental education. This immediate cost is then introduced as a disruption in the learning progression of an individual represented by a calibrated Potential Pedagogical Function (PPF) as modelled by Kaffenberger and Pritchett (2020b, 2021). The difference between the stock of learning produced by the undisturbed PPF and the one produced by the disturbed PPF is the long-run cost of the shock in learning. This gap can be expressed in terms of school years of learning by dividing it by the year to year learning progression implied by the calibrated PPF.

In contrast with Neidhöfer et al. (2021), our estimates do not translate directly into the effect of the pandemic on intergenerational educational mobility, as they are in the learning dimension. Although both concepts are linked, estimates on intergenerational educational mobility based on attained school years can be affected by policy measures that grant progression through the school system regardless of learning attainments. Policies of this type can produce a divergence between estimates on educational mobility and learning attainments, leading the first to underestimate the long-run effects of the pandemic.

The pandemic and online learning

Given the ease of aerial transmission of the SARS-COV-2 virus, a common public policy to reduce the virus spread was the suspension of in-person classes to be replaced by some form of distance learning. The sudden transition to remote education forced households to adjust their routines and use available resources to accommodate the new circumstances. However, the capacity of adjustment is not evenly distributed across households and depends, among other variables, on the amount of resources available to the family. Households with fewer resources faced more complications transitioning to an entirely remote learning environment. In the cases of the US. and the UK, these complications arise primarily from the lack of a stable internet connection, a computer suitable enough for educational tasks (Bansak and Starr, 2021; Francis and Weller, 2021; Andrew et al., 2020) and from more considerable exposure to other types of shocks caused by the pandemic such as the shutdown of care facilities and income earner unemployment (Rodríguez-Planas, 2020). Similar factors have been found to diminish household coping capacity and accessibility to remote education in developing countries (Hossain, 2021).

Bansak and Starr (2021); Andrew et al. (2020); Dietrich et al. (2021) find that in the US, the UK and Germany, parents increased the time they spent with their children in study activities to compensate for the reduction in engagement with the schools. In addition, evidence for the US (Bacher-Hicks et al., 2021) and Denmark (Jæger and Blaabæk, 2020) suggests increased use of resources such as the internet and digital library loans to supplement school materials. However, the intensity in the use of these resources is strongly correlated with the parents' educational and economic resources, highlighting the disequalising impact effect of the pandemic on the schooling experience.

Schools and teachers have also had to adjust to the global pandemic, needing to refurbish their materials and courses to the new setting. As Jordan et al. (2021) suggest, this transition faces several challenges, linked to the availability of resources to scale up distance learning and the lack of experience by teachers, and the school system in general, of operating entirely through distanced platforms (Coolican et al., 2020). Here again, school's capacity to transition to online learning is strongly affected by the heterogeneous availability of economic resources across regions and communities. For the US case, Bacher-Hicks et al. (2021) show that schools made more intensive use of digital resources in communities with higher average income than in low-income communities. Likewise, Parolin and Lee (2021) find that exposure to distance learning during 2020 was heavily correlated with a community's income. This means that US schools in low-income communities remained closed for more prolonged periods than those in high-income neighbourhoods. Considering that these communities were also those in which households faced lower accessibility to distance learning technologies, the total effect of this type of pattern is a widening of the educational gaps by income level.

More than a year into the pandemic, it is now possible to assess the initial effects of the displacement from in-person to distance instruction in students' learning and academic progression, as well as in their time-use patterns. Unfortunately, this body of research remains heavily focused on developed countries, as the necessary information is not available for most developing countries. In the case of the Netherlands, Belgium and Brazil, the evidence provided by Engzell et al. (2021); Maldonado and Witte (2021); Lichand et al. (2021) suggests that the suspension of in-person classes had a negative effect on standardised test scores both on mathematics and language components. In the Netherlands, the effect detected was relatively small (0.08 standard deviations), in part due to the short period of class suspension, which was eight weeks (Engzell et al., 2021). In contrast, in Belgium and Brazil, where the suspension of in-person classes was more prolonged, the effect was substantially more significant. In Belgium, Maldonado and Witte (2021) identified that the effect was 0.26 standard deviations, which would be equivalent to the loss of half of a school year. In Brazil, the effect detected by Lichand et al. (2021) was 0.32

standard deviations, or equivalent to the loss of the learnings from two-thirds of a school year. In both cases, the effect was larger among children with less household economic resources.

Besides the direct effects on test score performance, the pandemic has also affected the time dedicated by children and teenagers to study. Grätz and Lipps (2021) found that in Switzerland, teenagers between 14 to 25 years old reduced their study time by 12 hours per week when teaching shifted from in-person to remote learning. The effect observed in Germany is relatively similar, where Grewenig et al. (2021) identified that during the spring of 2020, children and teenagers reduced their learning time by 18 hours per week or 3.8 hours per day. The fall was more significant among students with grades below the median (20 hours per week or 4.1 hours per day) than those above the median (18.5 hours per week or 3.7 per day). Andrew et al. (2020) detect a similar effect for the UK.

The Mexican Education System and the COVID-19 Pandemic

Sanitary authorities detected the first COVID-19 case in Mexico by late February of 2020, and by mid-March of the same year, they detected community transmission of the disease in the country . In response, the federal government declared the suspension of all non-essential activities, including in-person classes in all educational levels, starting on March 23 and until the end of the spring recess on April 17 of 2020 (Diario Oficial de la Federación, 2020a). Due to the evolution of the pandemic in Mexico , authorities opted to keep instruction remote until it was safe for teachers and students to congregate in the classrooms (Diario Oficial de la Federación, 2020c). As a result, the last quarter of the 2019-2020 cycle and the totality of the 2020-2021 educational cycle were conducted remotely.

As in Mexico, the course curricula and general evaluation criteria are set at the federal level, the Federal Ministry of Public Education (Secretaría de Educación Pública) produced a series of materials to act as a guide for teachers throughout the country in terms of the pace they should be following in their classes. Among these instruments, the main one was the TV program "Aprende en Casa II" which covered all courses' syllabi at the primary and secondary levels. The objective of the program was to act as support material for professors in all regions of the country, leaving schools and professors to determine the exact form of the classes (Ramírez-Raymundo et al., 2021). The guiding principle was that schools had a better idea of the resources available in their communities than the federal government. However, the latter's lack of a compensating investment led to significant inequalities in the type of instruction throughout the country. Furthermore, the Ministry of Education changed the evaluation criteria in account of the pandemic, requiring professors to assign a passing grade to all students who remained in contact with the teacher and postponing the evaluation of those who failed to do so (Diario Oficial de la Federación, 2020b, 2021).

The literature that analyses the effects of the displacement to distanced learning on Mexican children and teenagers remains scarce. Boruchowicz et al. (2022) identified that teenagers between 12 to 18 years old reduced their time studying from 40 hours per week before the pandemic to 27 hours per week during the pandemic in 2020. They also identify an increase in the variability of hours dedicated to study, suggesting an increase in the educational inequality in the country. In a different dimension, Cabrera and Padilla-Romo (2020) identify a drop at the national level in the reports of child maltreatment. The reduction is more prominent for females and inhabitants of communities with a high poverty incidence. As the authors point out, this reduction, more than identifying a drop in the incidence of child maltreatment, is a consequence of a reduction in the number of children's interactions with household outsiders, and thus, a lower probability of an adult detecting the signals of abuse. To this, it is necessary to add that a recent analysis by Hillis et al. (2021) estimates that 141,132 children have lost at least one of their caregivers due to the pandemic.

Estimating the immediate learning cost of the pandemic shock

Given the lack of real-time data on the pandemic's impact on learning in Mexico, estimating these costs represents a methodological challenge, even when they represent crucial information for the correct educational policy design. The first part of this challenge is that the total net effect of the pandemic on the learning of the cohort affected by it will only be observable at the end of the academic career of the cohort. Unfortunately, that is several years into the future. Thus, a partial solution is to simulate the pandemic's impact on a relatively young cohort which most of its members have already finished their academic careers. In the Mexican case, this cohort is composed of individuals between 25 and 30 years old.

As Monroy-Gómez-Franco (2021a) distinguishes, any educational dislocation can produce two types of learning shocks. Firstly, an instructional shock can produce immediate losses in learning associated with the direct effects of the shock and the palliative measures taken to deal with them. These losses can be considered the "short-run" costs of the shock. Secondly, the immediate losses can trigger a cumulative process that aggravates them as the student progresses in her academic career without keeping pace with the content progression implied in school programs. Monroy-Gómez-Franco (2021a) calls the gap between the learning stock achieved by the student at a specific point of her school career and the expected learning stock implied in the course curricula the "long-run" cost of the shock.

Estimating the immediate learning cost

Following Monroy-Gómez-Franco (2021a), the first step in analysing the learning effects of the pandemic is to calculate the immediate loss in learning experienced by the cohort that experienced the shock. Following previous work by Neidhöfer et al. (2021), we model this immediate loss as depending on the effects of the shock on school attendance and on the public and private measures designed to attenuate the effects of the shock. Let k_i be the effective immediate loss in learning experienced by child i, measured as the share of a school year's learning that the student did not attain due to the instruction disruption. Then, formally we will have. Then, formally we will have.

$$k_i = \alpha_i \times C_r \tag{1}$$

In which the attenuation capacity of the parents is given by α_i and the gross share of a school year lost due to the in-person instruction disruption are represented by C_r. The parents' attenuation capacity, α_i, refers to each household's capacity to compensate for the effects of the shock on learning through investments of their own. Thus, it depends on the educational and economic resources of the household. In contrast, the gross share of a school year lost, C_r, is defined at the regional level as it depends on the average capacity to engage in remote instruction, the effectiveness of remote instruction compared to in-person learning and the incidence of the shock, in our case, COVID-19.

We model the attenuation capacity of parents following Monroy-Gómez-Franco (2021a), who modifies the approach followed by Neidhöfer et al. (2021) to include the effect of parental economic resources on the investments performed by parents on the education of their children. The empirical literature on parental educational investments identifies that households with more income and more education perform larger investments in their children and compensate the effects of shocks more aggressively than those with lower income or formal education (Heckman and Mosso, 2014; Prix and Erola, 2017).

Monroy-Gómez-Franco (2021a) expresses the attenuation capacity as the weighted average of the household educational and economic resources. In both cases, the variables are measured relative to the rest of the members of the distribution. This implies assuming that members at the top of both distributions can completely compensate for the effects of the share of schooling lost. Reciprocally, it implies that individuals at the bottom of both distributions cannot offset the shock's effects on learning. This is expressed formally in equation 2:

$$\alpha_i = 1 - \left[\theta \frac{e_i}{\max(e)} + (1-\theta)\left[1 - \frac{|\max(w) - w_i|}{|\max(w) - \min(w)|}\right]\right] \tag{2}$$

in which e_i represents the average school years of both parents, $\max(e)$ is the maximum of the average parental school years observed in the data, w_i is the value of the household assets index for the origin household, and $\max(w)$ and $\min(w)$ are the observed maximum and minimum of this variable. θ, with $0 \leq \theta \leq 1$, corresponds to each component's weight in the parents' attenuation capacity. We assume that $\theta = 0.5$ so that the parents' educational attainment and the household's economic resources play an equal role in the parental investments in the child's education.

In the case of the gross share of a school year lost, C_r, its' expected value is defined as a function of the share of a school year during which classes took place remotely, the accessibility to distance learning technologies, and the incidence of the virus. Formally, Neidhöfer et al. (2021) propose the following expression for this expected value:

$$E[C_r] = \frac{d}{D}\left(1 - [(\delta \times \gamma_r) + (\psi \times \kappa_r \times j)]\right) + \frac{\tau_r}{D} \tag{3}$$

in which d represents the number of effective days in which in-person instruction has been suspended, D is the number of days in a typical school year (190). Thus, the elements inside the brackets represent the public policy interventions to attenuate the loss of school days, the accesibility and effectiveness of those interventions, while the second component, τ_r, represents the share of a school year lost due to the incidence of COVID-19.

In the case of the elements inside the brackets, γ_r represents the probability that a household in region r has access to a digital t.v. set, κ_r is the joint probability of a household having a device that allows to connect to the internet (laptop or desktop computers, tablets, smartphones) and having internet access. Access to either one of these technologies would allow the student to follow the transmissions of the program “Aprende en casa II” (Learning from home II), the primary measure taken by the Federal Government to coordinate distanced learning. To this measure of access to remote instruction technologies, we add the probability that a student maintained contact with their teachers at least once per week, represented by j. Both components, γ_r and $\kappa_r \times j$, are weighted by the effectiveness of remote instruction through each one of them in substituting in-person instruction. The weights are represented by δ and ψ.

It is worthwhile discussing the role of the weights at length. If remote instruction were as effective as in-person instruction, the values of each weight would be $\delta = 0.5$ and $\psi = 0.5$. This set of values implies that if access to remote learning were universal, children would not lose a school day from the displacement from in-person to remote learning. This is the first scenario that we consider for our analysis. A second scenario considers $\delta = 0.25$ and $\psi = 0.25$. This assumption implies that a day of distance learning is equivalent to half a day of in-person learning. This assumption implies that distanced learning is only partially equivalent to in-person learning. Finally, a third scenario we consider is $\delta = 0.00$ and and $\psi = 0.00$. This implies that the interventions are completely ineffective, and remote instruction is equivalent to not attending school. The three scenarios encompass the range of possible effects the interventions might have had on students’ learning process.

The parameter representing the loss of school days due to the incidence fo COVID-19, τ, is defined following the literature on the loss of a parent and the economic and educational outcomes of the children who experience them[2]. Neidhöfer et al. (2021) provide the formal definition of this parameter:

$$\tau = \tau^q \times P\,(q = 1) + \tau^d \times P\,(d = 1) \tag{4}$$

in which τ^q represents the costs associated with the death of a household member and τ^d corresponds to the cost associated with the sickness of a member of the household. In both cases, the costs are expressed in terms of days of schooling lost due. We use the same values as Neidhöfer et al. (2021) for both costs. In the case of the cost of a household member being sick in terms of school days lost is $\tau^d = 5$, which corresponds to the average length of COVID-19 symptoms in the case of a mild infection. In the days lost due to the death of a household member, we assume $\tau^q = 15$, following Neidhöfer et al. (2021).

[2]See, among others Corak (2001); Gertler et al. (2004); Amato and Anthony (2014); Prix and Erola (2017); Steele et al. (2009); Cas et al. (2014)

Modelling the cumulative learning costs

The second step proposed by Monroy-Gómez-Franco (2021a) is to embed the share of a school year learning lost, (k_i), in a simulated academic trajectory in order to calculate the cumulative effects of the educational disruption. Consequently, this requires a model of how learning accumulates through a student's progression through the school system. The most common approach to this is through the estimation of an "Educational Production Function" (EPF), in which learning is conceived as the product of a production process that requires multiple inputs, including family background, peers, school-related factors and the innate abilities of the child.

Among the family of EPFs, the Potential Pedagogical Function (PPF) proposed by Kaffenberger and Pritchett (2020b, 2021) seeks to capture the dynamics of school progression, and learning accumulation and how they can interplay with each other and progress in one dimension (schooling) might not translate into the other (learning) . A series of recent empirical findings on how increases in the years of schooling are not translating into a more extensive learning stock in students motivates this insight (for Indonesia (Beatty et al., 2021), for Rwanda (Crawfurd, 2021), for multiple African countries (Pritchett and Sandefur, 2020; Kaffenberger and Pritchett, 2020a)and Pakistan (Bau et al., 2021)). This would explain why transitory shocks to education have persistent effects in the scholastic careers of those who suffered them, as a growing body of literature identifies (see, among others, Andrabi et al. (2021); Belot and Webbink (2010); Ichino and Winter-Ebmer (2004); Jaume and Willén (2019); Marcotte and Hemelt (2008); Meng and Zhao (2021); Sacerdote (2012)).

The abovementioned literature implies that not all children in the educational system are effectively receiving the learning treatment implied by the instructional process. Given that some, but not all, students report null gains from progressing through school, this means that for the treatment to be effective (namely, the instruction experience corresponding to one school grade), the students need to fulfil two conditions. One is that students require a certain amount of background knowledge to make sense of the curriculum contents of a specific grade. The lack of this knowledge makes it impossible for the students to acquire the expected learning for the grade in question. The second condition is that the student's knowledge should not be above the coverage of the course contents. If so, then the student will not learn anything new. Both conditions define a region of previous knowledge over which the curriculum of a school grade is designed to be effective. Movements outside that region reduce the gains in terms of learning obtained from taking that school grade.

A formal representation of this process is the PPF proposed by Kaffenberger and Pritchett

(2020b, 2021). This function describes the average learning that child i with learning stock s_i would obtain if she attended grade G. This a function of the learning profile of the grade, LP^G and the skill of the student. Formally, they define the PPF as follows:

$$PPF(LP(w,z,v,\pi^G),s_i) = \begin{cases} 0 & \text{if } s_i < \pi^G - \frac{w}{2} \\ z_{min} + v(s_i - (\pi^G - \frac{w}{2})) & \text{if } \pi^G - \frac{w}{2} < s_i < \pi^G + \frac{w}{2} \\ 0 & \text{if } s_i > \pi^G + \frac{w}{2} \end{cases} \quad (5)$$

In which z_{min} is the minimum level of learning attained if the instructional process G is effective, w is the range of G in terms of the initial skills of the students to which the course is directed and π^G is the skill level for which the course is centred. v represents the focus of the instructional process. If $v > 0$, then G is biassed in favour of the students with a higher initial learning stock; they will learn more than those with a lower initial stock. If $v < 0$, then students with a lower stock of learning will learn more. And if v=0 everyone would learn the same, z_{min}. Formally, v is defined as

$$v = \frac{z_{max} - z_{min}}{w} \quad (6)$$

In which z_{max} is the maximum learning expected in grade G.

The first condition of the trapezoidal PPF, 0 if $s_i < \pi^G - \frac{w}{2}$ implies that students taught a material that is too advanced would gain nothing from the instructional process G. In the same way, students taught material that is well below their current proficiency would not gain anything from attending the course (hence, the third case in the PPF, $s_i > \pi^G + \frac{w}{2}$. Only students that have the initial skills on the range for which the instructional process G was designed, $\pi^G - \frac{w}{2} < s_i < \pi^G + \frac{w}{2}$, will gain from attending classes.

We assume that $z_{min} \neq 0$ for students with initial skills that fulfil this last condition, such treatment is inherently effective provided the students are in the region of treatment. In plain words, this implies assuming the instructional experience is not entirely ineffective per se. Given the existing evidence, we consider this to be a realistic assumption.

Equation 5 represents the effects on learning of one year of instruction. To simulate a trajectory, it is necessary also to define the pace ρ at which the scope of the PPF shifts with every year of progression. The pace ρ refers to the change in the skill level on which the instruction is centred. As Kaffenberger and Pritchett (2020b, 2021) show, this implies

$$(\pi^G + \rho) - \frac{w}{2} = \pi^{G+1} - \frac{w}{2} \tag{7}$$

$$(\pi^G + \rho) + \frac{w}{2} = \pi^{G+1} + \frac{w}{2} \tag{8}$$

The last parameter in the model is the dropout rate. We follow Kaffenberger and Pritchett (2020b, 2021) and model the dropout rate as a function of students' skills, where the bottom of the skill distribution drops out of the instructional process. Formally, the dropout function, $q(s_i)$, is defined as follows

$$q(s_i) = \begin{cases} 1 & \text{if } s_i \leq \phi^G \\ 0 & \text{if } s_i > \phi^G \end{cases} \tag{9}$$

In which ϕ^G is the grade-specific cut-off value of the skill distribution below which students drop out. We select ϕ^G so that the model replicates the dropout rates observed in the Mexican education system before the pandemic .

Data

Our approach to estimating the short and long-run costs of the pandemic in learning requires us to use multiple data sources, as we require information on household economic and educational resources and the epidemiological conditions of the country. In particular, we need information on the parental resources available to each household to attenuate the pandemic shock. Furthermore, for the calibration of the Potential Pedagogical Function, we require information on learning progression, operationalised through a standardised test. Finally, to capture as much regional variation as possible, we would also require data sources that are representative at the lowest possible level of disaggregation.

With regards to the information on the economic and educational resources of the household, our primary data source is the *Espinosa Rugarcia Social Mobility in Mexico Survey 2017* (ESRU-

EMOVI 2017). The ESRU-EMOVI 2017 is a survey designed explicitly for the study of social mobility in Mexico. It contains ample information on the economic conditions of the household inhabited by the respondent when she was 14 years old and information on the educational attainment of both parents. The information in the survey is representative at the national and regional level of the Mexican population between 25 and 64 years old[3]. However, our interest is on the youngest cohort (25 to 30 years old) as the conditions of their households of origin are closest to the characteristics of the current Mexican households. This implies restricting our sample from 17,665 to 2,474 observations. Although their conditions when 14 years old are not necessarily equivalent to those of the current cohort, they represent the youngest available cohort in any survey that includes conditions of origin without restricting to only co-residents. Thus, it is a sample that allows us to attenuate any concern linked to co-residence bias. Table 1 shows the sample's composition in terms of sex, indigenous status, urban residence, and the years of school attainment of the respondent and both parents.

Table 1: Descriptive Statistics of ESRU-EMOVI 2017

Variable	National	North	North West	Center North	Center	South
Years of school of interviewee	11.8453	11.4828	12.2753	11.6716	12.5643	10.9050
(Regional mean)	(0.1291)	(0.2009)	(0.2578)	(0.2989)	(0.2396)	(0.3074)
Years of school of the father	7.5955	8.1953	6.7974	7.4127	8.8788	5.4875
(Regional mean)	(0.2115)	(0.2568)	(0.3377)	(0.3669)	(0.4235)	(0.3393)
Years of school of the mother	7.2009	8.1953	7.1173	7.2323	8.2746	5.0322
(Regional mean)	(0.1896)	(0.2473)	(0.3714)	(0.3350)	(0.4236)	(0.3308)
Female population	0.5219	0.5307	0.5139	0.5217	0.5160	0.5284
(Share of regional population)	(0.0141)	(0.0242)	(0.0321)	(0.0352)	(0.0273)	(0.0264)
Urban community of origin	0.7297	0.8955	0.6048	0.7135	0.8374	0.4953
(Share of regional population	(0.0327)	(0.0244)	(0.0566)	(0.0449)	(0.0298)	(0.0411)
Indigenous population	0.1085	0.0525	0.0254	0.0641	0.0656	0.2618
(Share of regional population)	(0.0123)	(0.0161)	(0.0146)	(0.0169)	(0.0149)	(0.0326)
Regional population		0.1588	0.0673	0.1310	0.4137	0.2292
(Share of national population)		(0.0184)	(0.0098)	(0.0163)	(0.0361)	(0.0213)

Notes: Data from ESRU-EMOVI 2017 for respondents between 25 and 30 years old. Standard errors in parenthesis.

We employ a household asset index as a summary measure of the economic resources available to the household of origin. Household asset indices have been employed in the development literature for analysis regarding the distribution of economic resources when other variables such as income are not available (see Filmer and Pritchett (2001); McKenzie (2003); Wittenberg and Leibbrandt (2017); Poirier et al. (2020)). To construct the index, we employ Multiple Correspondence Analysis (MCA) as the variables in ESRU-EMOVI 2017 only record ownership of the assets, thus producing binary variables instead of continuous ones. Briefly, MCA uses

[3]The North region consists of Baja California, Sonora, Chihuahua, Coahuila, Nuevo León and Tamaulipas; North West consists of Baja California Sur, Sinaloa, Nayarit, Durango and Zacatecas; the Center North region is form by Jalisco, Aguascalientes, Colima, Michoacán and San Luis Potosí; the Center region is formed by Guanajuato, Querétaro, Hidalgo, Estado de México, ; Mexico City, Morelos, Tlaxcala, and Puebla; the South region is formed by Guerrero, Oaxaca, Chiapas, Veracruz, Tabasco, Campeche, Yucatán and Quintana Roo

relative frequencies across the binary variables considered to identify structure in terms of ownership, which can be used to rank individuals according to the availability of economic resources. We use the parents' average years of education and the index value to calculate both components of the household attenuation measure described in equation 2.

Table 2 shows the assets that we employ to construct the asset index of the parental household, which will be used to calculate the capacity of attenuation of each household.

Table 2: Binary variables for the parental household asset index.

The household has access to the water supply	The household has a washing machine
The household has an oven	The household has a landline telephone
The household has a television	The household has a computer
The household has a refrigerator	The household has a VHS
The household has a microwave	The household has cable television
The household owns a water heater	The household owns a vacuum cleaner
A member of the household owned the housing facilities inhabited	A member of the household owns a car
A member of the household has a bank account.	A member of the household owns a credit card.
The household hires a domestic worker.	

In order to calculate the short-run costs of the pandemic, we need information about the availability of televisions, internet connections and digital devices at the regional level to calculate γ and κ in equation 3. Therefore, to have the most up to date information, we employ the data from the 2020 National Population Census to calculate the rates of access to the internet, digital devices, and televisions in each of the regions for which ESRU-EMOVI 2017 is representative. The second component of the short-run cost requires us to calculate both the probability of infection and the probability of death by COVID-19 in a given region (equation 4). For this calculation, we use the data in the website deployed by the National Council for Science and Technology (Consejo Nacional de Ciencia y Tecnología, CONACyT) to record the incidence of the pandemic across the Mexican territory. We present the values of these variables in appendix A.

In the following section, we describe the sources employed in the calibration of the Potential Pedagogical Function and the values obtained for the parameters.

Calibration of the PPF

We calibrate the model described by equations 5-9 to replicate the mean and standard deviation of the math component of the Mexican standardised knowledge test for students of the third

Table 3: Calibrated parameters for Mexican regions

	North	North West	Center North	Center	South
Average math score (Total sample)	492	494	513	511	483
Standard deviation of math score (Total sample)	124	113	115	117	107
w	153	153	153	153	153
$z_{\min}$	37.7	37.7	38.4	38.4	36.5
$z_{\max}$	71.63	71.63	72.96	72.96	69.35
ρ	54	54	54	54	54
v	0.2218	0.2218	0.2259	0.2259	0.2147
Initial distribution	N(0,20)	N(0,20)	N(0,20)	N(0,20)	N(0,20)

grade of secondary school (equivalent to ninth grade). The exam, part of the National Plan for the Evaluation of Learnings, was a governmental effort to evaluate students' performance in their learning careers from initial education up to the end of high school. We employ the results of the latest available version of the test, which corresponds to 2017[4].

As we are interested in capturing the regional variance in the effects of the COVID-19 shock, we need to calibrate a model for each one of the regions. However, it is important to note that the educational curriculum goals for each grade of education are set at the national level, not at the regional one. For this reason, we first calibrate a model to the national parameters, of which we obtain the pacing parameter to be used and held constant in the regional models. As a result, our models imply that learning differs across regions due to differences in the minimum and maximum learning goals obtained at each grade, not because of differences in curriculum pacing.

In table 3, we show the parameter values for each of the regional models, while in Appendix C, we present the results for the national model and a detailed description of the calibration process. We follow the same criteria as Kaffenberger and Pritchett (2020b) for the determination of z_{min}, z_{max} and in the selection of the initial distribution.

To provide an estimation of the long-run or accumulated cost of the pandemic for each observation in our sample, we use the individual immediate or short-run effective loss of school days due to the pandemic, k_i, to introduce a shock into the learning gain corresponding to the sixth and seventh year of individual i educational trajectory. Then, the average accumulated

[4]In Spanish this is the Plan Nacional para la Evaluación de Aprendizajes, PLANEA. The tests were designed and processed by the National Institute for the Evaluation of Education (INEE), which disappeared in 2019. This caused that subsecuent evaluations were designed by the Ministry of Education and the results no longer were standardized to be comparable across schools and states. For this reason, we employ the results from the 2017 test and not those corresponding to the 2019 evaluation.

learning by the end of the ninth grade under this simulated educational trajectory is assigned to individual i, corresponding to her long-run accumulated learning under the pandemic scenario. This method allows us to capture in a better way the heterogeneity of impacts caused by the pandemic than only considering a homogenous shock to all the observations in our sample.

Our main results are based on the case of a student that was in sixth grade when the pandemic hit. We focus on this case as the progression between sixth and seventh grade in the Mexican educational system implies the transition between *primaria* (elementary school) and *secundaria* (middle school). This means that these students had to transition between school levels during the pandemic, making them particularly vulnerable to being affected by the displacement to remote learning due to the pandemic.

To translate the model results from scores in the PLANEA test for ninth-grade students to school years of learning, we calculate the gap, Δ, between the simulated accumulated learning under the pandemic scenario with that of the no-pandemic scenario. That is, those used as a baseline for the calibration of the model. After calculating this gap in scores, we divide them by the grade progression value calibrated for the national model, $\rho = 54$. This allows us to express the gap in learning stocks in terms of years of learning progression. In other words, in terms of school years of learning. This is the long-run educational cost of the pandemic (LCRP) in terms of school years of learning lost. The value of the LCRP represents the years of learning progression that a specific student is behind with respect to those she would have reach under normal circumstances.

$$LCRP = \frac{\Delta}{54} \tag{10}$$

It is important to emphasise that the models only can simulate the learning progression under the assumption that contextual factors such as other types of interactions between students and professors remained constant. Although the pandemic makes this assumption unrealistic, the lack of information on the developmental effects of the interruption of these in-person interactions forces us to maintain it. For the same reason, we retain the dropout rates constant at the levels observed before the pandemic. Although these are likely to change, we do not have a reliable estimate of the dropout rates by school grade during the pandemic. Thus, our results should only be interpreted as a lower bound of the pandemic effects on learning and schooling for both reasons.

Results

Immediate learning costs

The sources and values for each parameter are specified in Appendix A. Table 4 details the expected value of the short-run cost for each of the regions under analysis.

Table 4: Average effective immediate cost in each scenario.

Region	Average effective immediate cost (Scenario 1)	Average effective immediate cost (Scenario 2)	Average effective immediate cost (Scenario 3)
National	0.3239 (0.0008)	0.6836 (0.0011)	1.0432 (0.0015)
North	0.2411 (0.0009)	0.6218 (0.0022)	1.0026 (0.0036)
North West	0.3234 (0.0012)	0.6944 (0.0025)	1.0653 (0.0037)
Center North	0.2908 (0.0097)	0.6696 (0.0023)	1.0484 (0.0035)
Center	0.2685 (0.0008)	0.6244 (0.0017)	0.9802 (0.0031)
South	0.4730 (0.0010)	0.8041 (0.0019)	1.1351 (0.0026)

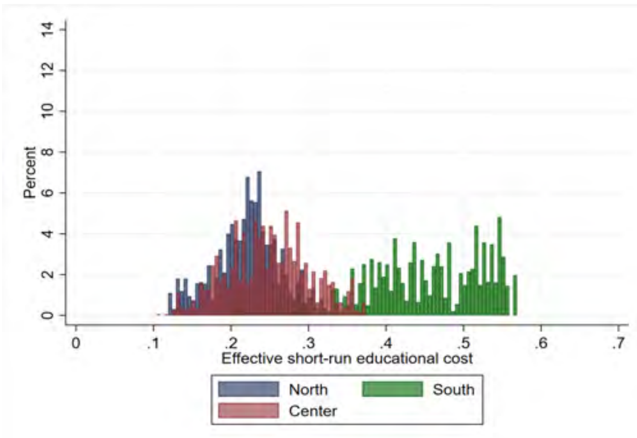
Note: Authors' calculations corresponding to equation 1. The estimated values of the gross educational cost for each region are presented in appendix A.

As shown in Table 4, the gap between the scenarios modelled is 0.70 of a school year of learning. In the best-case scenario, where remote instruction was a perfect substitute for in-person instruction, the transition to distance learning represented a loss of a third of a school year in terms of learning at the national level and almost half of the school year learning in the south. In the results corresponding to the worst-case scenario, the average loss at the national level is of an entire school year of learning and more than one school year of learning in the south of the country. Finally, we add the effective costs for the 2019-2020 and 2020-2021 school years to calculate the total cost expressed as school years of learning progression lost due to the direct impact of the pandemic.

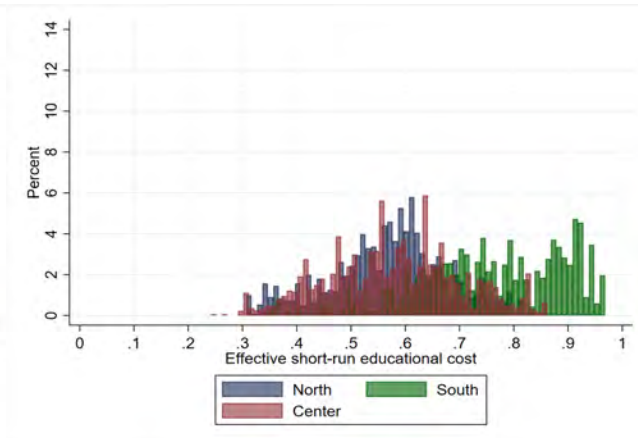
In all scenarios, the region most affected by the transition from in-person to remote learning is the South of the country. In contrast, the less affected region is the North. This consistent pattern across simulations highlights the close relationship between the private attenuation capabilities of the households (dependent on school years and economic resources) and the availability of digital devices and computers that allow the public interventions to be effective.

Although the mean effect represents the differences between the country's regions in terms of the short-run educational cost experienced, it fails to capture the within-regional variability of the shock. This within region variability is generated by the inequality in educational attainment and economic resources inside each region. Figure 1 shows the distribution of the effective short-run costs inside the north, centre, and south of the country under the three scenarios analysed. The distributions for the five regions under the three scenarios are in Appendix B.

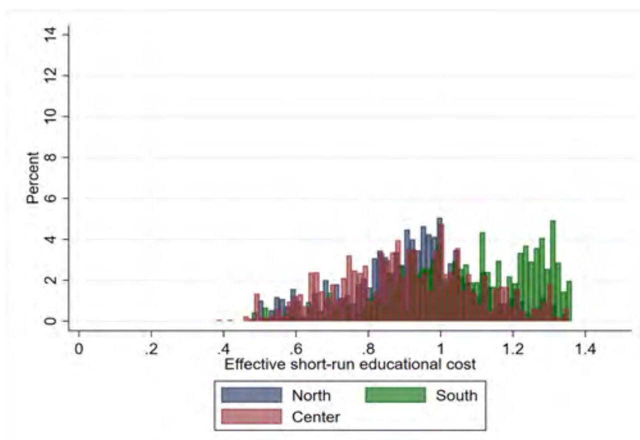
Figure 1. Distribution of the short-run educational costs of the pandemic across three different regions
(Fraction of a school year of learning lost)



(a) Scenario 1



(b) Scenario 2



(c) Scenario 3

Note: Authors calculations. The effective immediate cost corresponds to the share of a school year of learning lost due to the transition to remote learning. Scenario 1 corresponds to the assumption of $\delta = 0.5$ and $\psi = 0.5$; Scenario 2 to $\delta = 0.25$ and $\psi = 0.25$ and Scenario 3 to $\delta = 0$ and and $\psi = 0$

The first element highlighted by figure 1 is that depending on the assumption made about the effectiveness of remote instruction, the average cost and the within region distribution of the cost changes. Figure 1.a, in which we assume that the public interventions made distance learning equivalent to in-person learning, shows a relatively low dispersion of the immediate learning costs in the Center and North compared to that experienced in the South. In contrast, figure 1.c shows that the cost dispersion across the three regions has the same range of values under the assumption of entirely ineffective educational interventions. However, the south still exhibits a more significant concentration of the population in the right tail of the learning cost distribution.

The impact of an effective public intervention to attenuate the learning cost of the pandemic is not trivial and is highlighted by two factors. The first one is the difference in the range of values of the short-run educational cost observed between scenarios one and three. In the scenario with effective attenuation, the range for the short-run costs in the Centre region lies between 0.1 and 0.35 of a school year. In contrast, under the assumption that the interventions were ineffective, the range is between 0.37 and 1.37 school years. This third scenario corresponds to one in which attenuation is performed entirely by the household and the educational and economic resources available to its members. From the literature on the inequality of opportunity in Mexico (Vélez-Grajales et al., 2018; Monroy-Gómez-Franco et al., 2021; Plassot et al., 2021), we know that educational and economic resources are very unevenly distributed across Mexican households. This results in significant variability in the effective immediate learning costs when only these resources are available to attenuate the shock.

The second factor that highlights the importance of effective public interventions is the difference in mean and range between the Center and the South of the country under scenario one. Due to differences in availability of internet access and digital devices between both regions, even when we assume complete effectiveness of policy interventions, these measures affect a smaller share of the population in the South than in the rest of the country. As a result, the distribution of short-run or immediate costs in the South is wider than in the rest of the country, as private resources play a larger role in attenuating the learning costs.

This brings to the forefront our first major result: in the absence of a targeted policy and without considering the long-run educational costs of the pandemic, it is likely that the regional gap in learning and educational attainment between the South and the rest of Mexico will increase as a consequence of COVID-19. This pattern implies a reversal of the process of regional convergence in education during the XXth and the early XXIsth centuries (de la Torre and Vélez-Grajales, 2016).

In the absence of any compounding effects, the learning costs identified in this section would represent the associated learning costs of the pandemic. This is the approach taken by Neidhöfer et al. (2021) to estimate the effects of the pandemic in educational attainment and intergenerational mobility for a large set of Latin American countries. However, the literature on learning profiles and academic progression suggests that even temporary shocks can be amplified by an educational system that fails to adjust to the students' learning profiles. In the following section, we take this insight into consideration and model how the short-term costs of the pandemic would play out under the learning progression implied by the Mexican curricula.

Long-run or cumulative learning cost

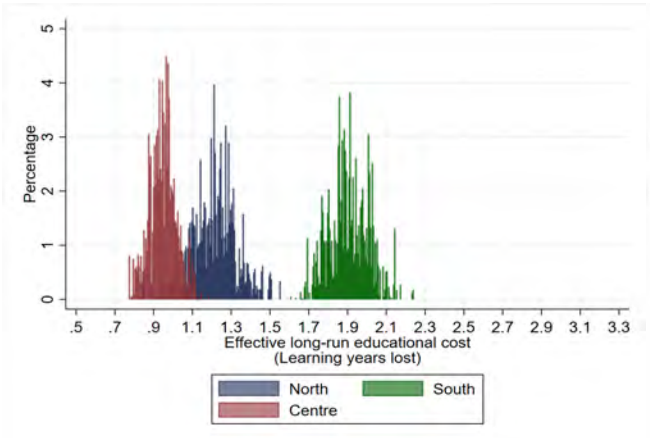
Figure 2 shows the distribution of the long-run or cumulative learning costs in the North, Center and South regions of the country under the three scenarios of the effectiveness of the distance learning interventions. A first element to note is that even under the best-case scenario, the compounding effect makes a short-run cost of 0.2 school years translate into a long-run cost of almost an entire school year of learning lost in the case of the Center region. This highlights the importance of considering how transitory shocks to the academic career can have more significant effects over the long run as the gap produced by the shock widens through time.

It is also important to note that, when comparing across scenarios, the compounding effect differs from region to region. For example, whereas under scenario 1 in the Center region, the compounding effect increases the short-run cost by a factor of 3.5, the equivalent number in the South is 4.6. Under scenario 3, the corresponding factors are 5.9 for the Center and 7.3 for the South. It is also important to note that the widening of the distribution of short-run costs observed when comparing the distributions from scenarios one, two, and three is not observable in the long run cost for the South and Centre region. However, it remains observable in the case of the North.

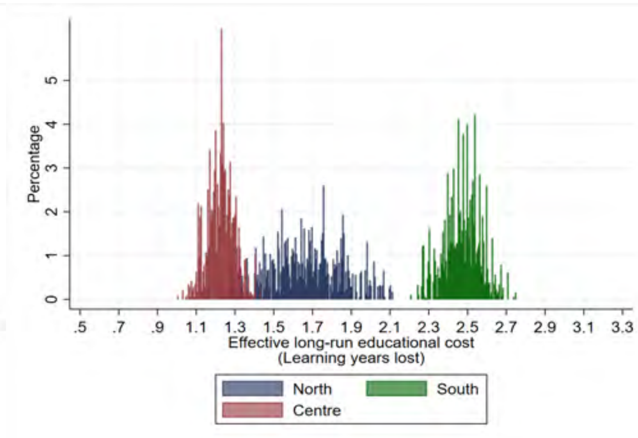
Table 5 shows the average long-run learning cost of the pandemic in terms of the number of years of learning progression a student is behind the expected learning stock at ninth grade for all regions and at the national level. In all three scenarios, the region where we expect the highest cost is in the South and the lowest cost in the Center. This is related to higher dropout rates in the South and the lower attenuation capacity of the households in the region. Moving from scenario 1 to scenario III implies assuming different grades of effectiveness of the public sector attenuation measures. As they diminish, the role of the family attenuation capacities increases, explaining the large (small) increase in the long-run costs in the case of the South (Center) when moving across the three scenarios.

In table 5 we show the average ratio between the long-run and the short-run costs at the national level and for each region. This figure indicates the regional differences in the compounding effect of the short-run costs due to the differences between each region's learning process.

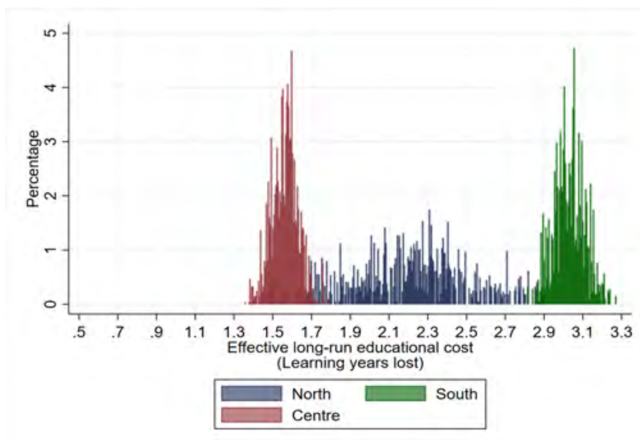
Figure 2. Distribution of cumulative learning costs of the pandemic across three different regions (years of learning progression a student is behind the expected learning stock at 9th grade.)



(a) Scenario 1



(b) Scenario 2



(c) Scenario 3

Note: Authors calculations.The effective long-run cost corresponds to the number of years of learning progression a student is behind the expected learning stock at ninth grade. Scenario 1 corresponds to the assumption of $\delta = 0.5$ and $\psi = 0.5$; Scenario 2 to $\delta = 0.25$ and $\psi = 0.25$ and Scenario 3 to $\delta = 0$ and and $\psi = 0$

In the three scenarios, the same pattern occurs, where the compounding effect is largest in the South and smallest in the Center of the country. Together, they imply an increase in the inequality of learning across Mexico in the absence of any compensating measure. In particular, they suggest an increase in the gap between the South and the rest of the country, as not only has the region suffered the most significant shock in the short-run but also, given its characteristics, is the region where this type of shocks have more significant consequences down the road.

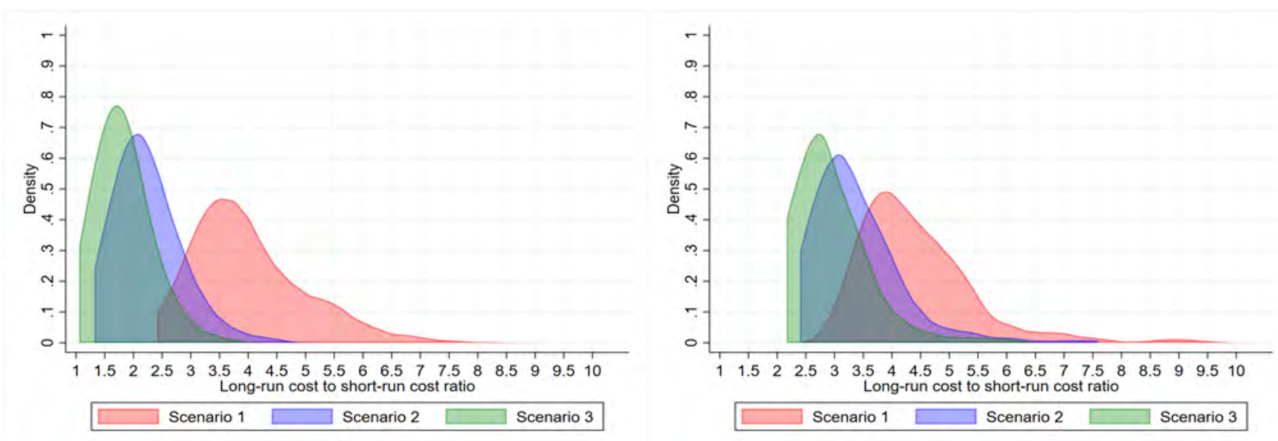
In figure 3, we show that the difference in the long to short-run cost ratios is not only

Table 5: Average long-run cost and long-run to short-run ratio

Region	Average long-run cost			Average long-run to short-run ratio		
	Scenario I	Scenario II	Scenario III	Scenario I	Scenario II	Scenario III
National	1.2908 (0.0224)	1.6813 (0.0294)	2.1409 (0.0369)	4.6215 (0.0626)	2.7578 (0.0392)	2.2962 (0.0354)
North	1.2155 (0.0064)	1.6644 (0.0122)	2.2043 (0.0164)	5.6993 (0.0674)	2.9866 (0.0209)	2.4356 (0.0112)
North West	1.3780 (0.0133)	1.6113 (0.0132)	2.1480 (0.0258)	4.7587 (0.0464)	2.5887 (0.0272)	2.2294 (0.0119)
Center North	1.2828 (0.0069)	1.7069 (0.0138)	2.2352 (0.0253)	5.0914 (0.0731)	2.9027 (0.0332)	2.4072 (0.0190)
Center	0.9473 (0.0039)	1.2269 (0.0037)	1.5628 (0.0039)	4.0077 (0.0746)	2.2367 (0.0428)	1.8135 (0.0353)
South	1.9065 (0.0059)	2.4726 (0.0053)	3.0311 (0.0052)	4.4333 (0.0746)	3.3859 (0.0558)	2.9371 (0.0477)

Notes: Authors' calculations. Standard errors in parenthesis

Figure 3. Distribution of the compounding effect in the centre and the south.
(Ratio between cumulative and immediate learning costs)



(a) Center

(b) South

Note: Authors calculations.The compounding effect is defined as the ratio between the cumulative years of learning progression lost and the share of a school year lost due to the displacement to remote learning. Scenario 1 corresponds to the assumption of $\delta = 0.5$ and $\psi = 0.5$; Scenario 2 to $\delta = 0.25$ and $\psi = 0.25$ and Scenario 3 to $\delta = 0$ and and $\psi = 0$

observable at the mean but is present for the whole distribution when comparing the two extreme regions: Center and South. In all three scenarios, most of the distribution of ratios of the Center is to the left of the South's distribution. If these differences translate into differences in income in the labour market, we would expect an increase in income inequality due to the long-run effects of the pandemic on learning.

When comparing the three scenarios, the decrease in the compounding effect observed in table 5 and figure 3 results from the increase in the magnitude of the short-run cost and the implicit lower bound in learning present in the PPF function. The latter refers to the fact that, even when a student stopped learning due to falling outside the range of the educational progress, the PPF assumes that the stock of knowledge attained does not depreciates once all new learning stops occurring. This produces an upward bias in our estimates of the long-run costs and, consequently, the compounding effect. Thus, our results must necessarily be interpreted as a lower bound of the potential long-run costs of the pandemic in terms of school learning.

Final Remarks

The analysis presented in this paper contributes to our understanding of the effects of the Covid-19 pandemic on human capital accumulation in two ways. First, the methodology used allows for an estimation of the regional differences in impact, measured as the reduction in learning caused by the closure of schools in Mexico. Second, given the type of data used, the analysis also looks into regional and family-context conditions to estimate the cumulative effect of school closure, which translates into the long-term cost of school learning. Considering different degrees of effectiveness of the distance education model established by the government, the results show marked regional heterogeneity and a permanent cost in learning that is significantly higher than that estimated for the short term.

The low effectiveness of the distance educational model may result in a permanent cost, measured in terms of learning, of at least three years of schooling for the southern region shows the magnitude of the problem. Furthermore, it is important to stress that such estimate of the permanent shock is an average effect. The observed dispersion places the young members of households with fewer resources in a worse situation. In the absence of compensatory measures, we face a situation that might translate into a diminishing set of opportunities with a cost in terms of social mobility for the young Mexican cohorts affected by the closures.

To give a simple idea of what a three-year learning cost entails, we can use the average labour income by educational level before the shock of the pandemic as a reference. With data from the ENOE (Employment National Survey) for the first quarter of 2020 (in 2021 pesos), the difference in average labour income between workers with complete elementary and middle levels in the southern region was 656 Mexican pesos per month (5112 versus 5767 pesos, respectively). Assuming that the labour market can distinguish the difference in learnings without changes in the formal years of schooling, the lifelong cost would translate precisely into the 656 pesos mentioned above, a drop of over 11 per cent of the monthly labour income the population that has completed middle school. In addition, we should mention that this reduction may be more significant for people who have school diplomas from school systems that signal negatively in the labour market, with which the income gap can be more significant in a context of a lower average income level.

The results of this exercise constitute in themselves an urgent appeal to the Mexican government. When designing a return-to-school strategy, it is necessary to consider that the short-term costs will translate into a higher cost if there is no significant state effort to reduce the resulting learning gap. This effort necessarily has to act in three dimensions. The first dimension has to be the provision of infrastructure that makes the return to in-person classes as safe as possible and generates trust in both parents and children. A second set of policies must focus on the training of teachers to accelerate learning and provide remedial education for those students left behind in terms of learning milestones. And thirdly, it is urgent to invest in capacity, both in terms of infrastructure and teachers, to deal with this new set of demands. Until now, the Mexican education system has been forced to adjust without increasing resources, which is not sustainable in the long run. Our results indicate that this multidimensional set of policies must consider at least two inequalities: between regions and between households. Otherwise, the official school credentials will hide the actual learnings loss, but the same will not occur when the affected children enter the labour market. Even if it costs fiscal resources, we must avoid that the effect on productivity and earnings and the barriers to social mobility grow as a permanent effect of the Covid 19 crisis.

References

Amato, Paul R. and Christopher J. Anthony, "Estimating the Effects of Parental Divorce and Death With Fixed Effects Models," *Journal of Marriage and Family*, 2014, *76* (2), 370–386.

Andrabi, Tahir, Benjamin Daniels, and Jishnu Das, "Human Capital Accumulation and Disasters: Evidence from the Pakistan Earthquake of 2005," *Journal of Human Resources*, 2021.

Andrew, Alison, Sarah Cattan, Monica Costa Dias, Christine Faquharson, Lucy Kraftman, Sonya Krutikova, Angus Phimister, and Almudena Sevilla, "The gendered division of paid and domestic work under lockdown," *Covid Economics, Vetted and Real Time Papers*, 2020, *1* (39), 109–139.

Bacher-Hicks, Andrew, Joshua Goodman, and Christine Mulhern, "Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real time," *Journal of Public Economics*, 2021, *193*, 104345.

Bansak, Cynthia and Martha Starr, "Covid-19 shocks to education supply: how 200,000 U.S. households dealt with the sudden shift to distance learning," *Review of Economics of the Household*, 2021, *19* (1), 63–90.

Bau, Natalie, Jishnu Das, and Andres Yi Chang, "New evidence on learning trajectories in a low-income setting," *International Journal of Educational Development*, 2021, *84*, 102430.

Beatty, Amanda, Emilie Berkhout, Luhur Bima, Menno Pradhan, and Daniel Suryadarma, "Schooling progress, learning reversal: Indonesia's learning profiles between 2000 and 2014," *International Journal of Educational Development*, 2021, *85*, 102436.

Belot, Michèle and Dinand Webbink, "Do Teacher Strikes Harm Educational Attainment of Students?," *Labour*, 2010, *24* (4), 391–406.

Boruchowicz, Cynthia, Susan W. Parker, and Lindsay Robbins, "Time use of youth during a pandemic: Evidence from Mexico," *World Development*, 2022, *149*, 105687.

Cabrera, Francisco and María Padilla-Romo, "Hidden Violence: How COVID-19 School Closures Reduced the Reporting of Children Maltreatment," *Latin American Economic Review*, 2020, *29*, 105687.

Cas, Ava, Elizabeth Frankenberg, Wayan Suriastini, and Duncan Thomas, "The Impact of Parental Death on Child Well-being: Evidence From the Indian Ocean Tsunami," *Demography*, April 2014, *51* (2), 437–457.

Coolican, Mariana, Juan Carlos Borras, and Michael Strong, "Argentina and the COVID-19: Lessons learned from education and technical colleges in Buenos Aires Province," *Journal of Education for Teaching*, 2020, *46* (4), 484–496.

Corak, Miles, "Death and Divorce: The Long-Term Consequences of Parental Loss on Adolescents," *Journal of Labor Economics*, 2001, *19* (3), 682–715.

Crawfurd, Lee, "Accounting for repetition and dropout in contemporaneous cross-section learning profiles: Evidence from Rwanda," *International Journal of Educational Development*, 2021, *85*, 102443.

de la Torre, Rodolfo and Roberto Vélez-Grajales, *Informe sobre Desarrollo Humano. México 2016: Desigualdad y Movilidad*, United Nations Development Program, 2016.

Delajara, Marcelo, Raymundo M. Campos-Vazquez, and Roberto Velez-Grajales, "The regional geography of social mobility in Mexico," *Regional Studies*, 2021, pp. 1–14.

Diario Oficial de la Federación, "Acuerdo 02/03/20 por el que se suspenden las clases en las escuelas de educación preescolar, primaria, secundaria, normal y demás para la formación de maestros de educación básica del Sistema Educativo Nacional, así como aquellas de los tipos medio superior y superior dependientes de la Secretaría de Educación Pública," 03 2020. `https://www.dof.gob.mx/nota_detalle.php?codigo=5589479&fecha=16/03/2020`.

_ , "Acuerdo 26/12/20 por el que se establecen las orientaciones pedagógicas y los criterios para la evaluación del aprendizaje para la educación preescolar, primaria y secundaria en el periodo de contingencia sanitaria generada por el virus SARS-CoV2 (COVID-19) para el ciclo escolar 2020-2021.," 12 2020. `https://www.dof.gob.mx/nota_detalle.php?codigo=5608934&fecha=28/12/2020`.

_ , "Acuerdo por el que se establecen los Lineamientos Técnicos Específicos para la Reapertura de las Actividades Económicas.," 05 2020. `https://www.dof.gob.mx/nota_detalle.php?codigo=5594138&fecha=29/05/2020`.

_ , "Acuerdo 16/06/21 por el que se regulan las acciones específicas y extraordinarias relativas a la conclusión del ciclo escolar 2020-2021, en beneficio de los educandos de preescolar, primaria y secundaria ante el periodo de contingencia sanitaria generada por el virus SARS-CoV2 (COVID-19)," 06 2021. `https://www.dof.gob.mx/nota_detalle.php?codigo=5621985&fecha=22/06/2021`.

Dietrich, Hans, Alexander Patzina, and Adrian Lerche, "Social inequality in the homeschooling efforts of German high school students during a school closing period," *European Societies*, 2021, *23* (sup1), S348–S369.

Engzell, Per, Arun Frey, and Mark D. Verhagen, "Learning loss due to school closures during the COVID-19 pandemic," *Proceedings of the National Academy of Sciences*, 2021, *118* (17).

Filmer, Deon and Lant H. Pritchett, "Estimating Wealth Effects without Expenditure Data-or Tears: An Application to Educational Enrollments in States of India," *Demography*, 2001, *38* (1), 115–132.

Francis, Dania V. and Christian E. Weller, "Economic Inequality, the Digital Divide, and Remote Learning During COVID-19," *The Review of Black Political Economy*, 2021.

Gertler, Paul, David I. Levine, and Minnie Ames, "Schooling and Parental Death," *The Review of Economics and Statistics*, 02 2004, *86* (1), 211–225.

Grewenig, Elisabeth, Philipp Lergetporer, Katharina Werner, Ludger Woessmann, and Larissa Zierow, "COVID-19 and Educational Inequality: How School Closures Affect Low- and High-Achieving Students," *European Economic Review*, 2021.

Grätz, Michael and Oliver Lipps, "Large loss in studying time during the closure of schools in Switzerland in 2020," *Research in Social Stratification and Mobility*, 2021, *71.*

Heckman, James J. and Stefano Mosso, "The Economics of Human Development and Social Mobility," *Annual Review of Economics*, 2014, *6* (1), 689–733.

Hillis, Susan, Juliette. T. Unwin, Yu Chen, Lucie Cluver, Lorraine Sherr, Phillipp Goldman, Olivier Ratmann, Christl Donelly, Samir Bhatt; Andrés Villaveces, Alexander Butchart, Gretchen Bachman, Laura Rawlings, Phil Greena, Charles A. Nelson III, and Seth Flaxman, "Global minimum estimates of children affected by COVID-19-associated orphanhood and deaths of caregivers: a modelling study," *The Lancet*, 2021, *398* (10298), 391–402.

Hossain, Mobarak, "Unequal experience of COVID-induced remote schooling in four developing countries," *International Journal of Educational Development*, 2021, *85*, 102446.

Ichino, Andrea and Rudolf Winter-Ebmer, "The Long-Run Educational Cost of World War II," *Journal of Labor Economics*, January 2004, *22* (1), 57–86.

Instituto Nacional de Estadística y Geografía, "Encuesta para la Medición del Impacto del COVID-19 en la Educación. Presentación de resultados," Technical Report, Instituto Nacional de Estadística y Geografía, Aguascalientes 2021.

Jaume, David and Alexander Willén, "The Long-Run Effects of Teacher Strikes: Evidence from Argentina," *Journal of Labor Economics*, 2019, *37* (4), 1097–1139.

Jordan, Katy, Raluca David, Toby Phillips, and Arnaldo Pellini, "Educación durante la crisis de COVID-19: Oportunidades y limitaciones del uso de Tecnología Educativa en países de bajos ingresos," *Revista de Educación a Distancia (RED)*, ene. 2021, *21* (65).

Jæger, Mads Meier and Ea Hoppe Blaabæk, "Inequality in learning opportunities during Covid-19: Evidence from library takeout," *Research in Social Stratification and Mobility*, 2020, *68*, 100524.

Kaffenberger, Michelle and Lant Pritchett, "Aiming higher: Learning profiles and gender equality in 10 low- and middle-income countries," *International Journal of Educational Development*, 2020, *79*, 102272.

_ and _ , "Failing to Plan? Estimating the Impact of Achieving Schooling Goals on Cohort Learning," Working Paper 20083, Research on Improving Systems of Education 2020.

_ and _ , "A structured model of the dynamics of student learning in developing countries, with applications to policy," *International Journal of Educational Development*, 2021, *82*, 102371.

Lichand, Guilherme, Carlos Alberto Dória, Onicio Leal Neto, and Joao Cossi, "The Impacts of Remote Learning in Secondary Education: Evidence from Brazil during the Pandemic," Working Paper IDB-TN-02214, Inter-American Development Bank 2021.

Maldonado, Joana Elisa and Kristof De Witte, "The effect of school closures on standardised student test outcomes," *British Educational Research Journal*, 2021.

Marcotte, Dave E. and Steven W. Hemelt, "Unscheduled School Closings and Student Performance," *Education Finance and Policy*, 07 2008, *3* (3), 316–338.

McKenzie, David J., "How do Households Cope with Aggregate Shocks? Evidence from the Mexican Peso Crisis," *World Development*, 2003, *31* (7), 1179 – 1199.

Meng, Xin and Guochang Zhao, "The long shadow of a large scale education interruption: The intergenerational effect," *Labour Economics*, 2021, *71*, 102008.

Miranda-López, Francisco, "Infraestructura escolar en México: brechas traslapadas, esfuerzos y límites de la política pública," *Perfiles Educativos*, sep. 2018, *40* (161), 32–52.

Monroy-Gómez-Franco, Luis, "Modeling the learning impacts of educational disruptions: The short and long-run of it.," mimeo 2021.

— , "A note on ex-ante inequality of opportunity across Mexican regions.," mimeo 2021.

— , Roberto Vélez-Grajales, and Gastón Yalonetzky, "Layers of Inequality: Unequal Opportunities and Skin Color in Mexico," *The Review of Black Political Economy*, 2021.

Neidhöfer, Guido, Nora Lustig, and Mariano Tommasi, "Intergenerational transmission of lockdown consequences: Prognosis of the longer-run persistence of COVID-19 in Latin America," *Journal of Economic Inequality*, 2021.

Parolin, Zachary and Emma K. Lee, "Large socio-economic, geographic and demographic disparities exist in exposure to school closures," *Nature Human Behaviour*, 2021.

Plassot, Thibaut, Isidro Soloaga, and Pedro J. Torres, "Inequality of Opportunity in Mexico and its Regions: A Data-Driven Approach," mimeo 2021.

Poirier, Mathieu J. P., Karen A. Grépin, and Michel Grignon, "Approaches and Alternatives to the Wealth Index to Measure Socioeconomic Status Using Survey Data: A Critical Interpretive Synthesis," *Social Indicators Research: An International and Interdisciplinary Journal for Quality-of-Life Measurement*, February 2020, *148* (1), 1–46.

Pritchett, Lant and Justin Sandefur, "Girls' schooling and women's literacy: schooling targets alone won't reach learning goals," *International Journal of Educational Development*, 2020, *78*, 102242.

Prix, Irene and Jani Erola, "Does death really make us equal? Educational attainment and resource compensation after paternal death in Finland," *Social Science Research*, 2017, *64*, 171–183.

Ramírez-Raymundo, Rodolfo, José Gutiérrez-García, and Magdalena Rodríguez de la Huerta, "Orientaciones para apoyar el estudio en casa de niñas, niños y adolescentes. Educación preescolar, primaria y secundaria," Technical Report, Secretaría de Educación Pública, Mexico City 2021.

Rodríguez-Planas, Nuria, "Hitting Where It Hurts Most: COVID-19 and Low-Income Urban College Students," IZA Discussion Papers 13644, Institute of Labor Economics (IZA) August 2020.

Sacerdote, Bruce, "When the Saints Go Marching Out: Long-Term Outcomes for Student Evacuees from Hurricanes Katrina and Rita," *American Economic Journal: Applied Economics*, January 2012, *4* (1), 109–35.

Secretaría de Educación Pública, "Principales Cifras del Sistema Educativo Nacional 2019-2020," Technical Report, Dirección General de Planeación, Programación y Estadística Educativa, Secretaría de Educación Pública, Mexico City 2020.

Steele, Fiona, Wendy Sigle-Rushton, and Øystein Kravdal, "Consequences of family disruption on children's educational outcomes in Norway," *Demography*, 08 2009, *46* (3), 553–574.

Vélez-Grajales, Roberto and Luis Monroy-Gómez-Franco, "Movilidad Social en México: Hallazgos y pendientes," *Revista de Economía Mexicana. Anuario UNAM*, 2017, *2* (1), 97–142.

Vélez-Grajales, Roberto, Luis Monroy-Gómez-Franco, and Gastón Yalonetzky, "Inequality Of Opportunity In Mexico," *Journal of Income Distribution*, 2018, *27* (3-4), 143–157.

Wittenberg, Martin and Murray Leibbrandt, "Measuring Inequality by Asset Indices: A General Approach with Application to South Africa," *Review of Income and Wealth*, 2017, *63* (4), 706–730.

A Sources for estimation of immediate learning loss and parameter values.

Table A.1: Sources for each parameter value.

Parameter	Source	Value
d	Several decrees published by the Federal Government in the Official Diary of the Federation.	64 days of the 2019-2020 school year. 190 days of the 2020-2021 school year.
D	Official basic education calendar, Secretary of Public Education (SEP).	190 days per school year
δ	Assumption for scenario 1.	0.50
	Assumption for scenario 2	0.25
	Assumption for scenario 3	0.00
ψ	Assumption for scenario 1	0.50
	Assumption for scenario 2	0.25
	Assumption for scenario 3	0.00
τ^q	Neidhoefer, Lustig and Tommasi (2021)	15
τ^d	Neidhoefer, Lustig and Tommasi (2021)	5
j	Survey to Measure the Effect of the Pandemic on Education	0.927

Table A.2 Sources for each parameter value

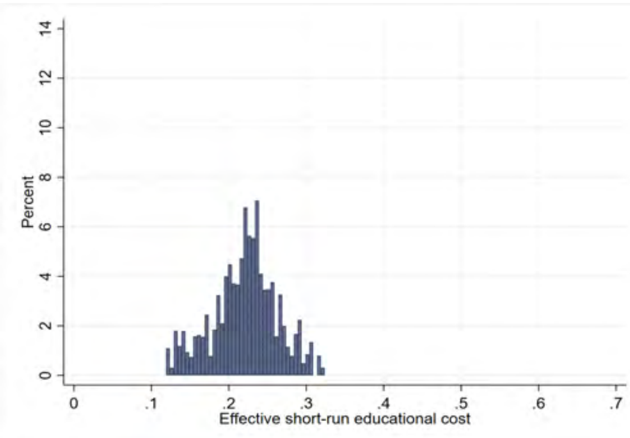
Parameter	Source
γ	National Population Census of 2020
κ	National Population Census of 2020
$P(q=1)$	CONACyT-CentroGeo-GeoInt-DataLab `https://datos.covid-19.conacyt.mx/#DOView`
$P(d=1)$	CONACyT-CentroGeo-GeoInt-DataLab `https://datos.covid-19.conacyt.mx/#DOView`

Table A.3: Values of the cost parameters

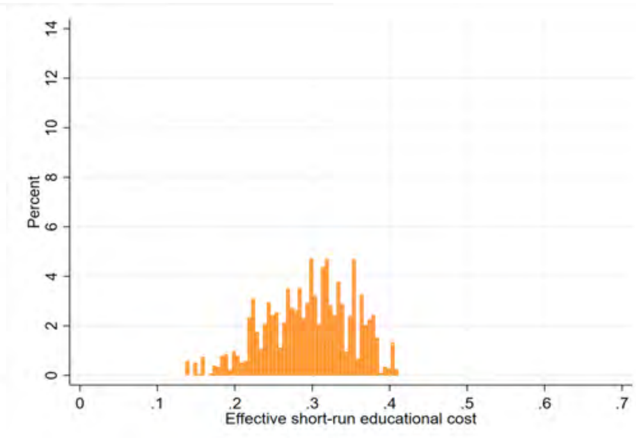
Parameter	Region	Value for year 1	Value for year 2
γ	North	0.9418	0.9418
	North West	0.9225	0.9225
	Centre North	0.9325	0.9325
	Centre	0.9338	0.9338
	South	0.8283	0.8283
κ	North	0.665	0.6265
	North West	0.5119	0.5119
	Centre North	0.5567	0.5567
	Centre	0.5646	0.5646
	South	0.3717	0.3717
$P(q=1)$	North	0.0031	0.0368
	North West	0.0035	0.0470
	Centre North	0.0021	0.0364
	Centre	0.0042	0.0447
	South	0.0068	0.0847
$P(d=1)$	North	0.0007	0.0032
	North West	0.0005	0.0039
	Centre North	0.0003	0.0044
	Centre	0.0007	0.0049
	South	0.0011	0.0042

B Effective immediate learning costs.

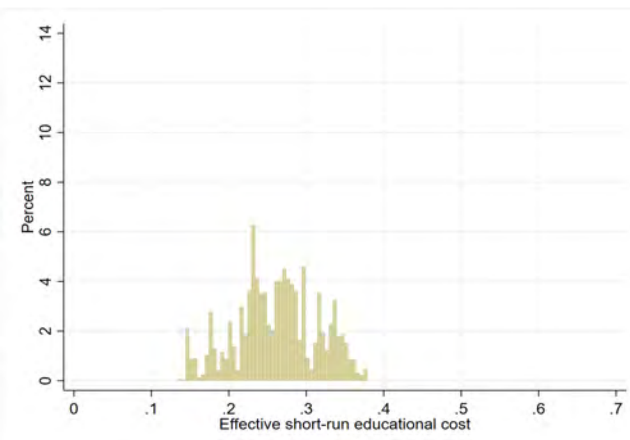
Figure 4. Regional distributions of effective immediate learning costs.
(scenario 1)



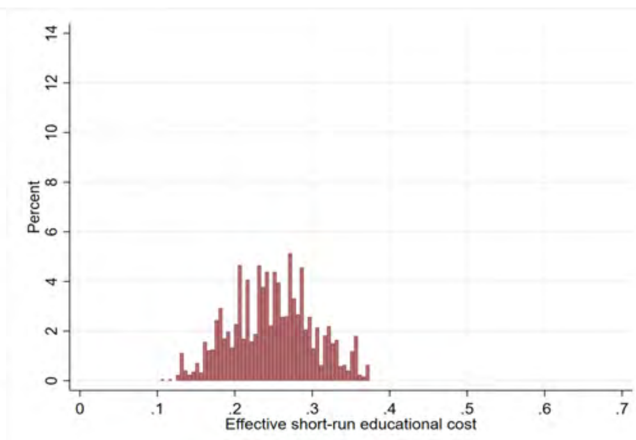
(a) North



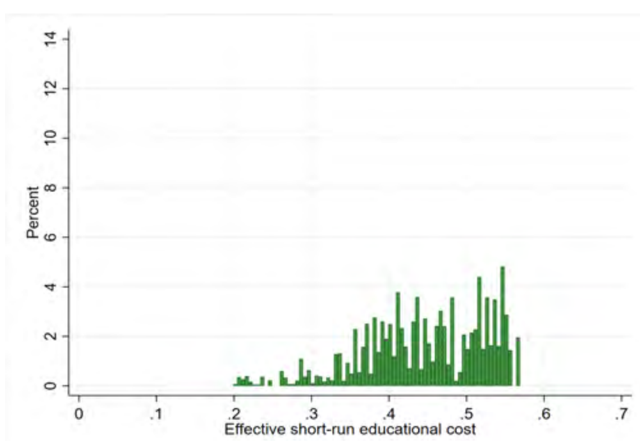
(b) North West



(c) Centre North



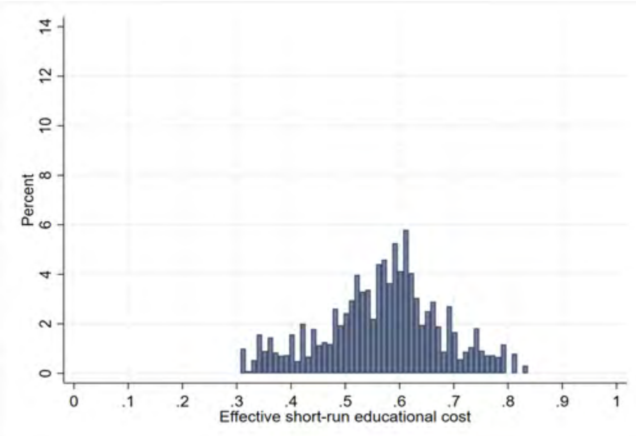
(d) Centre



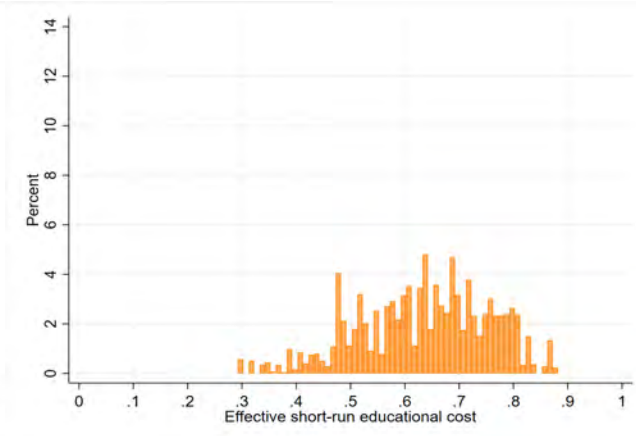
(e) South

Note: Calculations from the authors. The effective short-run cost corresponds to the share of a school year of learning lost due to the transition to remote learning.

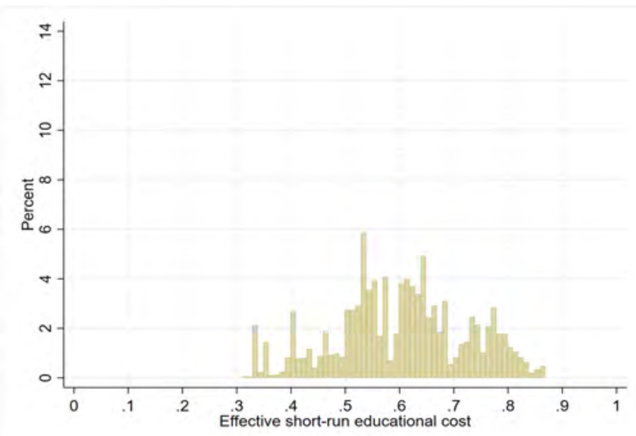
Figure 5. Regional distributions of effective immediate learning costs.
(scenario 2)



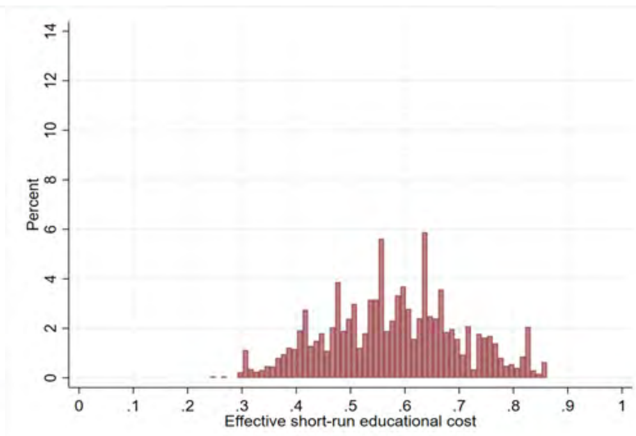
(a) North



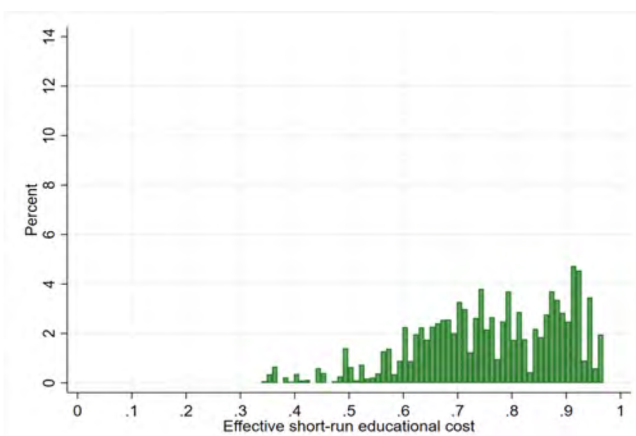
(b) North West



(c) Centre North



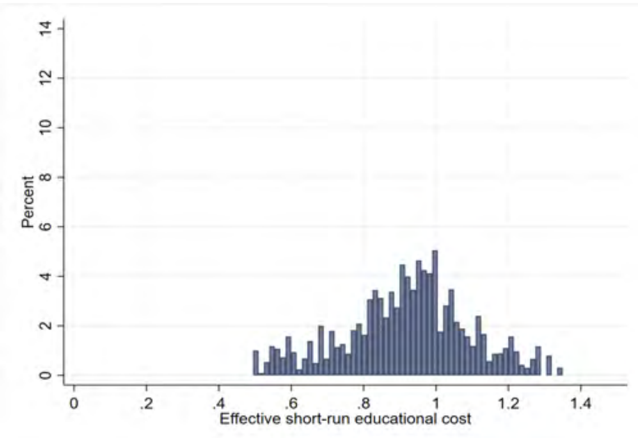
(d) Centre



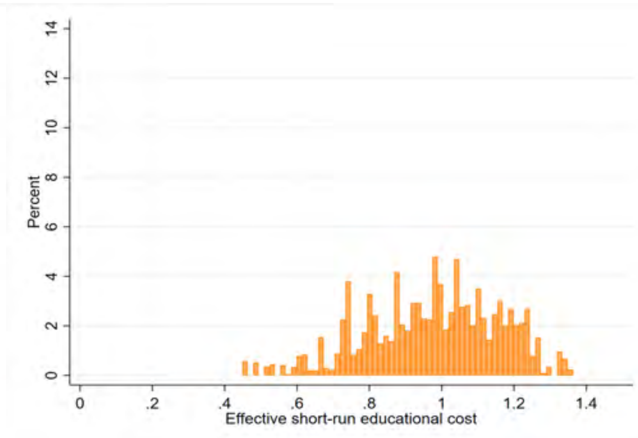
(e) South

Note: Calculations from the authors. The effective short-run cost corresponds to the share of a school year of learning lost due to the transition to remote learning.

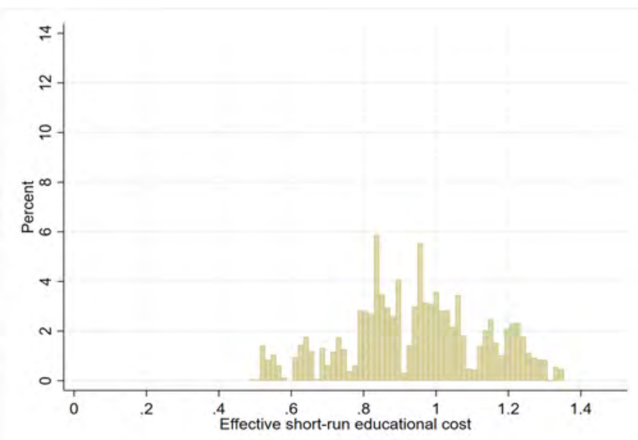
Figure 6. Regional distributions of effective immediate learning costs.
(scenario 3)



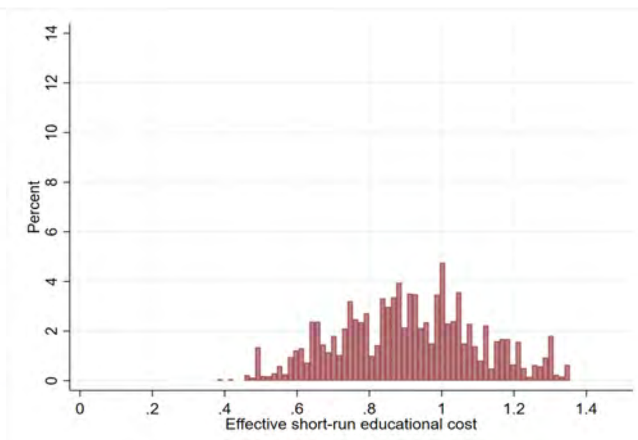
(a) North



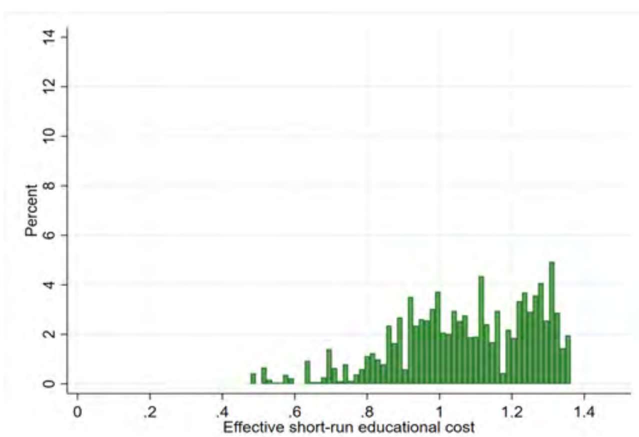
(b) North West



(c) Centre North



(d) Centre



(e) South

Note: Calculations from the authors. The effective short-run cost corresponds to the share of a school year of learning lost due to the transition to remote learning.

C Calibration parameters of the Potential Pedagogical Function

To calibrate the Pedagogical Production Function (PPF) parameters described by equation six, we employ the 2017 results of the math component of the standardised test of the National Plan for the Evaluation of Learnings (Programa Nacional para la Evaluación de los Aprendizajes, PLANEA in Spanish). Specifically, we use the results from the ninth-grade test applied nation-wide to all students at the specified grade. This makes the data from the tests representative at the national and regional dimensions. Furthermore, as the National Institute for the Evaluation of Education (Instituto Nacional para la Evaluación de la Educación) was in charge of the design, the institute's closing in 2019 made the 2017 results the latest available.

Our calibration strategy consists of two steps. The first stage implies calibrating a PPF to the national scores to identify the pacing of learning across grades from first to the ninth grade. For the decision on the type of initial distribution and its variance, we follow Kaffenberger and Pritchett (2020b). Table C.1 shows the resulting values for the parameters of the national PPF. We use the estimated ρ for each one of the five regional PPF. The parameter values for these PPF are shown in table 3 of the main text.

Table C.1: Calibrated parameters for National results

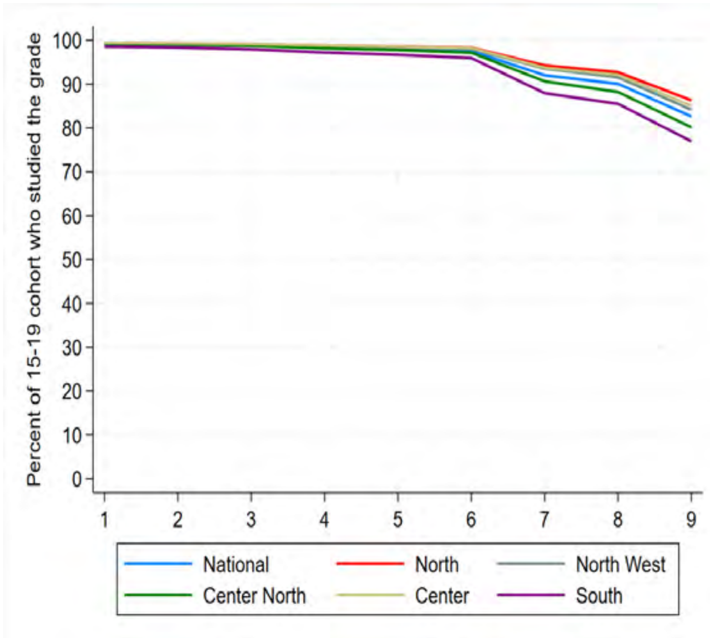
Parameters	National
Average math score (Total sample)	497
Standard deviation of math score (Total sample)	120
w	153
$z_{\min}$	37.50
$z_{\max}$	71.25
ρ	54
v	0.2205
Initial distribution	N(0,20)

Note: Data from the PLANEA 2017 ninth-grade test and authors' calibrations.

To simulate a cohort's educational trajectory, we also require the grade to grade dropout rates for each country region. Finally, we use the 2020 National Population Census information to estimate the attainment rates for each school grade and show them in figure 7.

Using this information, we iteratively estimate the parameter values of the regional models such that they replicate the regional average math scores in PLANEA 2017 and the dropout rates of each region. Finally, we show the parameter values that fulfil both conditions in table 3 of the main text.

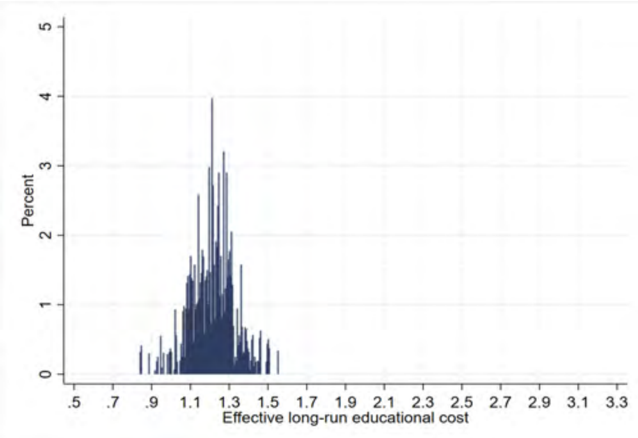
Figure 7. Grade attainment profile for Mexico for a recent cohort of 15-19-year-olds



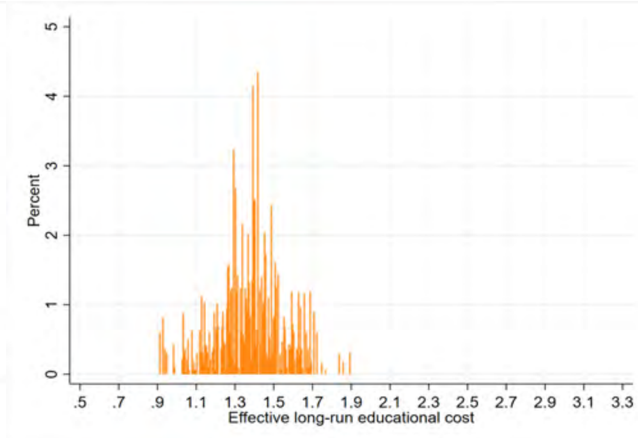
Note: Data from the Mexican National Population Census of 2020.

D Effective long-run or cumulative learning costs

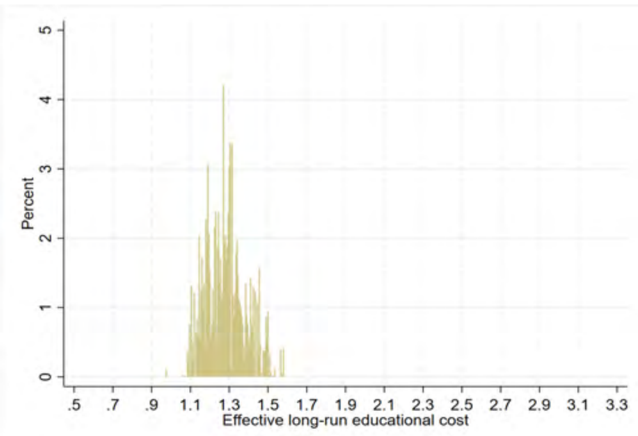
Figure 8. Regional distributions of effective cumulative learning costs.
(scenario 1)



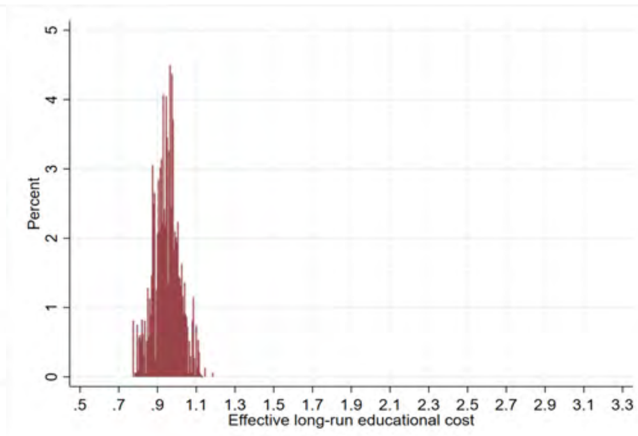
(a) North



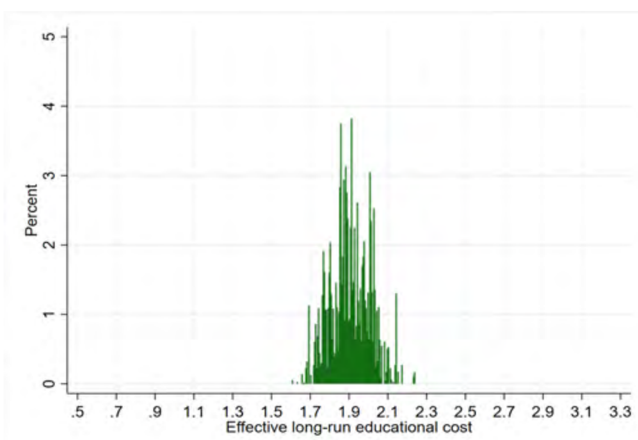
(b) North West



(c) Centre North



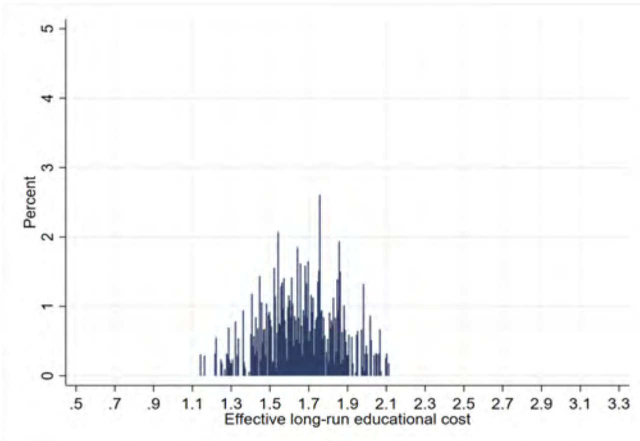
(d) Centre



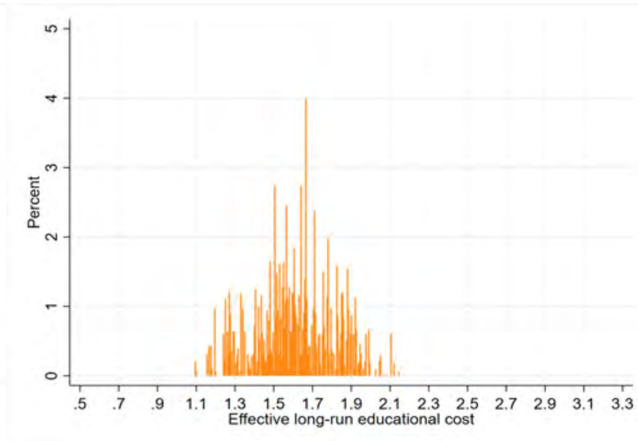
(e) South

Note: Calculations from the authors. The effective long-run cost corresponds to the number of years of learning progression a student is behind the expected learning stock at ninth grade.

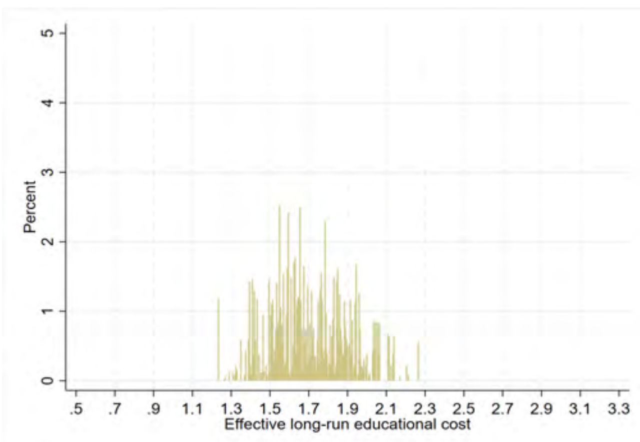
Figure 9. Regional distributions of effective cumulative learning costs.
(scenario 2)



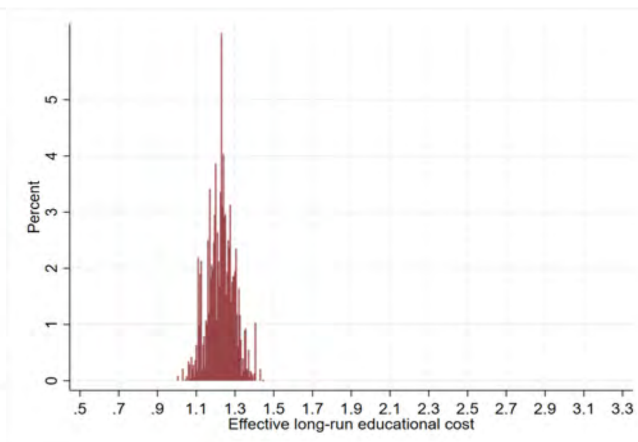
(a) North



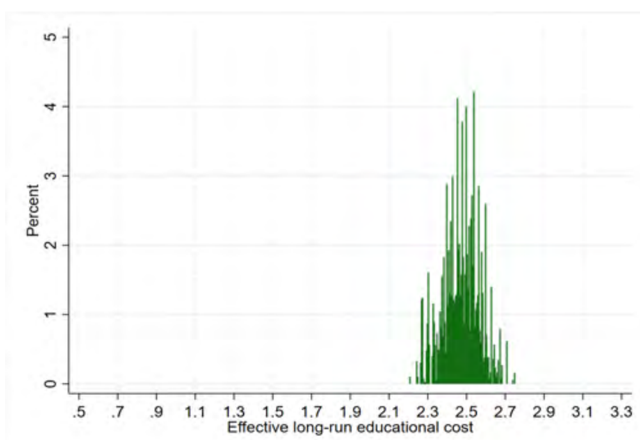
(b) North West



(c) Centre North



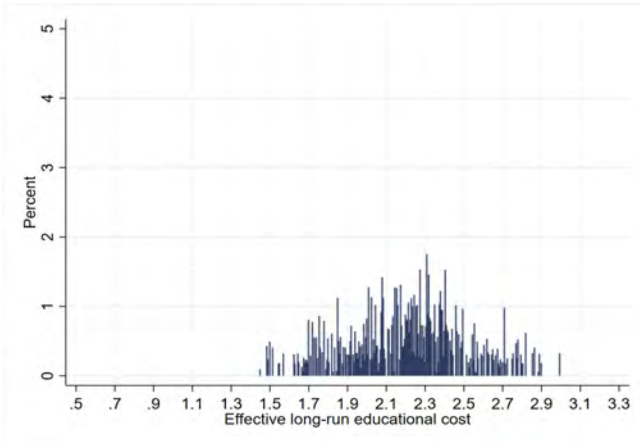
(d) Centre



(e) South

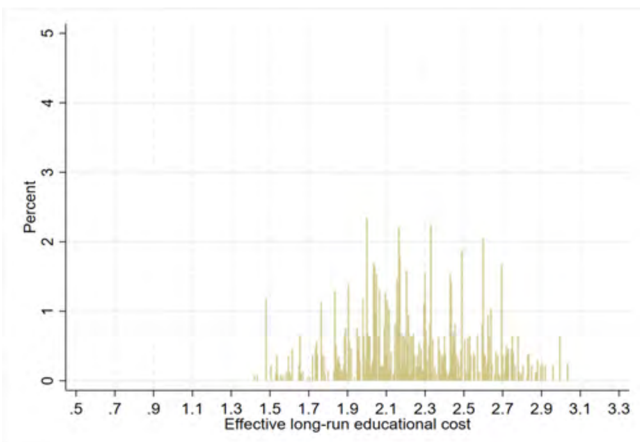
Note: Calculations from the authors. The effective long-run cost corresponds to the number of years of learning progression a student is behind the expected learning stock at ninth grade.

Figure 10. Regional distributions of effective cumulative learning costs.
(scenario 3)

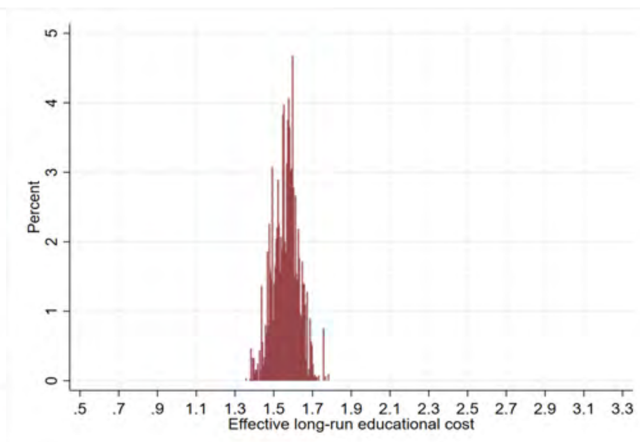


(a) North

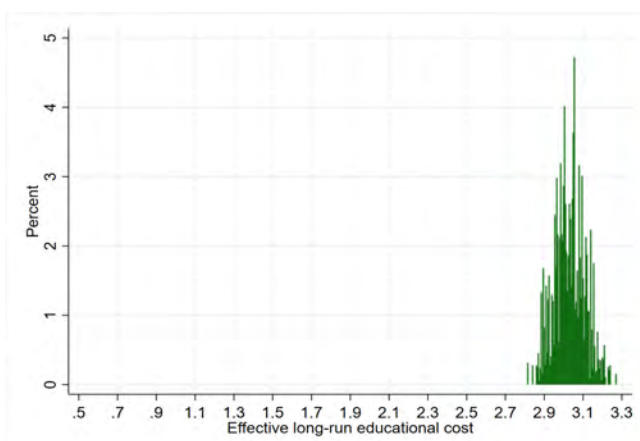
(b) North West



(c) Centre North



(d) Centre



(e) South

Note: Calculations from the authors. The effective long-run cost corresponds to the number of years of learning progression a student is behind the expected learning stock at ninth grade.