NBER WORKING PAPER SERIES

AGGREGATE AND DISTRIBUTIONAL EFFECTS OF 'FREE' SECONDARY SCHOOLING IN THE DEVELOPING WORLD

Junichi Fujimoto
David Lagakos
Mitchell VanVuren

Working Paper 31029
http://www.nber.org/papers/w31029

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2023

For helpful comments we thank Ufuk Akcigit, Marco Bassetto, Natalie Bau, Kevin Donovan, Raquel Fernandez, Chang-Tai Hsieh, Aubhik Khan, Jeremy Pearce, Karthik Muralidharan, Michael Peters, Tommaso Porzio, Pascual Restrepo, Todd Schoellman, Julia Thomas and seminar/conference participants at Berkeley, Chicago, City University of Hong Kong, the Development Bank of Japan, the Federal Reserve Banks of Minneapolis and St. Louis, Georgetown, Kyoto, NYU, Ohio State, Penn State, Riverside, the Cowles Foundation Conference on Structural Change, and the NBER Summer Institute. We thank the IGC, JSPS (Kakenhi No. 18K01505 and No. 22K01399), and GRIPS Policy Research Center for financial support and Ariel Akerman, Soala Ekine, and Isaac Otoo for outstanding research assistance. All potential errors are our own. The authors have no relevant financial disclosures. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Aggregate and Distributional Effects of ‘Free’ Secondary Schooling in the Developing World
Junichi Fujimoto, David Lagakos, and Mitchell VanVuren
NBER Working Paper No. 31029
March 2023
JEL No. E24,O11

ABSTRACT

This paper analyzes the aggregate and distributional effects of publicly funded merit-based (‘free’) secondary schooling in the developing world. Our analysis is based on an overlapping-generations model of human capital accumulation in which households face borrowing constraints that can lead to misallocation of talent in equilibrium. We estimate the model to match a randomized controlled trial that provided poor but talented children in Ghana with scholarships for secondary education. The model predicts that a nationwide free secondary schooling policy is largely redistributive in nature, with modest gains in GDP and average welfare. Policies that spend the same amount on improving education quality result in larger welfare gains for households of all income levels

Junichi Fujimoto
GRIPS
Tokyo, Japan
j-fujimoto@grips.ac.jp

David Lagakos
Department of Economics
Boston University
270 Bay State Road
Boston, MA 02215
and NBER
lagakos@bu.edu

Mitchell VanVuren
Yale University
34 Hillhouse Avenue
P.O. Box 208206
New Haven, CT 06520-8206
mitchell.vanvuren@yale.edu

1. Introduction

It has been said that talent is universal but opportunity is not. This saying could hardly apply better than to the millions of school-aged children throughout the developing world that are not enrolled in school. Whereas enrollment rates for primary education have seen dramatic increases in recent years, only around one in three secondary-school aged children in the developing world actually attends secondary school. Recent evidence suggests that a main reason many young people do not attend secondary schooling is credit constraints that prevent their families from borrowing. Keeping bright young people out of secondary school may lead to a significant misallocation of talent in the education system, which can reduce aggregate productivity levels (Hsieh, Hurst, Jones, and Klenow, 2019). More generally, low levels of human capital in developing countries are thought to be one of the most important proximate causes of their low income levels (e.g. Hall and Jones, 1999; Bils and Klenow, 2000; Manuelli and Seshadri, 2014; Hendricks and Schoellman, 2018) as well as a barrier to structural change (Porzio, Rossi, and Santangelo, 2022; Buera, Kaboski, Rogerson, and Vizcaino, 2022).

A number of developing countries have recently adopted 'free' schooling policies with the hopes of raising secondary enrollments. These policies have nearly all had two features in common: taxpayer funding of education, with little or no direct costs for students, and merit requirements, typically in the form of an entrance exam. Free schooling policies have generally been viewed as successes by policymakers, largely because they have been followed by noticeable increases in secondary enrollments (Center for Global Development, 2022). Yet the overall impacts of free schooling policies are not as easy to measure. The human capital gains from expanding school enrollments matter more than the enrollments themselves, and the direct costs to taxpayers have to be taken into consideration. So do the opportunity costs of lost working years, especially since most secondary-school aged individuals in poor countries are in, or approaching, their prime working years.

In this paper, we estimate the aggregate and distributional effects of free secondary schooling in the developing world. To our knowledge, ours is the first macroeconomic study of these issues. Our analysis is based on a general-equilibrium overlapping generations (OLG) model of human capital accumulation through schooling. In the model, parents choose their children's education level and face credit constraints that prevent borrowing against future income. Learning ability is passed down stochastically from parent to child. Population growth is endogenous and depends on the child's education, with lower fertility for those with higher education levels. Parents base the educational choice for their children on their family's income, assets, and taste for schooling, plus the child's score on a qualifying exam, which is modeled as a noisy signal of ability. Misallocation of talent arises in the model when children with high ability complete less schooling than they otherwise would have because of their parents' low income and asset levels.

We estimate the model using experimental evidence from a long-term study that offered secondary school scholarships to a randomly selected set of poor but high-ability children in Ghana (Duflo, Dupas, and Kremer, 2021). Those offered the scholarships were about 25 percent more likely to finish secondary school than a control group four years hence. Scholarship winners performed about 0.2 standard deviations higher on tests of literacy and mathematics, which is comparable to the effects found in other successful education interventions. Earnings for scholarship winners were higher, though imprecisely estimated, and fertility rates were significantly lower than in the control group.

We use the estimated model to simulate the effects of a nationwide free secondary schooling policy in general equilibrium. The model predicts an increase in the number of secondary school graduates by around 10 percent. The policy has negligible effects on GDP per capita, which rises by just 0.1 percent in the long run, and somewhat larger effects on average welfare, which rises by around 2 percent in consumption equivalents. The policy is not budget neutral, and pays for only about ten percent of its cost in the long run. While adult earnings do rise for those treated by the policy, the gains are largely offset by lost earnings during schooling years, signaling an important role for opportunity cost in holding back secondary attendance. We show that the welfare gains from free schooling largely accrue to the poorest households, who see the largest increases in schooling completion and relative wages, but who pay little of the resulting tax increases.

To understand why the model predicts such modest increases in GDP per capita, we conduct a series of alternative estimations of the model which target more favorable experimental moments than the ones actually estimated. We find that no single estimation target is responsible for the model's pessimistic conclusions, but that a combination of substantially different experimental outcomes could have led the model to predict GDP gains of around 10 percent from free schooling. For such a large predicted gain, the model requires much higher schooling quality and costs, and significantly less ability for households to save around borrowing constraints.

We show that this combination of parameters implies considerable misallocation of talent in equilibrium, but predicts counterfactually large treatment effects on schooling completion and test scores, and inaccurate predictions about which part of the test score distribution reacts most when offered scholarships.

As a frame of reference, we compare the effects of free secondary schooling to an economy-wide improvement in schooling quality, which could represent pay-for-performance incentives for teachers (Muralidharan and Sundararaman, 2011; Duflo, Hanna, and Ryan, 2012; Mbiti, Muralidharan, Romero, Schipper, Manda, and Rajani, 2019), additional teachers in the classroom (Banerjee, Cole, Duflo, and Linden, 2007) or other interventions shown to bolster student academic performance. We find that school quality improvements – costing the same amount as the free schooling policy – are significantly more effective at raising average income and welfare levels. A nationwide school quality improvement raising test scores by 0.1 standard deviations, and costing the same per student on average as in the experiments above, leads to a GDP increase of 4 percent and an average welfare gain of around 5 percent. This time the biggest winners in welfare terms are the top quartile of the income distribution, which have the highest secondary enrollment rates to begin with, though even the bottom quartile gains more under the school quality improvement policy than through free schooling.

Our quantitative analysis suggests that free secondary schooling policies are predominantly redistributionary in nature, at least at the current schooling quality levels. While a nontrivial number of students are misallocated under a privately funded system, our estimated model implies that the majority of those not attending secondary school would have low returns to education and high opportunity cost from lost work years. This suggests that low secondary school enrollments are largely an efficient response to low quality school options. This implication is broadly in line with the conclusions of the macro development literature emphasizing low schooling quality, rather than low average years of schooling, as a determinant of income levels (Hanushek and Woessmann, 2007; Schoellman, 2012).

We conclude by providing survey evidence supporting our model's prediction that free schooling raises welfare more for less-educated households then more-educated ones, even though the former are less likely to attend secondary school. To do so we ran a nationally representative survey of 3,500 households in Ghana about their attitudes toward free secondary schooling, which was enacted in Ghana in 2017. We find that, just as our model predicts, the least-educated households are those that support free schooling the most, even with its merit requirements. This is in contrast to patterns found in the United States, where the most educated typically show the strongest support for public tertiary education (see e.g. Fernández and Rogerson, 1997). We argue that the difference arises from the fact that the least-educated households in Ghana pay a much smaller share of taxes for education than their counterparts in the United States.

Related Literature. Our quantitative exercises build on the large literature on the macroeconomic effects of credit constraints in education in advanced countries, such as the seminal work of Lochner and Monge-Naranjo (2011). Our paper is most closely related to the studies by Abbott, Gallipoli, Meghir, and Violante (2019) and Daruich (2020), both of whom study expansions in publicly funded education in the United States. Both studies reach fairly positive conclusions about the effects of expanding public education, unlike our study, which arguably reflects disparities in school quality between rich and poor countries. As in Daruich (2020), we discipline our model using experimental evidence from a randomized controlled trial.[1]

Our paper also builds on the recent literature attempting to quantify the extent to which credit market imperfections drive misallocation in developing countries. Bassi, Muoio, Porzio, Sen, and Tugume (2022) and Caunedo and Kala (2022) show how rental markets for large indivisible capital goods can reduce capital misallocation, and Moll (2014) and Midrigan and Xu (2014) find a significant ability for firms to save their way around credit constraints.

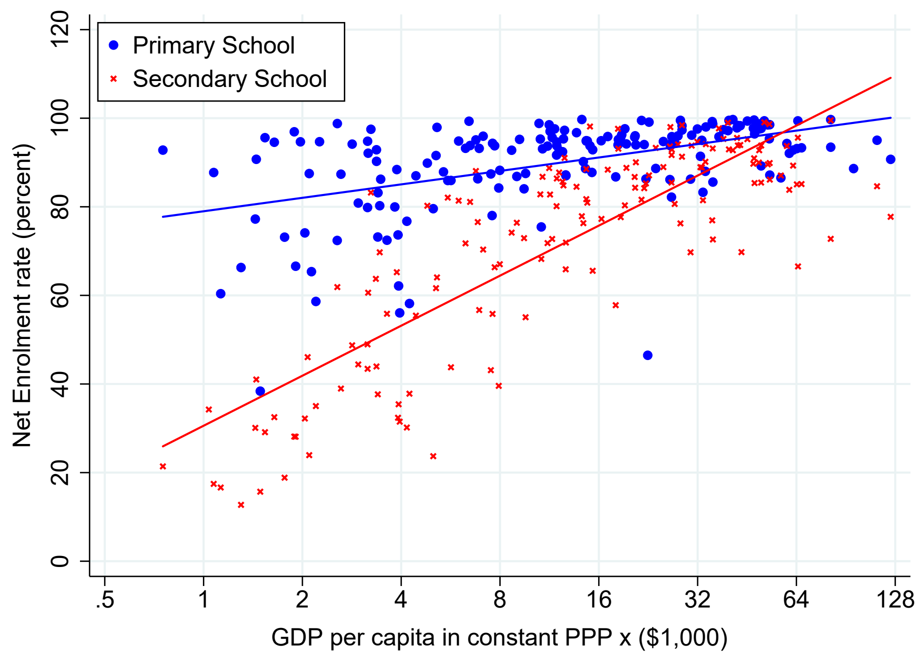
In estimating our model to a field experiment, we build on a growing body of macroeconomic research on development that uses randomized experiments in order to guide general-equilibrium counterfactuals (Buera, Kaboski, and Townsend, Forthcoming). Ours is the first to take this approach when studying education policy in the developing world. Other studies using this methodology have studied small business investment (Kaboski and Townsend, 2011), occupational choice (Buera, Kaboski, and Shin, 2021), infrastructure investments (Brooks and Donovan, 2020), rural-urban migration (Lagakos, Mobarak, and Waugh, Forthcoming), and firm training programs (Akcigit, Alp, and Peters, 2021).

[1]Our work builds on the seminal macro studies of income distribution and human capital, such as Galor and Zeira (1993), Banerjee and Newman (1993), and Bénabou (2002). Among more recent studies, Celik (2023) and Akcigit, Pearce, and Prato (2020) show that, in advanced economies, misallocating the talent of those with high ability for innovation can substantially reduce growth. Our work also builds on the large literature on intergenerational mobility, and has a model setup that is closely related to that of Hassler, Rodríguez Mora, and Zeira (2007).

2. Secondary Schooling in the Developing World

We begin by summarizing the main facts about secondary schooling outcomes and policies in the developing world. These help motivate our modeling choices and counterfactual simulations in the sections that follow.

Figure 1: Primary and Secondary School Enrollment Rates



Aggregate data on schooling enrollment show plainly that developing countries mainly lag behind richer ones when it comes to secondary schooling (as opposed to primary schooling). Figure 1 plots net enrollment rates in primary school (blue dots) and secondary school (red x's) in 2019 against GDP per capita using data from the World Bank. Net enrollment rates are defined as the number of people enrolled in school relative to the population of school-aged individuals. In the world's poorest countries, roughly four out of five children of primary-school age are enrolled in school, compared to nearly every child in the richest ones. For secondary schooling, the differences are much starker. At the bottom of the world income distribution, only around one-third of those of secondary-school age are enrolled in secondary school, whereas at the top, enrollment rates are again near one hundred percent.

One salient difference between rich and poor countries in terms of education policy is that richer countries are much more likely to publicly finance secondary education. It is not surprising then that many developing countries have recently considered implementing 'free' schooling policies, in which the government finances school fees for at least some secondary-age students (Center for Global Development, 2022). One main rationale for publicly funded schooling is to help raise average schooling levels, which is agreed to be a key determinant of GDP per capita. A second rationale is to make secondary education more accessible to poorer households, consistent with redistributionary motives. These two objectives are not necessarily in contrast with one another, since raising average years of schooling is likely to require expanding schooling access to poorer households that were previously unable to pay for secondary school fees.

One common feature of these free secondary school policies is a merit requirement, usually coming in the form of an eligibility exam. In a set of thirteen countries recently enacting free secondary education, we found that all but one required an eligibility test of some kind (see Appendix Table A.1). For example, Kenya requires that students pass their Certificate of Primary Education Exam, and Rwanda requires a 'high' score or better on their Ordinary Level Exam. These merit requirements likely serve two basic purposes. First, they leave governments with an additional lever to control the inflow of new secondary school enrollees each year, which helps control costs. Second, they focus the new secondary admissions on the most able students, which are those that are most likely to be misallocated to begin with.

Recently, a number of micro studies have estimated the impacts of merit-based scholarship programs in developing countries, though with mixed results. Brudevold-Newman (2021) found, using a difference-in-difference approach, that free secondary schooling in Kenya increased educational attainment, reduced fertility, and increased the likelihood of skilled work. Using a regression discontinuity design in Cambodia, Filmer and Schady (2014) found that scholarships increased educational attainment but did not increase earnings, fertility, or test scores. Both studies highlight credit constraints as a reason more students were not already enrolled.

Duflo, Dupas, and Kremer (2021) conducted the first long-run RCT evaluation of a merit-based scholarship program for secondary school. Their study is set is Ghana, where the education system consists of primary school and junior high school (JHS) until age 14, at which point students are required to pass the Basic Education Certification Examination (BECE) in order to attend senior high school (SHS). The authors identified approximately two thousand students who had passed the BECE in 2008 but had not enrolled in SHS by the deadline for the next school year. Among these students, one-third were randomly selected to receive a four-year scholarship covering one hundred percent of tuition and fees.

Students who received a scholarship were substantially more likely to complete SHS relative to those in the control group, although it is worth noting that many students in the control group did eventually graduate. They also exhibited higher human capital, as measured by math and reading tests, and lower medium-run fertility (for females), having fewer children after 12 years. We view this experiment as the most comprehensive and credible evaluation of free secondary schooling to date. Consequently, we use these experimental moments to parameterize our model, which we develop in the next section.

3. Overlapping Generations Model

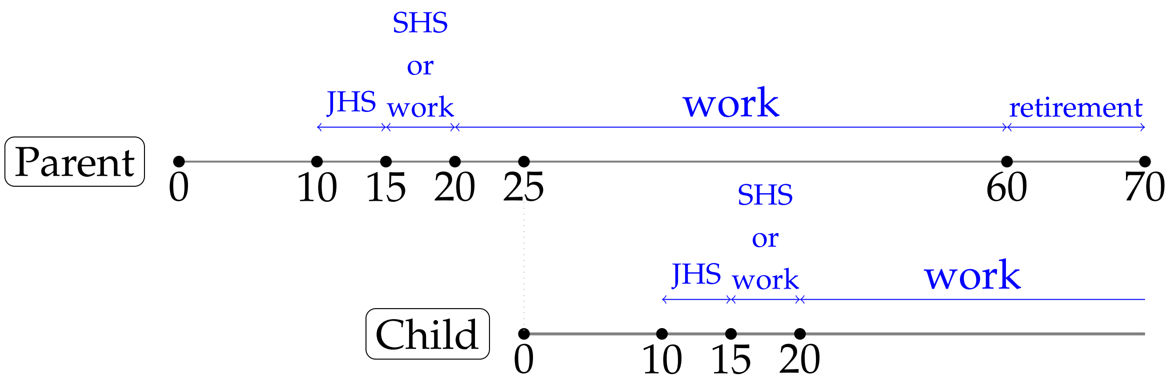
We now describe the model, which we tailor to be able to match the key features of the experimental evidence described above. Importantly, we allow for misallocation of talent through borrowing constraints but also include other factors that keep young people out of secondary school, in particular low education quality, which depresses the returns to secondary education, and opportunity cost of lost work years, which raises the cost of schooling. Including multiple explanations for low secondary enrollments allows us to speak to the relative importance of each in driving the aggregate effects of education expansions.

We also include several other features that are relevant for the question at hand. Consistent with the evidence of Khanna (2023), we model the labor of different education types as imperfect substitutes, so that an increase in the supply of educated workers depresses their relative wages. As in a growing literature in macro development we allow for saving constraints, which help match the low average levels of liquid asset holdings, and impede households from simply saving around borrowing constraints (see e.g. Donovan, 2021). Following the literature on public finance and development we posit a tax system in which a narrow base of high earners pay the majority of the taxes used to finance public expenditures (see e.g. Jensen, 2022).

3.1. Environment

Time is discrete and goes from 0 to infinity. There is a single good which can be used for consumption, savings, and investment in education. The economy is populated by overlapping generations of households that are heterogeneous in their parental human capital, child ability, taste for schooling, and savings. The timeline of events for these households is shown in the graphic below.

Individuals live for 14 periods, where each period corresponds to 5 years. For their first five periods of life (ages 0–24) children live with their parents. In the third period (ages 10–14), all children attend Junior High School. We abstract from the choice of attending school at this age based on the evidence of the previous section that virtually all children already attend Junior High School. In the fourth period, (ages 15–19), children either attend Senior High School or work. This is the key schooling choice in the model. The fact that a household must give up a period of a child's work, and thus income, in order to attend secondary school captures the notion that, even in the case where schooling is made free, there remains an opportunity cost. In the fifth period, all children work using their respective education level, which is fixed for the remainder of their life.



At the beginning of period six, when turning age 25, children leave their parents, have children, and become parents themselves. We abstract from household formation decisions since they do not seem crucial for our task at hand. Instead, we model households as continuous dynasties that do not mix. These new parents then work from age 25 to 60, at which point they retire, and die at age 70. This is roughly the average life expectancy in Ghana, for example, whose features we will use to parameterize our model in the following section.

Each new household consists of a parent aged 25 and newborn children. The model features population growth, and the number of children, denoted as $1+\nu_{s_p}$, is allowed to depend on the parent's schooling level s_p. As a result, policies that change an individual's level of schooling will also change their fertility, consistent with a variety of evidence.

Individuals are heterogeneous in learning ability $z \in Z = \{z^1, z^2, \ldots, z^N\}$. The ability within a household follows a first-order Markov chain which mimics the AR(1) process:

$$\log z_c = \rho \log z_p + \epsilon, \;\; \rho \in (0,1)\,. \tag{1}$$

Here, z_p and z_c denote the parent and children's ability. Throughout, variables with superscript p and c pertain to parents and children, respectively. The shock ϵ is a zero mean i.i.d. random variable. Thus, ability is transmitted within each household but only imperfectly, and is identical across siblings. Following the evidence in e.g. Cunha and Heckman (2007), we interpret ability to be a function of inherited capabilities and parental inputs.

All household decisions are made by parents, who derive flow utility $U(c) = \log(c)$ from household consumption $c \geq 0$ and discount the future at rate $\beta \in (0, 1)$. The assumption that parents make educational decisions is consistent with evidence that parents in low-income countries predominantly take an authoritarian approach to parenting, dictating decisions directly rather than trying to reach an agreement with children (Doepke and Zilibotti, 2017). Parents and children (from ages 15 to 25) have a single unit of time each period which they supply inelastically to wage work or education. Parents are imperfectly altruistic toward children and therefore derive utility also from children's well-being (as in e.g. Laitner, 1997).

Parents make schooling decisions for their children when the children turn age 15, after observing the children's ability and test scores as well as the children's realization of a schooling taste shock. More precisely, children enjoy random utility (internalized by the parent through imperfect altruism) δ_s from schooling level $s \in S = \{J, S\}$ (JHS, SHS), where δ_s follows a standard Gumbel distribution with scale parameter θ. Parents must forgo a period of children's income to send their children to an additional period of school, and further, providing children final schooling level $s \in S$ requires goods costs Ψ_s. These goods costs represent school fees and satisfy $\Psi_S > \Psi_J = 0$, where the equality reflects the free primary education that prevails in most developing countries. Thus when deciding whether to send their children to school, parents must consider the opportunity cost of their children's work as well as the explicit goods cost.

Households face incomplete markets as in Aiyagari (1994), Bewley (1977), and Huggett (1993) and cannot borrow but can save at an exogenous rate r. While households do face idiosyncratic income risk, the most important feature of this borrowing constraint is that it prevents parents from borrowing against their child's future income in order to fund school attendance. This allows for the possibility that a high ability child, whose return to additional schooling far exceeds the cost, may not attend if born to a poor parent, resulting in misallocation.

To capture the fact that one must pass an entrance test to enter secondary schooling in most developing countries (as discussed in Section 2), we set a threshold test score for entering SHS. One's test score $\tilde{z}$ is related to ability as

$$\tilde{z} = z + \varepsilon, \tag{2}$$

where the noise ε follows a normal distribution with mean zero and standard deviation σ_ε. The human capital of an individual with ability z and schooling level s is given by

$$h(z,s) = \begin{cases} 1 & \text{if } s = J, \\ z \cdot \eta_S & \text{if } s = S, \end{cases} \tag{3}$$

where $\eta_S > 0$ and represents the efficiency, or quality, of schooling. Thus, ability affects human capital only for those with SHS education, and the resulting human capital of a secondary education depends on the product of the student's ability and the schooling quality.

Markets are competitive and the aggregate production function, operated by a representative profit-maximizing firm, is given by:

$$Y = AK^\alpha \left[(N_J)^\lambda + (N_S)^\lambda \right]^{\frac{1-\alpha}{\lambda}}, \ \alpha, \lambda \in (0,1). \tag{4}$$

Here, A is aggregate productivity, K is physical capital, and N_s is aggregate efficiency units of labor of individuals with schooling level s. The firm rents physical capital from households or foreign investors at an exogenous international market rate r^*. Due to savings frictions, however, the return to physical capital for households is lower, at $r = r^* - \chi < r^*$. This lower return to capital helps us match the low savings rates among households in low-income economies (as in Donovan, 2021).

The labor income y of an individual equals the product of three terms. The first term is the wage rate per efficiency units of unskilled ($s = J$) or skilled ($s \in S$) labor, denoted as w^U or w^S, respectively. The second term, ζ, represents idiosyncratic shocks to labor productivity. The third term is human capital $h(z,s)$, given by (3). For example, the labor income of an individual with education level S is given by

$$y(z,S,\zeta) = w^S \zeta h(z,S) = w^S \zeta z \eta_S. \tag{5}$$

3.2. Parents' Problems

Parents make consumption and saving decisions in each period, and additionally, schooling decisions when their children reach the age for secondary school. We discuss below the parents' problems in the key periods in the life-cycle; we omit the description of their problems in other periods, which are standard consumption-savings problems. In addition to individual state variables described below, the parent's problems depend on the p.d.f. f describing the distribution of households across individual states and the aggregate population level P.

When $\tau = 9$, and children turn 15, parents observe the realizations of the schooling taste shocks (δ_J, δ_S), children's ability and test score $(z_c, \tilde{z}_c)$, and their own and children's labor productivity (ζ_p, ζ_c). Then, if $\tilde{z}_c$ weakly exceeds the threshold test score $\bar{z}$, parents have an option to send children to an additional period of schooling $(s = S)$. The value function of such parents with ability z_p, schooling level s_p, and assets a is given by

$$V_9(a, z_p, s_p, \zeta_p, \delta_J, \delta_S, z_c, \zeta_c; f, P | \tilde{z}_c \geq \bar{z}) = \max_{c \geq 0, a' \geq 0, s'_c \in \{J,S\}} \log(c) + \delta_J \mathbf{I}\left(s'_c = J\right) + \delta_S \mathbf{I}\left(s'_c = S\right) + \beta \mathrm{E}\left[V_{10}(a', z_p, s_p, \zeta'_p, z_c, s'_c, \zeta'_c; f', P')\right]$$

where the maximization is subject to the flow budget constraint

$$a' + c + \left(1 + \nu_{s_p}\right) \mathbf{I}\left(s'_c = S\right) \Psi_S = y_p(z_p, s_p, \zeta_p) + (1+r)a + \left(1 + \nu_{s_p}\right)(1 - \mathbf{I}\left(s'_c = S\right)) y_c(z_c, J, \zeta_c) - T(z_p, s_p, \zeta_p, z_c, J, s'_c, \zeta_c) \tag{6}$$

and the perceived laws of motion for the aggregate state variables f and P, given by $f' = F(f, P)$ and $P' = H(f, P)$, respectively. Here, the prime denotes values of variables in the next period and T is total amount of taxes paid by the household, which depends on the parent and children's labor income, and is therefore a function of $(z_p, s_p, \zeta_p, z_c, J, s'_c, \zeta_c)$.[2] We suppress the dependence of y_p, y_c, and T on f and P except where it is necessary to make that dependence explicit.

When $\tau = 10$, children live one final period with their parents and work with the human capital given by their education decision the previous period. The value

[2] Note that T depends on both the children's current schooling level s_c $(= J)$ and next period's schooling level s'_c. This is because the labor income depends on educational attainment, and only the children who do not go to school $(s'_c = s_c)$ earns the labor income in the current period.

function of such parents is expressed as

$$V_{10}(a, z_p, s_p, \zeta_p, z_c, s_c, \zeta_c; f, P) = \max_{c \geq 0, a' \geq 0} \; \log(c) + \beta \mathrm{E}\left[V_{11}(a', z_p, s_p, \zeta_p'; f', P')\right] + \beta b \left(1 + \nu_{s_p}\right) \mathrm{E}\left[V_6(0, z_c, s_c', \zeta_c'; f', P')\right] \tag{7}$$

subject to

$$a' + c = y_p(z_p, s_p, \zeta_p) + (1+r)a + \left(1 + \nu_{s_p}\right) y_c(z_c, s_c, \zeta_c) - T(z_p, s_p, \zeta_p, z_c, s_c, s_c', \zeta_c), \tag{8}$$

$f' = F(f, P)$, $P' = H(f, P)$, and $s_c' = s_c$. On the right-hand side of (7), V_{11} denotes the parent's value function in the following period, which no longer depends on the ability and schooling of children who become independent from parents. The last term on the right-hand side of (7) denotes utility that imperfectly altruistic parents derive from their children's well-being, where $b > 0$ is the altruism parameter and $V_6(0, z_c, s_c', \zeta_c'; f', P')$ is the value function of children who form new households with zero assets.

3.3. Government, Taxes and Equilibrium

The government collects tax revenue from households which it then spends on "public goods," and, in the policy counterfactuals, free secondary schooling. The government budget constraint in per capita terms is given by:

$$G + \xi \int \mathbf{I}\left(s_c' = S \wedge \tau = 9\right) df = \int T(z_p, s_p, \zeta_p, z_c, s_c, s_c', \zeta_c) df \tag{9}$$

where G is spending on public goods per capita and ξ is expenditure on free secondary education per capita. Since the paper is about public financing of secondary education, and not other public expenditures, we abstract from how G affects households or producers in the economy. When we simulate the effects of free public schooling, we assume that G remains constant, so that any schooling subsidy must be funded through per-period adjustments in the tax function T.

We focus our quantitative analysis on the balanced growth path of the economy. We relegate the full definition of recursive competitive equilibrium and the balanced growth path to Appendix B. In essence, the balanced growth path is the equilibrium in which the aggregate population level grows at a constant rate, but the relative distribution of households across individual states is constant. In this situation, household behavior does not depend on the aggregate population level. In all of our analyses, we assume the economy starts on a balanced growth path. To examine the effects of a policy change, we introduce the policy into the balanced growth path of the economy and compute transition dynamics by calculating sequences of population growth rates and prices that converge to the new balanced growth path.

3.4. School Attendance Decisions and Misallocation of Talent

We now present some examples to illustrate how the model captures misallocation of talent in education, and how the extent of misallocation can be recovered from experimental moments such as the ones discussed in the previous section.

The key decision for a household is whether or not their children attend secondary school. The benefit of attending is higher future wages, increasing in proportion to the ability level of the child. The costs of schooling are the goods cost, Ψ_S, and the opportunity cost, represented by foregone earnings. Neither of these costs depend directly on household characteristics, but, due to borrowing constraints, the utility cost ends up being higher for households with low income and assets, who have higher marginal utility of consumption. Intuitively, misallocation arises whenever a poor family chooses not to send a high ability child to secondary school even though doing so would increase lifetime household income.

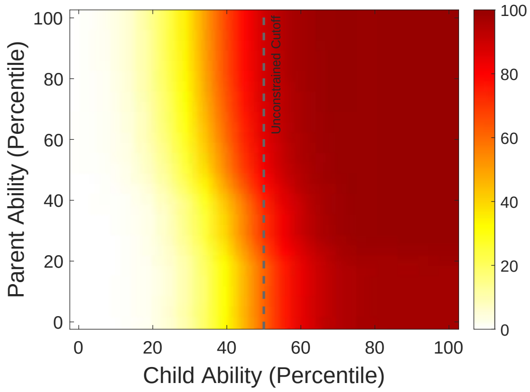
Consider two example cases of our model that vary in the goods cost of schooling, Ψ_S, the quality of secondary education, η_S, and the savings wedge, χ. The first economy, which we call the *low misallocation economy*, features a relatively low cost of schooling, relatively low schooling quality, and a small savings wedge. As we show below, borrowing constraints bind for few households in this economy, and many households choose not to send their children to secondary schooling because the returns are low. The second economy, namely the *high misallocation economy*, has a higher cost of school, higher schooling quality, and a larger savings wedge. We take other parameters to be the same across economies.[3]

Figure 2 panels (a) and (c) plot the probability that a child attends secondary school – conditional on passing the entrance exam – as a function of the child's ability and their parent's ability (a proxy for parental income and wealth). Panel (a) represents the low misallocation economy, and panel (c) represents the high misallocation case. The dotted gray line labeled *unconstrained cutoff*, marks the child

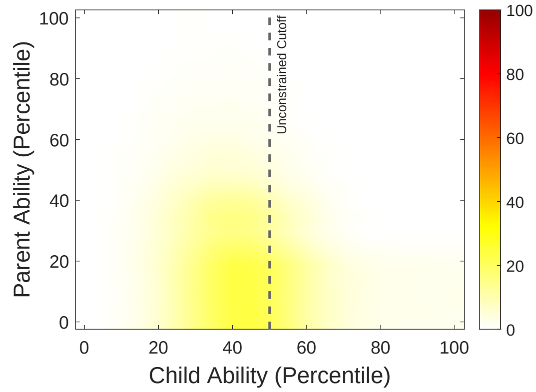
[3] In particular, the low misallocation economy features values (Ψ_S, η_S, χ) of $(5.5, 1.5, 0.1)$ while the high misallocation economy has values $(14.0, 5.0, 0.2)$. The rest of the parameter values, which are not crucial for the conclusions in this section, can be found in Appendix Table A.2.

Figure 2: Child SHS Attendance Probability

Low Misallocation

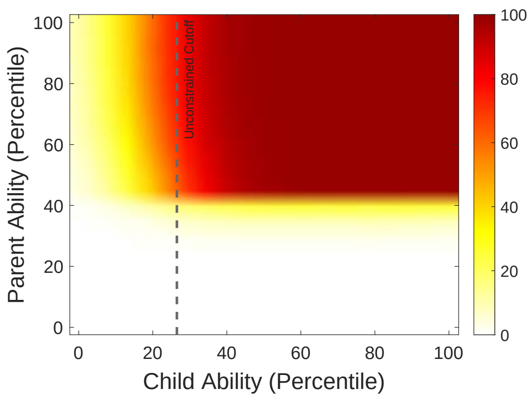


(a) SHS Attendance Probability

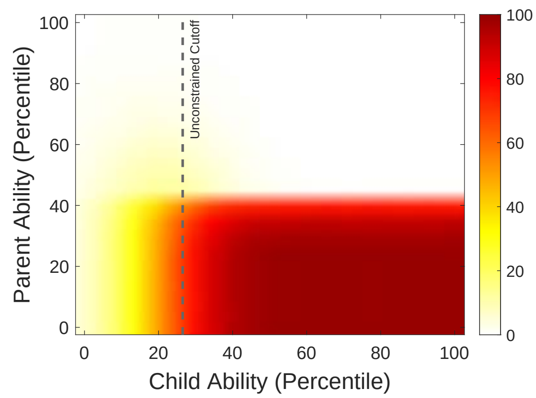


(b) Increase in Attendance from Free SHS

High Misallocation



(c) SHS Attendance Probability



(d) Increase in Attendance from Free SHS

ability level at which the net-present-value of the additional period of schooling is exactly equal to the total cost of attendance. If households faced no borrowing constraint, this is the ability level above which all children would attend schooling, and below which none would attend (assuming the average taste shock). In Panel (a), children's attendance probabilities are roughly in line with this cutoff: those above the cutoff largely attend regardless of their parent's ability level, and those below largely do not.

In Panel (c), in contrast, children born to sufficiently high ability parents attend school roughly according to the unconstrained cutoff, and children born to low-ability parents are unlikely to attend school regardless of their ability level. In this economy, there is substantial misallocation in the sense that many children for whom the net-present-value of education outweighs the costs of education do not attend. As a result, one can imagine substantial scope for gains in output from relaxing borrowing constraints.

It is important to note that aggregate data alone are not necessarily informative about the reasons why school attendance rates are so low. Although the two economies in Figure 2 are very different, they both have aggregate school attendance rates of around 30 percent. In the Panel (a) economy, attendance is low simply because the return to schooling is low on average relative to the cost, as shown by the high unconstrained cutoff. In Panel (c), however, attendance is low in large part because many high-ability children face borrowing constraints and hence forgo secondary school. This inspires our use of experimental moments, in addition to aggregate moments, in estimating the model.

Panels (b) and (d) illustrate how experimental data on the effects of scholarships can be used to distinguish between the two economies. In both panels, we simulate secondary scholarships for students above a test cutoff, as in the experiment of Duflo et al. (2021), described above. The panels plot the increase in the probability that a child attends school when experimentally offered a scholarship. Here, several differences between the two economies are apparent. First, the increase in attendance due to the scholarship is much smaller in the low misallocation economy (panel b) than the high misallocation one (panel d). Second, in the low misallocation economy, the scholarship has very little effect on the attendance of high ability children, as the vast majority of them are already attending school. Instead, only children of relatively marginal ability are induced to attend. In the high misallocation economy, in contrast, the scholarships increase school attendance across all ability levels.

These rightmost panels of Figure 2 highlight how treatment effects on schooling, both overall and by ability level, are informative about the underlying parameters of the economy that govern misallocation in education. Additionally, the experimental returns to education are larger in a high misallocation economy, since this economy has a higher school quality and since the experiment differentially affects higher ability children. The upshot of Figure 2 is that several experimental moments are informative about the overall level of misallocation in education in the economy; for this reason we focus our model estimation strategy around these moments.[4]

[4]One complication is that the experiment of Duflo et al. (2021) was necessarily conducted on a non-representative set of individuals, whereas the examples above cover the entire population. When we simulate their experiments in the following section, we select individuals from the larger

4. Model Estimation

While our estimation is largely focused on moments of the experiment described above, we first choose a handful of parameters directly, either as normalizations or to match standard values from the literature. We then estimate the rest using simulated method of moments.

4.1. Directly Chosen Moments and Aggregate Moments

We start by normalizing aggregate productivity, A, to be one, and the mean of the log ability process, μ, to be zero. We set capital's share in production, α, to be 0.33, and the discount factor to be 0.96^5, which are standard values. The international market interest rate r^* is chosen to generate a (depreciation-inclusive) user cost of capital equal to 10 percent per year.

We set the income tax function to match Ghana's statutory income tax rates at the time of the experiment, summarized in Appendix Table A.3, which focuses on a narrow tax base consisting of only the highest income earners. This specification is consistent with the overall view that taxation in the developing world is highly progressive and absent for the poorest households (see e.g. Jensen, 2022). We then set per capita government spending on public goods G such that the government budget is exactly balanced each period along the balanced growth path.

We choose the parameter governing the substitutability of skills, λ, to be 0.75, which generates an elasticity of substitution of 4. This is consistent with the long-run estimates of Bils, Kaymak, and Wu (2022) based on cross-country school attainment and wage data by attainment level. We are primarily interested in the long-run effects of schooling expansions, making a long-run elasticity of substitution appropriate for our study. We have experimented with lower values of this elasticity, down to a value of 1.4, but these do not affect substantively affect our conclusions.

We pick the standard deviation of the idiosyncratic income shock, σ_ζ, to be 0.32, which matches the values of the transitory income shock process estimated by Lagakos and Waugh (2013). In the model, this transitory component is calculated by computing the permanent component of the variance (explained further below) and subtracting this from the total variance of income.

4.2. Simulated Method of Moments

We estimate the remaining parameters of the model using the Simulated Method of Moments (SMM). There are ten such parameters, which we estimate using ten

population in the same way as in the experiment.

moments. Formally, we solve for the parameter vector:

$$\Theta = \{\nu_J, \nu_S, \eta_S, \Psi_S, b, \sigma_\epsilon, \theta, \chi, \rho, \sigma\} \tag{10}$$

that minimizes the sum of squared difference between the moments in Table 1 and their model counterparts. We also compute 95-percent confidence intervals for our parameters through bootstrapping, treating non-experimental moments (those above the line in Table 1) as fixed values, and re-sampling the experimental moments.

The first five moments we target do not use experimental variation; these are listed in the top portion of Table 1. The first is a population growth rate of 2.2 percent per year, which is the value estimated by the World Bank for Ghana. The next three targets are the secondary school completion rate in the aggregate – meaning for all individuals of school age – and the secondary school completion rates in the top and bottom test score quartiles of the control group. The final non-experimental moment we target is the standard deviation of the permanent component of wages. We measure this as $Cov(\log(w_{i,t}), \log(w_{i,t-1}))$ in the model for a panel of households taken from the balanced growth path of the model. We target a value of 0.22 from the estimate of Lagakos and Waugh (2013), which is in line with other estimates found in the literature.[5]

The remaining five moments come primarily from the experiment of Duflo et al. (2021), described in Section 2. To match these moments, we need to be able to replicate their experiment within our model. We describe how we do this in the following subsection. The last moment we target is the intergenerational correlation of schooling in Ghana, taken from Azomahou and Yitbarek (2021), and computed from regressions of children's educational attainment on parents' educational attainment. We target these regression coefficients by running these same regressions in our model.

4.3. Running the Experiment in the Model

We replicate the experiment in partial equilibrium. Since the experiment affected just 2,064 students, we find it implausible that the experiment had any significant general equilibrium effects. We also abstract away from the difference between day schools, which are the subject of the experiment, and boarding schools, which may be of higher quality, since day schools are more likely to be the focus of secondary

[5] An important caveat here is that most estimates of the transitory and permanent components of wages use data from advanced countries, where large panels of wage earners are available.

schooling expansions in the future.

Importantly, we mimic the sample selection in the experiment, which consisted of picking "smart kids from poor families." To match the requirement that students in the sample have passed the BECE, we choose a test score cutoff so that only the top 42 percent of students in the model pass, consistent with the actual BECE passing rate. Selection into the experiment also required that students had not registered for secondary school in the fall semester following their exam, which is harder to match literally within the model (particularly since one period in the model represents five years). Our strategy is to choose a parental income cutoff such that, among the students passing the BECE in the model, the eventual secondary school completion rate for those below the cutoff is 47.5 percent, just as in the control group of the experiment. We then choose the experimental sample in our model to be a subset of those with test scores *above* the test score cutoff and income *below* the income cutoff.

We treat the experiment as unanticipated, and assume that model households know that the experiment ends after a single generation. Households selected into the control group solve their optimization problem as usual. Households selected into treatment experience an exogenous reduction in the goods cost of secondary school Ψ_S to 0 for the current period and then re-optimize. We construct simulated equivalents of the experimental moments by taking simple differences of average outcomes between the treated and control households in the model, which corresponds to the intent-to-treat estimates in the experiment.

We target the negative treatment effects on fertility and positive treatment effects on human capital in the experiment, which we view as the most important findings of the experiment. The treatment effect on fertility is large, and consistent with a 10.6 percent reduction in fertility after 12 years. The experimental effects on human capital are more nuanced. On the one hand, the experiment found substantial positive impacts on test scores in reading and math of 0.16 standard deviations, which are consistent with the impacts of other successful interventions found in this literature (e.g. Duflo et al., 2012; Mbiti et al., 2019). On the other hand, the treatment effect on earnings itself is imprecisely estimated, with the 95 percent confidence interval containing wage gains anywhere between -10 percent to +15 percent. The authors also find increases in tertiary education, though many of those induced to attend college are not yet in the labor market, suggesting that the estimated treatment effect on earnings may be an underestimate of the human capital gains from free schooling.

For these reasons, we choose to target the treatment effect on test scores of 0.16

standard deviations as the experimental effect on human capital. In the model, we convert this increase in test scores to an increase in wages by assuming that a 0.16 standard deviation increase in test scores for the treatment group relative to the control group corresponds to a 0.16 standard deviation increase in wages for the treatment group. In our quantitative model, this is equivalent to wage gains of 7.6 percent. This is higher than the point estimate of 2.5 percent from the experiment, but well within their confidence interval for earnings.

As in the simple illustration of Section 3.4, the treatment effects on school attendance are informative about the extent of misallocation in education. We hence target the experiment's treatment effect on school completion, which was 27 percent. Additionally, we target the treatment effect on secondary school completion in the top quartile of test scores relative to the bottom one. This difference is small, at 4 percent, meaning that the overall treatment effect on secondary school completion was not particularly skewed toward those with high test scores relative to those with low scores. As we highlighted in Section 3.4, this turns out to be an informative moment for the extent of misallocation in education implied by our model.

4.4. Model Fit and Validation

Table 1 reports the targeted moments and their values in the estimated model. We also report the 95-percent confidence intervals for the moments that we resample in the bootstrap procedure. The fit is good for most moments, but a bit off for several of them, it must be said. On the plus side, the model does well in matching the treatment effects on human capital (6.7 percent versus 7.6 percent in the data) and fertility (-11.6 percent versus -10.6 percent in the data). The population growth rate and the variance of the permanent component of income are matched more or less exactly, and the model's treatment effect on secondary school completion is only 3 percentage points higher for the top quartile of the test score distribution than the bottom quartile, which is close to the 4 percent in the data.

The model is less successful in matching the average secondary schooling completion rates (30 percent versus 34 percent in the data), and the model's completion rates are a bit too high in the top test quartile and a bit too low in the bottom test quartile. Though overall, the model captures the slight increase in completion rates by test score quartiles in both the control and treatment groups (see Appendix Figure A.1). The treatment effect on secondary school completion is too low in the model, and the same is true of the intergenerational schooling correlation. We return to this issue later, and show that in fact missing on these two moments is not

Table 1: Targeted Moments and Model Predictions

Moments	Data	Model
Aggregate Population Growth	2.2	2.2
Aggregate SHS Completion Rate	34	30
SHS Completion, Q4 of Test (Control Group)	53	65
SHS Completion, Q1 of Test (Control Group)	41	35
Var(Permanent Component of Income)	0.22	0.22
Treatment Effect on Human Capital	7.6 $(3.2,\ 12)$	6.7
Treatment Effect on Fertility	-10.6 $(-20.8,\ -0.4)$	-11.6
Treatment Effect on SHS Completion	27.0 $(22.7,\ 31.3)$	21.3
Treatment Effect on SHS Completion, Q4 - Q1	4 $(0,\ 8)$	3
Intergenerational Schooling Correlation	0.45 $(0.43,\ 0.47)$	0.32

Note: This table reports the moments targeted in the estimation and their values in the data and in the model. The range reported below each moment in the bottom half of the Table (below the line) is its 95 percent confidence interval.

important for our main conclusions.

The estimated parameter values, and their bootstrapped confidence intervals, are presented in Table 2. While there is certainly some uncertainty in the estimated values, the confidence intervals for each parameter are fairly reasonable, suggesting that the model is precisely estimated in a statistical sense.

The estimated parameters seem reasonable from an economic sense as well. The estimated fertility parameter ν_J implies that each less-educated family has 1+1.07 = 2.07 offspring, implying an average of 4.1 children in a less-educated family with two adults. Similarly, the estimate of ν_S implies that each more-educated family has 1+0.19 = 1.19 offspring, which corresponds to around 2.4 children per family with two adults. These predictions are quite similar to (non-targeted) averages from the Demographic and Health Surveys for Ghana, which show average fertility of 4.1 children for women with junior high school only, and 2.6 children for those with secondary education or more.

Table 2: Parameter Estimates and Confidence Intervals

Parameter	Description	Estimate (Confidence Interval)
ν_J	Fertility of primary school graduates	1.07 $(1.03,\ 1.17)$
ν_S	Fertility of primary school graduates	0.19 $(0.17,\ 0.21)$
η_S	Efficiency of secondary school	5.66 $(4.39,\ 6.14)$
Ψ_S	Goods cost of secondary school	1.56 $(1.48,\ 1.71)$
b	Intergenerational altruism factor	2.26 $(2.1,\ 2.45)$
σ_ε	Std. deviation of exam score noise	0.92 $(0.89,\ 1.04)$
θ	Gumbel scale parameter of taste shock	0.42 $(0.39,\ 0.46)$
χ	Savings wedge	0.09 $(0.09,\ 0.10)$
ρ	Persistence of ability process	0.79 $(0.77,\ 0.92)$
σ	Std. deviation of ability process	0.36 $(0.34,\ 0.39)$

Note: This table reports the estimated parameters. The confidence interval is the 2.5th and 97.5th percentiles of 100 bootstrapped parameter estimates.

The estimated efficiency of schooling, η_S, is hard to interpret directly but implies (with all the other parameters) an annual return to education of 7.9 percent per year for this experimental sample. This is generally in line with other estimates of returns to education in developing countries, and if anything is on the high side. Schoellman (2012), for example, estimates returns of around 4 percent in Ghana and values generally under 5 percent for Sub-Saharan Africa (with large confidence intervals). The cost of schooling, Ψ_S, amounts to 25 percent of GDP per capita, which is close to the 21 percent reported by Duflo et al. (2021).

To better understand how plausible the estimate of b is, we compute the compensating variation of secondary schooling for all children in the model, at age 15 when their schooling decision is being made. We find that the average compensating variation is similar to the average cost of schooling (including opportunity cost), modestly lower for children who receive only a JHS education, and substantially larger for those whose parents choose a secondary education (see Appendix Figure A.2). These calculations imply that the children's valuation of schooling is mostly in line with that of their parents, suggesting that the value of b is reasonable.

The savings wedge, χ, has a value of 0.09, which implies that households save at around 11 percent per period, or 2 percent per year. This is a low return to savings but not as low as the negative returns posited by other similar incomplete-markets models estimated to data from developing countries (e.g. Lagakos, Mobarak, and Waugh, Forthcoming; Donovan, 2021).[6]

The estimated value for the intergenerational persistence of ability, ρ, is 0.79, implying a strong correlation between parents' and children's ability. This is broadly consistent with the recent conclusions of Lee and Seshadri (2019) that parental traits, summarized by ability in our model, explain a substantial amount of the variation in children's income levels. The estimated standard deviation of the ability process is 0.36. While not directly interpretable, this value (along with the other parameters) generates a Gini coefficient within the model of 0.31. This is somewhat lower than the Ghanaian value of 0.43 but well within the range of 0.3 to 0.6 reported in the World Development Indicators for other Sub-Saharan African countries.

4.5. Identification

An important question is which of the targeted moments are most informative for each of the estimated parameter values. To help answer this question, we follow Kaboski and Townsend (2011) and compute the percent change in each moment when each parameter is increased by one percent. While in general all moments jointly discipline all the parameters, some parameters correspond more closely to certain moments. For expositional purposes we present this Jacobian matrix in Appendix Table A.4, and summarize the main findings here.

The population growth parameters ν_J and ν_S are, perhaps unsurprisingly, significant determinants of the aggregate population growth rate and the treatment effects on fertility. The variance and persistence parameters of the ability process, σ and ρ, naturally increase the variance of the permanent component of income and the intergenerational schooling correlation, but also have sizable effects on many other moments in equilibrium.

[6]The average household asset-to-income ratio in the model is 0.5. This is broadly in line with other estimates from low-income countries, such as Samphantharak and Townsend (2018), who find a ratio of around 0.6 in Thai villages. Unfortunately, we know of no reliable household asset data in Ghana to which we can make a direct comparison.

The effectiveness of schooling, η_S, and the intergenerational altruism parameter, b, govern the benefits of schooling and thus result in similar changes, notably a sizable increase in aggregate secondary attendance. The key difference is that η_S increases the treatment effect on human capital while b has a minimal effect, as it only impacts the parent's valuation of better schooling. Intuitively, the cost of schooling, Ψ_S, decreases school attendance, increases the treatment effect on schooling, and consequently increases (in absolute value) the treatment effect on fertility.

Finally, the savings wedge, χ, and the variances of the test score noise and taste shocks, σ_ϵ and σ_ζ, all jointly impact secondary attendance in the top and bottom quartiles of the test score distribution as well as the difference in treatment effect between the quartiles. In fact this was the purpose of introducing these shocks into the model, and without them schooling completion and the treatment effect on schooling are always (counterfactually) much larger for those with higher test scores.

5. Simulating the Effects of Free Secondary School

Using the estimated model, we simulate the effects of a national free secondary schooling policy. We assume that households do not anticipate the policy and that the economy is on the balanced growth path at the time of implementation. The policy does not change the entrance exam score cutoff required to attend secondary school but does allow students who pass the entrance exam to attend for free. That is, the policy reduces the costs of secondary school from Ψ_S to zero.

Of course, such a policy is not truly free and must be funded by the government. We require that the government pay for the policy by raising taxes in proportion to the existing tax rates. Before the policy, each household paid taxes according to the tax function T which is a function of parent's and child's income. The post-policy tax function takes the form $(1+\tau)T$ where τ is the proportional increase in taxes. Taking this approach maintains the current structure of the labor tax schedule, and in particular the feature that the poorest half of households pay no taxes (see Appendix Table A.3).

We choose τ so that per period tax revenue along the post-policy balanced growth path is equal to per period tax revenue along the pre-policy balanced growth path plus the additional cost of the subsidy. In other words, we assume that the policy does not change per capita spending on public goods G.[7]

[7] We have experimented with alternative public finance arrangements but find that they make no substantive difference in our conclusions. For this reason we stick with the simpler assumption of period-by-period budget balance.

5.1. Quantitative Results

The general equilibrium effects of the policy are summarized in Table 3. We also report confidence intervals for each aggregate outcome using the bootstrapped parameter estimates summarized above. While this is a natural use of bootstrapped parameter estimates, it is not commonly done in G.E. counterfactual simulations in macroeconomics. The goal is simply to quantify the uncertainty in the model's counterfactual predictions arising from sampling uncertainty in the targeted moments — the experimental moments in particular as these are estimated with large standard errors.

The number of secondary schooling graduates increases by about 10 percent, under free schooling, from 30 percent of the population to 33 percent. This increase is only around a fourth of the potential increase of 12 percentage points (up to 42 percent, the percent passing the exam). The increase is also small relative to the changes in secondary school completion in the experiment, in large part because the experimental sample is highly selected relative to the general population. Fertility falls due to the schooling expansion, but this leads to a negligible decline in the population growth rates of around 0.1 percentage points. The reason fertility does not fall further is simply that schooling completion rises so little.[8]

Adult earnings increase by about 0.7 percent from the policy, stemming largely from the higher wages for the 3 percent of the population now receiving secondary education. This is offset in part by a 1.6 percent drop in the child earnings, representing the opportunity cost of the newly educated workers. Secondary-school aged children represent a smaller fraction of the population than the adults, so the net impact on wage earnings is positive.

The increase in GDP per capita from free schooling is negligible, at around 0.1 percent in the long-run. Our confidence interval excludes an increase of anything above around 0.3 percent. Thus, from the perspective of sampling uncertainty in the targeted moments, we can reject even a modest increase of say one percent of GDP. The long-run cost of the policy is 1.2 percent of GDP, implying that the program pays for only about one tenth of its cost. Capital per worker increases by less than 0.1 percent in response to higher human capital, indicating that only a small portion of the

[8] We refrain from trying to compare our model's predictions to aggregate outcomes in Ghana for the years following the inception of free SHS. In reality, numerous other policy changes went into affect around this time, making comparisons difficult. For example, the government cut taxes of various types, launched several initiatives aimed at boosting rural manufacturing, and made other changes to primary and secondary education systems.

Table 3: G.E. Effects of Free Secondary Schooling

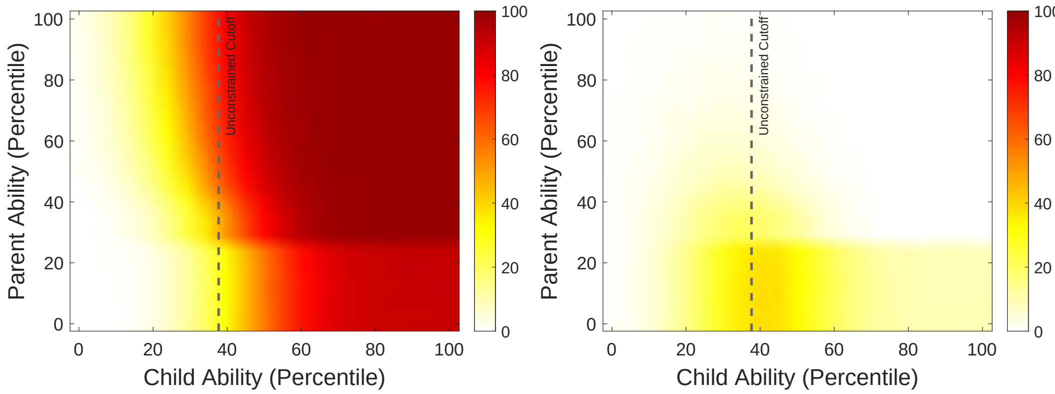
Statistic	Change Under Free Schooling
Secondary School Completion (p.p.)	3.0 (1.3, 3.6)
Population Growth Rate (p.p.)	-0.1 (-0.1, 0.0)
Adult Earnings (%)	0.7 (-0.9, 0.9)
Child Earnings (%)	-1.6 (-2.3, -0.4)
GDP per Capita (%)	0.1 (-0.8, 0.3)
Taxes per Capita (%)	1.2 (0.8, 1.3)
Skilled Wage/Unskilled Wage (%)	-2.5 (-6.0, -2.0)
Gini Coefficient	-0.01 (-0.02, -0.00)
Intergenerational Schooling Correlation	-0.04 (-0.06, -0.04)

Note: This table reports the estimated aggregate effects of free secondary schooling. The changes in secondary school completion rates and population growth rates are expressed in percentage points. The changes in the Gini coefficient and intergenerational schooling correlation are measured in levels. The changes in all other statistics are expressed in percentage changes. The range reported below each estimated value is its bootstrapped 95 percent confidence interval for the change.

GDP gains occur due to capital accumulation. Relative wages of the skilled fall by about 2.5 percent, pointing to clear distributional impacts of free schooling policies, even for those who remain unskilled after the policy change. Our predictions here are similar at least qualitatively to those of Khanna (2023), who finds substantial declines in the relative wages of skilled workers after an education expansion in India. His wage effects are larger than ours quantitatively, though his study focuses on the short run, where elasticities of substitution between low and high skilled workers are likely smaller.[9]

[9]Our analysis abstracts from several potentially important factors that are worth mentioning explicitly. Education expansions have been shown to reduce crime (e.g. Lochner and Moretti, 2004), create more informed voters, or raise the wages of others through externalities more generally (e.g. Lucas, Jr., 1988; Acemoglu and Angrist, 2000; Ciccone and Peri, 2006). We abstract from these channels largely due to a lack of evidence for our setting, though Duflo et al. (2021) found no evidence

Figure 3: Child SHS Attendance Probability in Estimated Model



(a) Attendance Probability (b) Increase in Attendance

The modest gains to GDP suggest that the estimated model may not feature high levels of misallocation in education. Recall that Figure 2 provided examples of economies with high and low misallocation by displaying the probability of SHS attendance as a function of child and parent ability. Figure 3, Panel (a), displays an identical plot using the fully estimated model. While it falls between the two extremes shown in Figure 2, it is clear that the estimated model corresponds more closely to the case with low misallocation; the highest ability children are fairly likely to attend SHS even if they are born to low ability parents. Attendance does have some dependence on parent ability; however, this effect is fairly minimal and the probability of attendance looks much closer to that of Panel (a) in Figure 2. Increases in attendance from free secondary schooling (Panel b) also more closely resemble those of the low-misallocation economy.

5.2. Welfare and Distributional Impacts

Consumption equivalent welfare calculations are performed by asking a household how much per period consumption they would be willing to give up to be indifferent between having the free schooling policy and the higher consumption. We compute the welfare equivalents along the full transition path for every individual that is alive at the time of the policy change. The distribution of welfare gains is shown in Figure 4. The dark red (leftmost) bars on the graph report the average welfare gain for the bottom quartile of the income distribution, the mean welfare

that increased school attendance altered voting behavior. Moreover, we conjecture that given our low estimated effects of free schooling policies on GDP and average wages, adding an external effect of human capital on the wages of others would be unlikely to have much additional impact.

change, and the average welfare gain for the richest quartile. On average, individual welfare increases by 2 percent of consumption.

Figure 4: Welfare Effects of Free Schooling by Parental Income

We are interested in the redistributive component of the policy; that is, how much of the welfare gains accrue to poor households relative to rich ones. As in Fernández and Rogerson (1995), rich households in our model are more likely to go to school. Thus a free SHS policy risks being regressive. Unlike Fernández and Rogerson (1995), who model schooling as funded through proportional taxation, this effect is mitigated by the fact that our tax schedule is strongly progressive. Thus the redistributive nature of the policy is a quantitative question. Examining Figure 4, we see that the welfare gains are 3.3 percent for the poorest quartile and that the richest quartile actually have small welfare losses. Thus, the policy helps the poorest households and is primarily redistributive.

To get some insights into the sources of the welfare gains, the remaining bars of Figure 4 decompose the gains into three channels: the effect of just the schooling subsidy, the effect of just the G.E. wage adjustments, and the effect of just the tax adjustment. We isolate the effect of the subsidy by offering the same free schooling to households but without requiring that it be funded, and restricting relative wages to be the same as in the estimated model. Similarly, the effect of G.E. wage adjustments represent the impact of the wage changes by themselves without any changes in the cost of schooling or taxes, and the effect of tax adjustments represents the impacts of just the tax increases, holding wages fixed and without actually using the tax increases to pay for free schooling.

The education subsidy by itself (the second of each group of four bars) are a boon for all income groups. The welfare gains for the top quartile of the income distribution are substantial at 3.3 percent. This is slightly larger than the average welfare gain of 3.2 percent, reflecting the fact that although the top quartile have higher incomes, they are also more likely to have children attend secondary schooling and collect the subsidy. The poorest quartile gains quite a bit at 4.5 percent as the subsidy represents a much larger portion of their income. Wage effects by themselves (third bar in each group) have modest effects, causing around 0.3 percent welfare gains for the poorest group and 0.7 percent welfare losses for the richest group. Higher taxes (the last bar in each group) are bad news for everyone, causing welfare losses of 1.5 percent for the poorest households and 3.3 percent for the richest quartile, who pay for most of the tax increase.

Finally, the Gini coefficient drops slightly as a result of the policy, indicating that overall inequality in the economy decreases. The correlation between parents' and children's years of schooling drops by 0.04, indicating a modest reduction in inequality across generations. Taken together, these two statistics point to an important redistributionary effect of the policy, which is the clear message from the welfare analysis as well.

5.3. Sensitivity of Conclusions to Estimation Moments

A natural question is which moments, or combinations of moments, are most influential for the model's conclusion that free secondary education leads to such small increases in GDP per capita. To answer this question we conduct an analysis similar in spirit to the one developed by Andrews, Gentzkow, and Shapiro (2017), where we re-estimate the model a number of different times under alternative assumptions about the values of the targeted moments. For each re-estimated version of the model, we then simulate the effects of free secondary schooling, in general equilibrium, just as in the main experiment above.

The results are summarized in Table 4. The first row reproduces the predictions of the main estimation, to serve as a comparison. The second row, labeled (i), reports

Table 4: Sensitivity of GE Counterfactuals to Estimation Moments

	GDP Gain (%)	Welfare Gain (%, Avg)	SHS Increase (pp)
Estimated model (Confidence Interval)	0.1 $(-0.9, 0.3)$	2.0	3.0 $(1.3, 3.6)$
(i) Treatment effect on human capital x 1.5	0.2	2.1	3.2
(ii) Treatment effect on fertility x 1.5	-0.1	2.5	2.4
(iii) Treatment effect on SHS completion x 1.5	-0.1	2.3	2.4
(iv) Treatment effect on SHS, Q4-Q1 x -1	0.3	2.4	3.2
(i), (ii), (iii) and (iv) combined	0.5	3.6	5.1

Note: This matrix displays the sensitivity of the model's predictions for the general equilibrium effects of free secondary schooling to alternative values of estimation moments. See the text for details. The next reports the effect when multiplying the difference in the treatment effects between Q4 and Q1 by -1. The last row shows the effects of changing all of the above.

the effects of free secondary schooling when the model is re-estimated targeting a fifty percent larger treatment effect on human capital. Specifically, this means targeting an 11.4 percent effect rather than the 7.6 percent used in the main estimation, keeping all other estimation targets the same. This alternative estimation produces a GDP gain of 0.2 percent, around double that of the estimated model, and marginally larger increases in welfare and secondary schooling. So higher values of this target do increase the importance of free schooling, but do not substantively affect the conclusions of the model. Free schooling is still not doing much to increase average living standards with this single higher target.

Row (ii) reports the results of re-estimating the model to target a fifty percent larger treatment effect on fertility, meaning a decrease in fertility of 15.9 percent rather than 11.4 percent, again with all other targets held fixed. The effect on GDP is now slightly negative, at -0.1 percent, with a slightly smaller impact on secondary completion rates. A very similar result comes in row (iii), which targets a fifty percent higher treatment effect on schooling attainment (i.e. 40.5 percent rather than 27 percent). These similar conclusions come about from different changes in parameter estimates (as we show in Appendix Table A.5), but give the same basic conclusion as the main estimation, and show that the results about free schooling are not particularly sensitive to changes in any one of these three key targets.

Row (iv) reports what happens when we re-estimate the model to target a treatment effect on schooling for students in Q4 of the test score distribution that is 4 percentage points *lower* than those in Q1, rather than 4 points higher as in the main estimation. This single change leads the model to predict a 0.3 percent in increase in GDP per capita, or around triple the baseline results. This alternative estimation infers a somewhat higher impact of free schooling mostly because it results in a substantially higher estimated cost of schooling in order to match this increase in treatment effect for the lowest ability students, who are more likely to come from poor households and thus are disproportionately impacted by an increase in Ψ_S.

The last row of Table 4 shows the effects of free schooling when we re-estimate the model to match all four of the previous changes simultaneously. The combined effect of these changes now leads to a 0.5 percent increase in GDP per capita, or five times the baseline effects, and a 3.6 percent increase in average welfare. These effects, which are certainly less pessimistic than the main results, come about from substantial movements in parameter estimates (displayed in Appendix Table A.5). This estimation features a higher ν_J and lower ν_S, which implies the larger effects on fertility, and substantially larger estimated values of η_S and Ψ_S. The result of these combined parameter changes is a model with more misallocated young people, whose parents cannot afford more expensive and higher quality schools (relative to the main estimation).

The upshot of this analysis is that while the model is not particularly sensitive to single changes in estimation targets, it is indeed sensitive to at least certain combinations of different targets. When the treatment effects on human capital, fertility and schooling completion are higher, and the latter particularly so for those with lower test scores, the model will predict higher GDP and welfare gains from publicly funded, merit-based schooling.

5.4. Could the Model Have Predicted Larger Effects?

Given the sensitivity results above, are there any parameter values that could have led to large impacts of free secondary schooling? Or is the finding of small effects simply hard-wired into the structure of the model regardless of parameter values? To address these questions, Table 5 reports the model's aggregate predictions when we make several large changes in parameters in a way that attempts to greatly increase misallocation of talent. In short, we do this by increasing the effectiveness of schooling, η_S, the cost of schooling, Ψ_S, and the savings wedge, χ. Unlike in the

Table 5: Effects of Free Schooling Under Alternative Parameter Choices

	GDP Gain (%)	Welfare Gain (%, Avg)	T.E. SHS Compl. (%)	T.E. Human Capital (%)	T.E. SHS Q4 - Q1 (%)
Data	-	-	27	8	4
(Confidence Interval)			$(23, 31)$	$(3, 12)$	$(0, 8)$
Estimated model	0.1	2	21	7	3
+ Schooling quality $\times$ 2.5	1.1	8	9	2	-11
+ Schooling cost $\times$ 3.5	7.9	24	44	26	-19
+ Savings wedge $\times$ 2	10.5	22	50	38	-17

Note: This table reports the gains in GDP and average welfare from free secondary schooling under alternative parameter choices. The last three columns report moments of the data and model (the treatment effect on SHS graduation, the treatment effect on human capital, and the treatment effect on SHS for the top quartile of test scores minus the bottom) under the benchmark estimated parameters and three alternative parameter choices. The first of these increases η_S by a factor 2.5 and keeps all other parameter values the same as in the estimated model. The second also increases Ψ_S by a factor 3.5, in addition to the factor 2.5 increase in η_S, and keeps all other parameter values the same as in the main estimation. The last one additionally increases the savings wedge, χ, by a factor 2.

previous subsection, we make no attempt to actually match the estimation targets, but simply report their values for the alternative parameters.

The first row of Table 5 reports the values of certain select moments in the data, and the second row simply reproduces the predictions of the main estimation. In the third, we report what happens when we increase the schooling quality, η_S by 2.5, keeping all other parameters equal. In this case, the higher-quality schooling leads to a larger GDP gain of 1.1 percent, and larger welfare gains of 8 percent. Notably, the treatment effects on schooling and human capital are smaller in this specification, despite the gains to GDP being an order of magnitude larger, highlighting the fact that there is not a simple correspondence between treatment effects and aggregate outcomes.

The fourth row of Table 5 presents the model's predictions for free schooling when the cost of schooling, Ψ_S, is multiplied by 3.5 (in addition to the higher productivity η_S, as before). In this case, the GDP and welfare gains are much larger at 7.9 percent and 24 percent, respectively. Intuitively, this model predicts much higher GDP and welfare gains from free schooling due to credit constraints proving much more binding for students who have a larger potential gain from schooling than before.

The bottom row of the table presents the model's predictions when the savings wedge χ is doubled, in addition to the higher values of η_S and Ψ_S. This puts more households up against credit constraints, since it makes it even harder to save. This model predicts a 10.5 percent increase in GDP from free schooling, accompanied by a 22 percent increase in average welfare. The large positive effects of free schooling in this version of the model stem from getting a large number of very credit constrained children into high quality schooling, which leads to substantial increases in average wages for all of them. But this model's prediction for the treatment effect on SHS graduation is about twice as high as actually observed (50 percent versus 27 percent in the data), and the treatment effect on human capital is about five times as high as in the data (38 percent versus 8 percent). Further, the selection on which students complete secondary schooling (by test score) is counterfactually negative. Comparing these moments to their true counterparts (in the first row) shows that this model's predictions are far outside their respective confidence intervals.[10]

The lesson of this subsection is that the model is able to deliver much larger impacts of free schooling on GDP and welfare with only a simple set of changes to parameter values. A ten percent increase in GDP per capita from pulling one single policy lever would be substantial indeed, if poor countries could in fact pull it. Yet the parameter changes implying such large increases in GDP per capita lead to strongly counterfactual predictions for key experimental moments. In other words, the conclusion of small GDP and welfare impacts of free schooling is not an artifact of the model structure per se, but a feature of the model once estimated to the targeted moments in question.[11]

6. Aggregate Effects of Alternative Policies

Are there any alternative policy levers that governments in low income countries can pull to bolster their education systems and raise their average income levels? Or

[10] Appendix Figure A.3 plots the probability of SHS attendance as a function of child and parent ability, equivalent to Figure 2 which compares these probabilities for example parameterizations featuring low and high misallocation. Examining the figure, it is clear that these alternative parameters represent an economy where misallocation is very high.

[11] In a similar vein, we address the concern that the estimated model modestly under-predicts the intergenerational schooling correlation by increasing ρ by 20 percent to 0.95. This parameterization results in a 2 percent GDP *loss* from the free schooling policy. The reason is that a higher ρ means less misallocation in the estimated model, as fewer high ability children are born to poor parents.

do they all lead to small changes in GDP per capita by virtue of being small? We use the estimated model to simulate some alternative policies in order to address these questions. Doing so also helps shed additional light on why the model predicts such modest impacts from free secondary schooling.

Table 6: Aggregate Effects of Alternative Policies

	GDP Gain (%)	Welfare Gain (%, Avg)	SHS Increase (pp)
Free Secondary School	0.1	2.0	3.0
Free Secondary School + Lower Test Cutoff	0.0	4.8	6.4
Universal Basic Income (costing same amount)	0.0	0.5	0.0
Raise Schooling Quality	4.1	4.6	3.1

Note: This table reports the gains in GDP, the gains in C.E. welfare and the increase in the SHS graduation rate under free schooling and several alternative policies (described in the text).

Table 6 summarizes the aggregate effects of various alternative policy counterfactuals. The first row reproduces three key aggregate statistics from the free schooling policy counterfactual from the previous section: the gains in GDP, the increase in average welfare, and the increase in secondary schooling completion. The second row reports the same outcomes from an alternative simulation where free secondary schooling is offered alongside a reduction in the test score cutoff allowing 62 percent of students to pass, up from 42 percent in the current system. This policy leads to no change in GDP per capita, but a more substantial 4.8 percent increase in average welfare. The increase in secondary graduates is roughly twice what it was under the baseline policy, at 6.4 percentage points, providing education for more households with more marginal abilities but strong tastes for schooling.

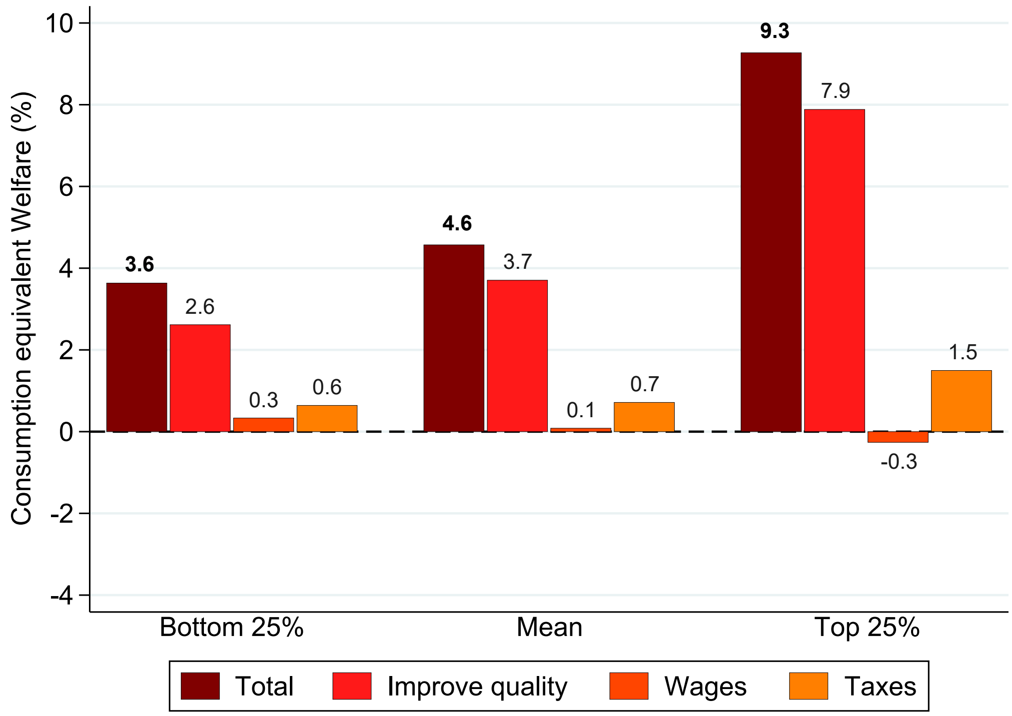
The third row simulates a simple universal basic income policy costing the same amount as the free schooling policy in the main experiment. We simulate universal basic income by simply increasing the tax rates proportionally, as before, but now redistributing the proceeds evenly to all households. The result is essentially no change in GDP and a 0.5 percent increase in average welfare. The welfare gains here stem purely from transferring consumption from those with low marginal utility to those with high marginal utility. Secondary high school completion rates are basically unchanged. Thus, pure redistribution is a substantial force for raising average welfare in our estimated model, accounting for around one-quarter of the welfare gains we predict from free schooling.

The last row of Table 6 summarizes the effects of improving schooling quality in such a way that average test scores rise by 0.1 standard deviations (compared to the benchmark estimation). This effect is conservative relative to the average effect size estimated in a number of different randomized interventions aimed at improving schooling quality in the developing world, many of which find effects of around 0.2 standard deviations or higher. One such intervention is to offer financial incentives to teachers based on the test scores of their students. Muralidharan and Sundararaman (2011) and Duflo et al. (2012) found that this raised test scores in India for example, while Mbiti et al. (2019) found effects of a similar size for teacher incentives plus block grants for schools in Kenya. Another successful schooling quality intervention is to increase the number of teachers in the classroom, as in the studies of Banerjee et al. (2007) and Muralidharan and Sundararaman (2013) in India. For our simulated intervention, we use the policy cost from Mbiti et al. (2019) who report that the cost of increasing test scores by 0.1σ per student was US $5.78.

Our model implies larger effects on GDP and welfare of improving schooling quality than providing free schooling. GDP rises by 4.1 percent and welfare increases by 4.6 percent under such an intervention. Even though this policy has no provisions aimed at expanding secondary enrollment directly, improved schooling quality raises school enrollments by 3.1 percentage points, almost exactly the same increase as in the free schooling policy. The implication is that many students were not attending secondary schooling to begin with because they felt the returns were not high enough to justify the costs (including opportunity cost).

Figure 5 plots the welfare gains from improving school quality across the income distribution. The dark red bars show that welfare gains are higher everywhere under schooling quality improvements than under free secondary school (i.e. those displayed in Figure 4). The largest welfare gains are for the richest quartile of the Ghanaian income distribution, since this group is most likely to have kids in secondary school already; however, the increase in welfare for the poorest households is still larger than under the free secondary school policy. The remaining bars break down the welfare gains into the pure effects of schooling quality improvements, relative wage effects, and taxes. One can see that the pure gains from quality improvements drive the welfare gains for each income group. Wage effects are modest, and tax effects are even positive now. The reason is that this policy raises human

Figure 5: Welfare Effects of Improving School Quality by Parental Income



capital enough so that tax rates can be lowered and still have enough funds to cover the cost of the schooling quality improvements.

In other words, schooling quality improvements with similar costs and benefits to the micro studies cited above pay for themselves in the long run, unlike free secondary schooling. The main factor accounting for why policies aimed at raising education quality perform better than free schooling policies is that free schooling ends up affecting just 3 percent of new secondary graduates, compared to the 33 percent of inframarginal students already attending school plus new attendees affected by the schooling quality improvements.

7. Support for Free Secondary School: Survey Evidence

The welfare results of Figure 4 predict that free secondary schooling in Ghana is progressive, with the poorest households gaining much more than the richest ones. Progressivity is not a necessary outcome of the model, as richer households are more likely to attend secondary schooling than poorer ones. Nevertheless, the higher tax burden for richer households in the estimated model outweighs the higher chances of attending secondary school for free, leaving them worse off in welfare terms.

In this section we provide an additional validation of the model by testing its prediction that free secondary schooling is better for poorer households than for richer ones. We do this by conducting a survey of Ghanaian households to gauge their attitudes towards the free secondary schooling policy that was enacted in 2017. Our survey covers a nationally representative sample of 3,500 households interviewed in August and September of 2022. Details about the survey and sample selection procedure can be found in Appendix C. From each household, an adult was interviewed and asked a variety of questions about which government programs and taxes should be scaled back or expanded. Relevant for our purpose, respondents were given a list of eight government expenditure items, including the recently implemented free SHS program, and asked whether they thought each item should be abolished, postponed, cut substantially, cut somewhat, maintained, or expanded. The answers to each of these questions allow us to examine how support for free secondary schooling varies as a function of demographic characteristics. In particular, we look at variation with respect to education, a proxy for permanent income.[12]

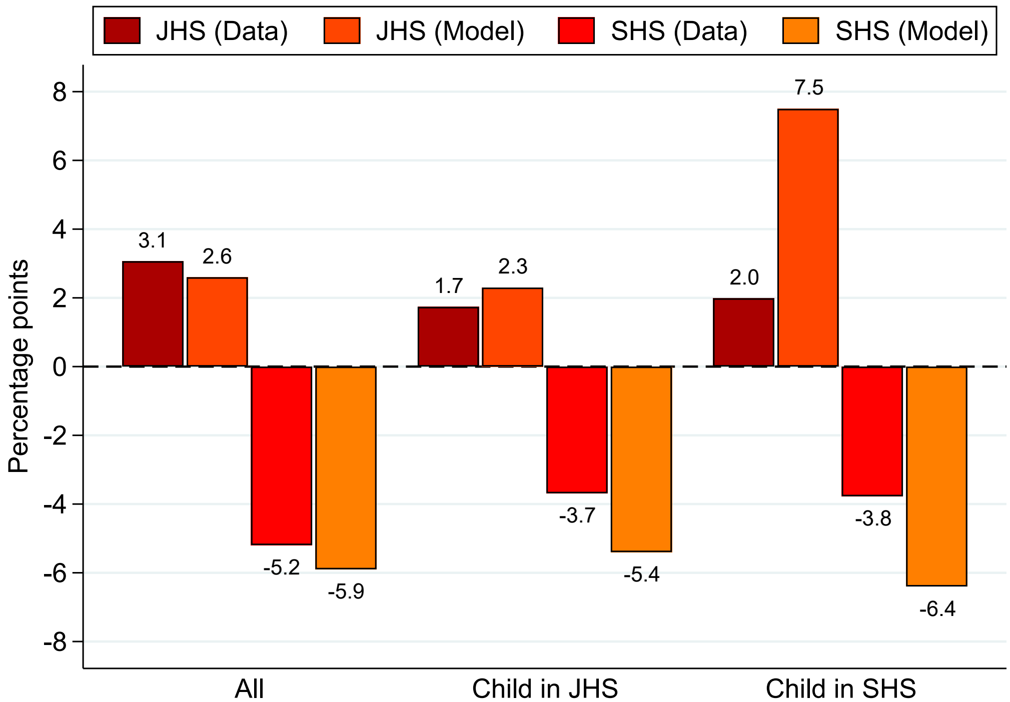
In order to compare between our survey and the model, we construct an analogous measure of support within the model. Starting from the stationary distribution of households on the pre-policy balanced growth path, we take a representative sample of households and, for each household, compute the change in welfare the household would experience if the free SHS policy were implemented.[13] Any household whose change in welfare is (weakly) positive is classified as a supporter (i.e. someone who would select "maintained" or "expanded" when asked about free SHS). Those who experience a loss in welfare are classified as non-supporters.

Figure 6 displays how support for the free SHS policy varies with respect to parental education level (different colored bars) and whether or not the household has a child enrolled in JHS and SHS. Each bar represents the difference in support between the represented group and the average level of support. Respondents with at most a JHS level of education were 3.1 percentage points more likely to support

[12]We find a wide range of support on average for different public expenditures, suggesting that household responses are informative about their viewpoints on different expenditure categories rather than simply a referendum on the current government in power. For example investments in electricity have an average support of 95 percent while salary increases for public servants has an average support of just 13 percent.

[13]We exclude any model households above the age of 60 as, being retired and having already experienced the impulse utility from their children, they are completely isolated from the rest of the economy and experience no welfare change. We do the same for our survey when comparing to the model.

Figure 6: Support for Free SHS Policy (Dev. from Average), by Parent Education



Note: This figure plots the relative support, defined as the deviation from the average level of support, for free SHS as a function of parental education (different bars) and whether or not the parent has a child enrolled in JHS or SHS (different sets of bars) measuring using our survey as well as the corresponding measure in the model.

the policy than average, while respondents with an SHS education or more were 5.2 percentage points less likely. These results are remarkably close to the model values of 2.6 and 5.9 respectively. We interpret this as validation of the quantitative prediction that gains from government-funded free SHS largely accrue to households with low education and low incomes.

The remaining two sets of bars display the levels of support for households that have a child enrolled in JHS or SHS, as these households may be particularly sensitive to schooling policy. Both the survey responses and the model respondents tell largely the same story as the aggregate result, namely that support for free SHS is stronger among those with low education. Those with a JHS level of education continue to be more likely to support the policy than those with SHS education, but the difference between the two in the data is smaller than the same comparison among all households. While the model results are fairly close to the data for the case of households with children in JHS, it over-predicts the support of JHS educated households with children in SHS.[14]

Overall, our survey results support the model's quantitative prediction that free SHS operates as a redistributive policy. Both the model and survey suggest that poorer and less-educated households benefit more from free schooling than richer households. This is notable as it stands somewhat in contrast to notions in the United States that funding for higher education is largely captured by the wealthiest households and can be regressive (see e.g. Fernández and Rogerson, 1995; Ansell, 2010; Catherine and Yannelis, 2023).

8. Conclusions

One of the main reasons income per capita is so low in the developing world is that human capital levels are so low (Manuelli and Seshadri, 2014; Hendricks and Schoellman, 2018). One of the potential paths these countries can take to raise human capital is to increase attendance levels in secondary school. Making secondary schooling free for students, and funding the costs through higher taxes, is a natural option to consider. Not surprisingly, many developing countries are currently considering or implementing free schooling policies of some kind.

In this paper we analyze the aggregate and distributional effects of free secondary schooling policies in the developing world, looking through the lens of an OLG model of human capital accumulation with credit constraints. We focus on the case of Ghana, for which we can draw on recent experimental evidence on the outcomes of students randomly assigned to receive free secondary schooling, leading to higher secondary school completion rates and higher average test scores (Duflo et al., 2021). Ghana is also a country that has recently adopted free secondary schooling, and the policy is viewed as a success there and in other developing countries (Center for Global Development, 2022).

Our conclusions are less optimistic. When we simulate the general equilibrium effects of free secondary school in our model, we find that it would have next to no impact on GDP per capita and increase welfare by only around 2 percent. The reason for these modest aggregate effects is that when estimated to match the experimental

[14] Appendix Table A.6 shows that these patterns of support for free SHS are present under a variety of sample restrictions, including only urban or rural respondents, and only male or female respondents, and when restricting the sample to only respondents from the Ashanti and Volta regions of Ghana, which are the regions with the most and least support for the current ruling party. Although the level of support changes between these regions, once again the pattern that JHS educated respondents are more supportive of the policy remains.

data, our model implies that most students eligible for secondary school, but not attending, choose not to attend due to low potential returns and high opportunity costs. The model implies that some students are misallocated, but credit constraints are not the main reason secondary enrollment rates are not higher.

We conclude that free secondary education policies are mostly redistributive in nature, rather than a path to economic growth, at least at current low levels of schooling quality. Spending the same amount on improving quality would lead to substantially higher GDP per capita and welfare across the income distribution. Improving schooling quality would also expand enrollments by around the same amount as free schooling policies. Thus, our analysis suggests that human capital levels would rise more in poor countries by raising the quality of existing schools than by giving away a mediocre education to more young people.

References

ABBOTT, B., G. GALLIPOLI, C. MEGHIR, AND G. L. VIOLANTE (2019): "Education Policy and Intergenerational Transfers in Equilibrium," *Journal of Political Economy*, 127, 2569–2624.

ACEMOGLU, D. AND J. D. ANGRIST (2000): "How Large are Human Capital Externalities? Evidence from Compulsory Schooling Laws," *NBER Macroeconomics Annual 2000*, 9–59.

AIYAGARI, S. R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, 109, 659–84.

AKCIGIT, U., H. ALP, AND M. PETERS (2021): "Lack of Selection and Limits to Delegation: Firm Dynamics in Developing Countries," *American Economic Review*, 111, 231–75.

AKCIGIT, U., J. G. PEARCE, AND M. PRATO (2020): "Tapping into Talent: Coupling Education and Innovation Policies for Economic Growth," NBER Working Paper 27862.

ANDREWS, I., M. GENTZKOW, AND J. M. SHAPIRO (2017): "Measuring the Sensitivity of Parameter Estimates to Estimation Moments," *Quarterly Journal of Economics*, 132, 1553–1592.

ANSELL, B. W. (2010): *From the Ballot to the Blackboard: The Redistributive Political Economy of Education*, Cambridge University Press.

AZOMAHOU, T. T. AND E. YITBAREK (2021): "Intergenerational Mobility in Education: Is Africa Different?" *Contemporary Economic Policy*, 39, 503–523.

BANERJEE, A. V., S. COLE, E. DUFLO, AND L. LINDEN (2007): "Remedying Education: Evidence from Two Randomized Experiments in India," *Quarterly Journal of Economics*, 122, 1235–1264.

BANERJEE, A. V. AND A. F. NEWMAN (1993): "Occupational Choice and the Process of Development," *Journal of Political Economy*, 101, 274–98.

BASSI, V., R. MUOIO, T. PORZIO, R. SEN, AND E. TUGUME (2022): "Achieving Scale Collectively," *Econometrica*, 90, 2937–2978.

Bénabou, R. (2002): "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica*, 70, 481–517.

Bewley, T. (1977): "The Permanent Income Hypothesis: A Theoretical Formulation," *Journal of Economic Theory*, 16, 252–292.

Bils, M., B. Kaymak, and K.-J. Wu (2022): "Labor Substitutability among Schooling Groups," NBER Working Paper 29895.

Bils, M. and P. J. Klenow (2000): "Does Schooling Cause Growth?" *American Economic Review*, 90, 1160–83.

Brooks, W. and K. Donovan (2020): "Eliminating Uncertainty in Market Access: The Impact of New Bridges in Rural Nicaragua," *Econometrica*, 88, 1965–1997.

Brudevold-Newman, A. (2021): "Expanding Access to Secondary Education: Evidence from a Fee Reduction and Capacity Expansion Policy in Kenya," *Economics of Education Review*, 83.

Buera, F. J., J. P. Kaboski, R. Rogerson, and J. I. Vizcaino (2022): "Skill-Biased Structural Change," *Review of Economic Studies*, 89, 592–625.

Buera, F. J., J. P. Kaboski, and Y. Shin (2021): "The Macroeconomics of Microfinance," *Review of Economic Studies*, 88, 126–161.

Buera, F. J., J. P. Kaboski, and R. M. Townsend (Forthcoming): "From Micro to Macro Development," *Journal of Economic Literature*.

Catherine, S. and C. Yannelis (2023): "The Distributional Effects of Student Loan Forgiveness," *Journal of Financial Economics*, 147, 297–316.

Caunedo, J. and N. Kala (2022): "Mechanizing Agriculture," Unpublished Working Paper, MIT.

Celik, M. A. (2023): "Does the Cream Always Rise to the Top? The Misallocation of Talent in Innovation," *Journal of Monetary Economics*, 133, 105–128.

Center for Global Development (2022): *Schooling for All: Feasible Strategies to Achieve Universal Education*, Center for Global Development.

CICCONE, A. AND G. PERI (2006): "Identifying Human-Capital Externalities: Theory with Applications," *Review of Economic Studies*, 73, 381–412.

CUNHA, F. AND J. HECKMAN (2007): "The Technology of Skill Formation," *American Economic Review Papers and Proceedings*, 97, 31–47.

DARUICH, D. (2020): "The Macroeconomic Consequences of Early Childhood Development Policies," Unpublished Working Paper, University of Southern California.

DOEPKE, M. AND F. ZILIBOTTI (2017): "Parenting with Style: Altruism and Paternalism in Intergenerational Preference Transmission," *Econometrica*, 85, 1331–1371.

DONOVAN, K. (2021): "The Equilibrium Impact of Agricultural Risk on Intermediate Inputs and Aggregate Productivity," *Review of Economics Studies*, 88, 2275–2307.

DUFLO, E., P. DUPAS, AND M. KREMER (2021): "The Impact of Free Secondary Education: Experimental Evidence from Ghana," Unpublished Working Paper, Stanford University.

DUFLO, E., R. HANNA, AND S. P. RYAN (2012): "Incentives Work: Getting Teachers to Come to School," *American Economic Review*, 102, 1241–1278.

FERNÁNDEZ, R. AND R. ROGERSON (1995): "On the Political Economy of Education Subsidies," *Review of Economic Studies*, 62, 249–62.

——— (1997): "Keeping People Out: Income Distribution, Zoning, and the Quality of Public Education," *International Economic Review*, 38, 23–42.

FILMER, D. AND N. SCHADY (2014): "The Medium-Term Effects of Scholarships in a Low-Income Country," *Journal of Human Resources*, 49, 663–694.

GALOR, O. AND J. ZEIRA (1993): "Income Distribution and Macroeconomics," *Review of Economic Studies*, 60, 35–52.

HALL, R. E. AND C. I. JONES (1999): "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics*, 114, 83–116.

HANUSHEK, E. A. AND L. WOESSMANN (2007): *Education Quality and Economic Growth*, World Bank.

HASSLER, J., J. V. RODRÍGUEZ MORA, AND J. ZEIRA (2007): "Inequality and Mobility," *Journal of Economic Growth*, 12, 235–259.

HENDRICKS, L. AND T. SCHOELLMAN (2018): "Human Capital and Development Accounting: New Evidence from Wage Gains at Migration," *Quarterly Journal of Economics*, 133, 665–700.

HSIEH, C.-T., E. HURST, C. I. JONES, AND P. J. KLENOW (2019): "The Allocation of Talent and U.S. Economic Growth," *Econometrica*, 87, 1439–1474.

HUGGETT, M. (1993): "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies," *Journal of Economic Dynamics and Control*, 17, 953–969.

JENSEN, A. (2022): "Employment Structure and the Rise of the Modern Tax System," *American Economic Review*, 112, 213–32.

KABOSKI, J. P. AND R. M. TOWNSEND (2011): "A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative," *Econometrica*, 79, 1357–1406.

KHANNA, G. (2023): "Large-Scale Education Reform in General Equilibrium: Regression Discontinuity Evidence from India," *Journal of Political Economy*, 131.

LAGAKOS, D., A. M. MOBARAK, AND M. E. WAUGH (Forthcoming): "The Welfare Effects of Encouraging Rural-Urban Migration," *Econometrica*.

LAGAKOS, D. AND M. E. WAUGH (2013): "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review*, 103, 948–980.

LAITNER, J. (1997): "Intergenerational and Interhousehold Economic Links," *Handbook of Population and Family Economics*, 1, 189–238.

LEE, S. Y. AND A. SESHADRI (2019): "On the Intergenerational Transmission of Economic Status," *Journal of Political Economy*, 127, 855?921.

LOCHNER, L. AND A. MONGE-NARANJO (2011): "The Nature of Credit Constraints and Human Capital," *American Economic Review*, 101, 2487–2529.

LOCHNER, L. AND E. MORETTI (2004): "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports," *American Economic Review*, 94, 155–189.

LUCAS, JR., R. E. (1988): "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22, 3–42.

MANUELLI, R. E. AND A. SESHADRI (2014): "Human Capital and the Wealth of Nations," *American Economic Review*, 104, 2736–62.

MBITI, I., K. MURALIDHARAN, M. ROMERO, Y. SCHIPPER, C. MANDA, AND R. RAJANI (2019): "Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania," *Quarterly Journal of Economics*, 134, 1627–1673.

MIDRIGAN, V. AND D. XU (2014): "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, 104, 422–458.

MOLL, B. (2014): "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?" *American Economic Review*, 104, 3186–3221.

MURALIDHARAN, K. AND V. SUNDARARAMAN (2011): "Teacher Performance Pay: Experimental Evidence from India," *Journal of Political Economy*, 119, 39–77.

——— (2013): "Contract Teachers: Experimental Evidence from India," NBER Working Paper 19440.

PORZIO, T., F. ROSSI, AND G. SANTANGELO (2022): "The Human Side of Structural Transformation," *American Economic Review*, 112, 2774–2814.

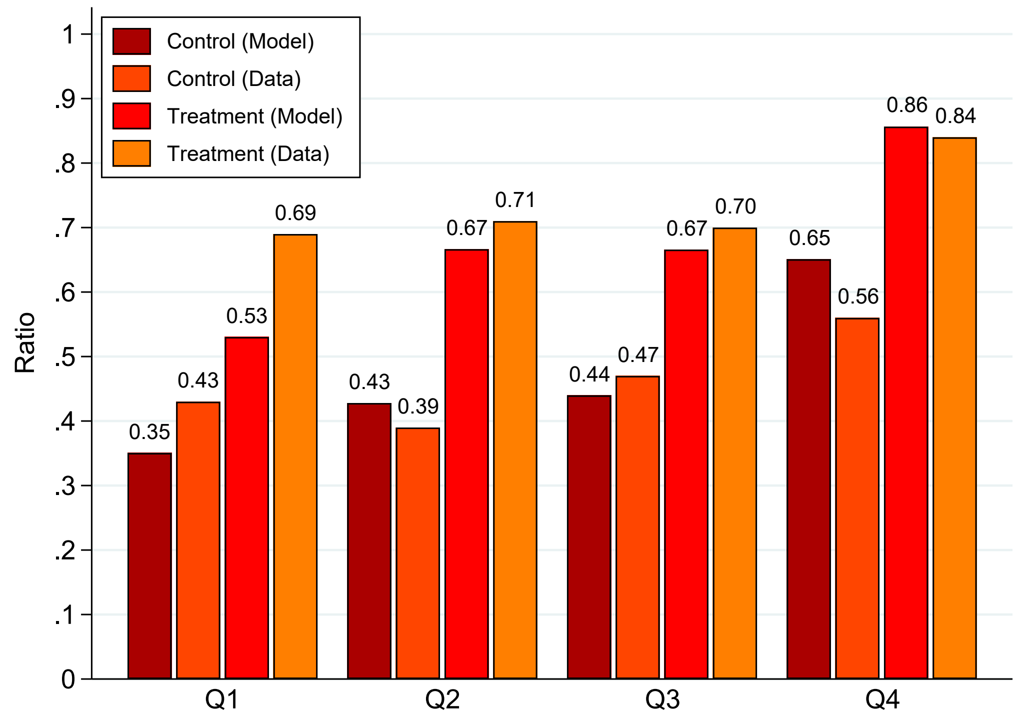
SAMPHANTHARAK, K. AND R. M. TOWNSEND (2018): "Risk and Return in Village Economies," *American Economic Journal: Microeconomics*, 10, 1–40.

SCHOELLMAN, T. (2012): "Education Quality and Development Accounting," *Review of Economic Studies*, 79, 388–417.

Appendix (for Online Publication)

A. Appendix Figures and Tables

Figure A.1: SHS Completion by Quartile of Test Score: Data vs Model



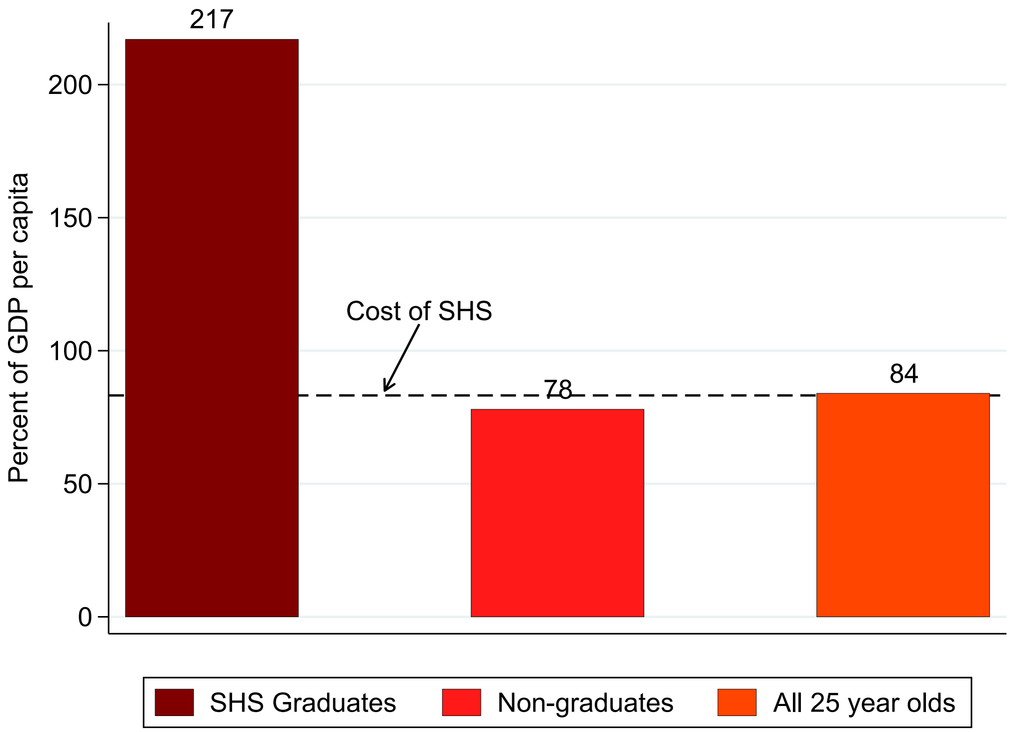
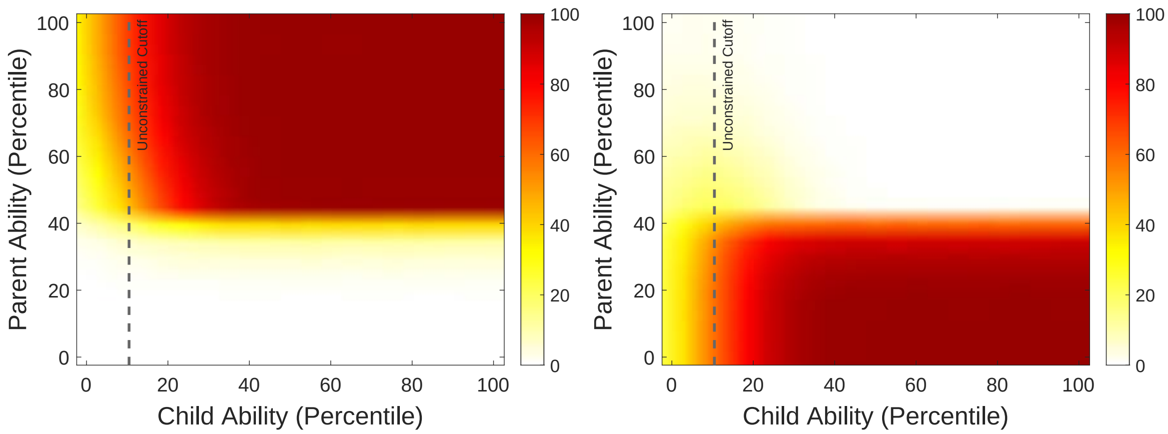


Figure A.2: Compensating Variation of Secondary School to Children



(a) Attendance Probability

(b) Increase in Attendance

Figure A.3: Child SHS Attendance Probability (Counterfactual Parameters)

Table A.1: Free Secondary Schooling Policies in Developing Countries

Country	Year	Requirement
Benin	2007	Pass Brevet d'Etudes du Premier Cycle
Gambia	2015	Pass Basic Education Certificate Exam
Ghana	2017	Pass Basic Education Certificate Exam
Kenya	2008	Pass Certificate of Primary Education Exam
Malawi	2019	Pass Primary School Leaving Certificate Exam
Mauritius	2016	Pass General Certificate of Education Exam
Nepal	2018	Pass final district-level exam
Philippines	1988	Do not fail in two consecutive years
Rwanda	2012	Score $\geq$ 'High' on O-level Test
Sierra Leone	2018	Score $\geq$ 6 on Basic Education Certificate Exam
Tanzania	2015	Pass Standard 7 Exam
Uganda	2007	Score $\geq$ 28 in Primary School Leaving Exam
Zambia	2022	Pass Baccalaureate Exam

Note: This table reports the year that each country adopted a free secondary schooling policy and the merit requirement to attend secondary schooling.

Table A.2: Parameter Values Used in Discussion

Description	Parameter	Misallocation Low	Misallocation High
Fertility of Primary School Graduates	ν_J	1	
Fertility of Secondary School Graduates	ν_S	0.2	
Intergenerational Altruism Factor	b	2.5	
Std. Dev. of Exam Score Noise	σ_ε	1	
Gumbel Scale Parameter of Taste Shock	θ	0.5	
Persistence of Ability Process	ρ	0.8	
Std. Dev. of Ability Process	σ	0.5	
Std. Dev. of Idiosyncratic Income Shock	σ_ζ	0.3	
Gains from Secondary School	η_S	5.5	14.0
Goods cost of Secondary School	Ψ_S	1.5	5.0
Savings Wedge	χ	0.1	0.2

Note: This table lists the parameter values used for creating the figures in subsection 3.4. The two versions of the model share many parameters and differing only on three key parameters. Parameters not listed here take the value given by Table 2 for both models.

Table A.3: Labor Income Tax Schedule in Ghana

Income	Tax Rates
First 1,008 GHC (=up to 42% of GDP p.c.)	0%
Next 240 GHC (=up to 52% of GDP p.c.)	5%
Next 720 GHC (=up to 82% of GDP p.c.)	10%
Next 14,232 GHC (=up to 675% of GDP p.c.)	17.5%
Exceeding 16,200 GHC ($\geq$ 675% of GDP p.c.)	25%

Note: The table reports the marginal labor tax schedule in Ghana in 2011. It shows, by income in Ghanaian Cedis (GHC), the marginal tax rate assessed on labor income, and the corresponding ratio of GDP per capita in Ghana in 2011.

Table A.4: Elasticities of Moments to Parameters

	ν_J	ν_S	σ	ρ	η_S	b	Ψ_S	σ_ε	σ_ζ	χ
Aggregate population growth	0.8	0.0	0.0	0.6	-0.2	-0.1	0.0	0.0	0.0	0.0
Aggregate SHS attendance	-0.5	0.1	-0.1	-2.2	0.9	0.3	-0.1	-0.1	0.0	0.0
Intergenerational school corr.	0.0	0.0	0.5	1.7	-0.7	-0.2	0.2	-0.4	0.0	0.0
Var(permanent income)	-0.4	0.1	0.8	-1.1	1.7	0.0	0.0	0.0	0.0	0.0
SHS in top quartile	0.0	0.0	0.2	-0.4	0.3	0.0	0.0	-0.2	0.0	-0.1
SHS in bot quartile	0.0	-0.2	-0.4	-1.0	0.6	0.3	0.0	0.2	0.0	0.0
TE on human capital	0.0	0.0	-0.1	-0.2	0.1	0.0	0.0	0.0	0.0	0.0
TE on fertility	-0.5	0.3	0.8	3.0	-1.9	0.1	-0.4	0.5	0.0	-0.1
TE on SHS completion	-0.4	0.0	-0.8	-2.8	1.8	-0.2	0.4	-0.5	0.0	0.1
TE on SHS, Q4-Q1 difference	1.0	-0.3	-0.5	12.8	-8.8	2.7	-1.5	3.6	-1.6	1.1

Note: This matrix represents the elasticities of each moment to each parameter. The entry in row r and column c represents the percentage change in model moment r resulting from a one-percent increase in model parameter c.

Table A.5: Sensitivity of Parameters to Moments

	ν_J	ν_S	σ	ρ	η_S	b	Ψ_S	σ_ε	σ_ζ	χ
(i) TE on human capital x 1.5	3.6	2.1	-1.0	0.0	-0.7	4.2	-1.7	2.8	1.3	3.0
(ii) TE on fertility x 1.5	8.4	-3.1	4.5	-0.4	6.9	-2.9	4.7	3.8	4.7	-2.2
(iii) TE on SHS completion x 1.5	6.4	3.1	-3.0	-1.1	3.3	0.2	1.2	-0.7	1.4	8.3
(iv) TE on SHS, Q4-Q1 difference x -1	2.3	1.8	1.0	0.2	-0.6	-1.9	7.1	2.2	-2.8	-7.6
(i), (ii), (iii) and (iv) combined	10.3	-5.9	-2.8	1.5	17.9	-4.8	28.9	7.0	-2.8	-4.4

Note: This matrix displays the percent change in each parameter when the mode is re-estimated to match different target moments. Row (i) is when we target a fifty percent higher target for the treatment effect on human capital but keep all other targets the same. Row (ii) is when we target a fifty percent higher treatment effect on fertility. Row (iii) is when we target a treatment effect on secondary school completion that is fifty percent higher. Row (v) is when we target a difference in treatment effects on schooling between Q4 and Q1 that is -4 instead of 4. The bottom row is when we re-estimate the model to match all of the higher targets..

Table A.6: Support for Free Secondary School

	Obs.	Percent of respondents All	JHS	SHS	Difference (JHS - SHS)	P-value
Sample restricted to only ...						
Urban respondents	1,829	69.9	74.0	65.2	8.8***	0.00
Rural respondents	1,086	70.1	72.0	63.5	8.4**	0.01
Male respondents	1,427	67.6	70.2	64.5	5.7**	0.02
Female respondents	1,488	72.2	75.2	65.2	9.9***	0.00
Respondents from the Volta region	265	42.3	45.0	37.2	7.8	0.22
Respondents from the Ashanti region	854	78.2	81.7	69.8	11.9***	0.00

Notes: This table reports the share of respondents who support the free secondary school program. More specifically, respondents were asked whether the "expenditures on the free secondary school program should either be abolished, cut substantially, cut somewhat, maintained, expanded, or delayed/postponed." The table shows the percent of respondents who agreed that the expenditures on free secondary school should either be maintained or expanded. To check for the robustness of the responses, we limited the sample to only respondents in urban and rural localities, male and female heads of households, and respondents from the Volta and Ashanti regions of Ghana. The JHS column shows the percent of respondents who supports the free secondary school program and had completed at most basic education (Junior High School or below). On the other hand, the SHS column reports the percent of respondents who support the free secondary and had completed at least secondary education (Senior High School or above). ***, **, and * represent significance at the 1%, 5%, and 10% levels respectively.

B. Model Appendix

In this appendix we define the concepts of recursive competitive equilibrium and balanced growth path for our model. Letting X denote the vector of individual state variables $(\tau, a, z_p, s_p, \zeta_p, \delta_J, \delta_S, z_c, s_c, \tilde{z}_c, \zeta_c)$, a recursive competitive equilibrium is defined as follows.

Definition: A **recursive competitive equilibrium** consists of

1. A price system $w_S(f, P)$, $w_U(f, P)$
2. Household value functions $V(X, f, P)$ and policy functions $a'(X, f, P)$, $c(X, f, P)$, $s'_c(X, f, P)$
3. Perceived laws of motion $f' = F(f, P), P' = H(f, P)$

such that

a) V, a', c, s'_c solve the household's optimization problem given w_S, w_U, F, G.

b) For all f, P,

$$w_S(f, P) = (1-\alpha) AK^\alpha (N_J)^{\lambda-1} \Big[(N_J)^\lambda + (N_S)^\lambda\Big]^{\frac{1-\alpha}{\lambda}-1},$$
$$w_U(f, P) = (1-\alpha) AK^\alpha (N_S)^{\lambda-1} \Big[(N_J)^\lambda + (N_S)^\lambda\Big]^{\frac{1-\alpha}{\lambda}-1},$$
$$r^* = \alpha AK^{\alpha-1} \Big[(N_J)^\lambda + (N_S)^\lambda\Big]^{\frac{1-\alpha}{\lambda}}.$$

c) Markets clear:

$$N_J = \Big[\int_{6\le\tau\le12, s_p=J} \zeta_p h(z_p, s_p) f(X) dX + \int_{9\le\tau\le10, s'_c(X,f,P)=J} \zeta_c h(z_c, s'_c) f(X) dX \Big] P,$$
$$N_S = \Big[\int_{6\le\tau\le12, s_p=S} \zeta_p h(z_p, s_p) f(X) dX + \int_{\tau=10, s_c=S} \zeta_c h(z_c, s'_c) f(X) dX \Big] P.$$

d) Perceived laws of motion for f and P coincide with those induced from household policy functions a', c, s'_c.

The balanced growth path is a particular type of recursive competitive equilibrium defined below.

Definition: A **balanced growth path** is a recursive competitive equilibrium that satisfies the following properties:

1) Aggregate population grows at a constant rate: $\frac{P'}{P} = \nu$ for some constant $\nu > 0$.

2) The distribution of X is stationary: $f' = f$.

3) The household value and policy functions do not depend on P.

Along the balanced growth path, aggregate population grows but the distribution of households across individual states remains stationary. Further, the household value and policy functions are independent of aggregate population, and thus household behavior remains the same over time conditional on the individual states.

Now we walk through the details of population growth within the model and discuss how model parameters translate to outcomes that are measured in data such as the aggregate population growth rate and the number of children per household. We start with the most general case that applies to any equilibrium whether it satisfies the properties of a balanced growth path or not. Later, we specialize to the case of the balanced growth path to provide more explicit formulas. By definition, the aggregate population growth rate is given by the formula

$$\text{Agg. Pop. Growth Rate} = \frac{\text{\# births} - \text{\# deaths}}{P} \tag{11}$$

Given the aggregate state variables of the economy f, P, we have the following accounting equations for births and deaths

$$\text{\# births} = \big[\nu_J \int_{s_p=J,\tau=5} f(X)dX + \nu_S \int_{s_p=S,\tau=5} f(X)dX\big]P \tag{12}$$

$$\text{\# deaths} = \big[\int_{\tau=14} f(X)dX\big]P \tag{13}$$

In any given period, the aggregate population growth rate can be computed from state variables as

$$\nu - 1 = \nu_J \int_{s_p=J,\tau=5} f(X)dX + \nu_S \int_{s_p=S,\tau=5} f(X)dX - \int_{\tau=14} f(X)dX \tag{14}$$

Note that as written, $\nu > 1$ is the aggregate population growth rate such that $P' = \nu P$. To compare to data, it must be converted to an annual percentage growth rate.

Recall that the aggregate population growth rate ν is constant along the balanced growth path by definition. By leveraging this assumption we can calculate the aggregate population growth rate as a function of educational shares along the balanced growth path analytically. This calculation provides insight into the changes in population dynamics that can be expected due to changes in education. Such changes are important for our general equilibrium analysis.

With the aggregate population growth rate fixed at ν, we know that the ratio of the population of households of age x and households of age y must be given by:

$$\frac{\int_{\tau=x} f(X)dX}{\int_{\tau=y} f(X)dX} = \nu^{y-x} \tag{15}$$

From that fact that $\tau \in \{1, \cdots, 14\}$ and $\int f(X)dX = 1$ because f is a pdf, we can derive that along the balanced growth path with aggregate population growth rate ν the following equations are true

$$\int_{\tau=14} f(X)dX = \frac{\nu - 1}{\nu^{14} - 1} \tag{16}$$

$$\int_{\tau=5} f(X)dX = \frac{(\nu - 1)\nu^9}{\nu^{14} - 1} \tag{17}$$

Finally, because household policy functions are invariant with respect to P and f is stationary along the balanced growth path we have that the share of the adult population with a given level of education is the same for all ages. In particular, this implies that the education shares of the parents giving birth this period can be replaced by the aggregate education shares $\hat{J}, \hat{S}$.

$$\hat{J} \equiv \frac{\int_{s_p=J,\tau\geq 5} f(X)dX}{\int_{\tau\geq 5} f(X)dX} = \frac{\int_{s_p=J,\tau=5} f(X)dX}{\int_{\tau=5} f(X)dX} \tag{18}$$

$$\hat{S} \equiv \frac{\int_{s_p=S,\tau\geq 5} f(X)dX}{\int_{\tau\geq 5} f(X)dX} = \frac{\int_{s_p=S,\tau=5} f(X)dX}{\int_{\tau=5} f(X)dX} \tag{19}$$

Combining equations (16) to (19) with equation (14) yields the following equation which describes the aggregate population growth rate along the balanced growth path as an implicit function of the education shares of the population:

$$\nu - 1 = \left[\nu^9\left(\nu_J \hat{J} + \nu_S \hat{S}\right) - 1\right] \frac{\nu - 1}{\nu^{14} - 1} \tag{20}$$

which can be reduced to

$$\nu^5 = \nu_J \hat{J} + \nu_S \hat{S}. \tag{21}$$

One wrinkle not yet addressed is the fact that, as written, the balanced growth path of the model is not an attractor. That is, the model does not necessarily converge over time to the BGP. To see why, consider a simplified model with two generations, each of whom do nothing other than live through their first period of life and, at the end of their second period of life, die and have ν children who become the new first generation. If the initial stocks of age 1 and age 2 agents are N_1 and N_2, the aggregate population growth rate will oscillate between $\frac{(\nu-1)N_2}{N_1+N_2}+1$ and $\frac{(\nu-1)N_1}{N_1+\nu N_2}+1$ indefinitely, never converging to a single constant rate, as there is no mechanism to close "gaps" in size between the initial stocks.

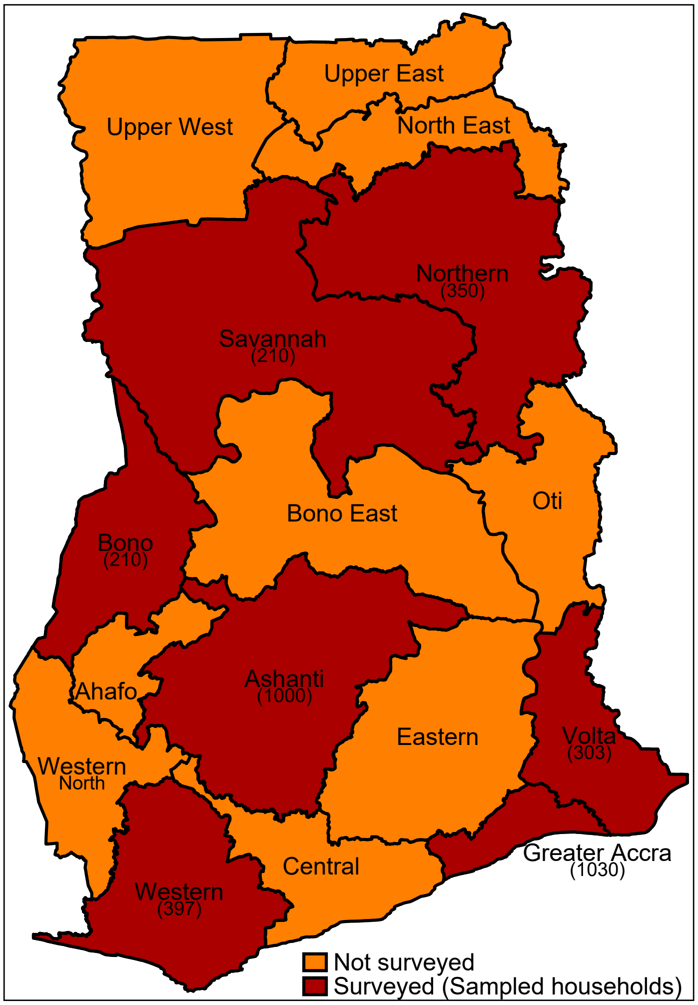
To address the computational issues arising from this fact, we assume that a neglibly small fraction of children leave their parents and have their own children one period earlier than the typical timing (that is, at age 20 rather than 25). This slight randomization in timing effectively mixes away any differences in the initial stocks of agents for each generation, ensuring that the model converges to the BGP over time regardless of the initial state. In our computations, we assume the probability that any given child leaves early is 0.1 percent, small enough to ensure that this outcome has minimal impact on parents' decisions.

C. Household Survey Appendix

We surveyed a nationally representative sample of Ghanaian households to explore residents' support for the free secondary school program. To do so, we first split the country into three zones: northern, middle, and coastal. Then we chose two regions from the northern zone (Northern and Savannah), two from the middle zone (Ashanti and Bono), and three from the coastal zone (Greater Accra, Volta, and the Western regions). According to the 2021 Population and Housing Census (PHC), the selected seven areas account for almost 61 percent of the entire population. Figure C.1 depicts a map of the studied regions. The dark grey areas are the regions we surveyed, whereas the light grey parts are the regions we did not survey. In addition, the values in parenthesis in the dark grey zones – the regions we surveyed – show the number of households sampled in each region.

Second, we obtained an exhaustive list of enumeration areas (EA) across the seven regions from the 2021 Population and Housing Census conducted by the Ghana Statistical Services. The list of EAs included details such as location, type of residence (urban or rural), and projected size (number of households). We divided the EAs into 14 strata based on geography (7) and kind of habitation (2). Following that, we selected a nationally representative sample of 3,500 homes from the list of EAs using a two-stage cluster sampling approach. In the first stage, we selected 151 EAs at random and independently using a probability proportional to size to allocate the total number of EAs per stratum. In the second stage, an average of 23 homes were chosen at random from each EA sampled in the first stage. Following that, twenty-four (24) enumerators were hired, trained, and tasked with interviewing an adult individual from each of the 3,500 chosen houses, preferably the head of the household.

Figure C.1: Sampled Households in each Region Surveyed



Notes: This figure plots the regions that were surveyed and the total number of households surveyed in each region. The numbers in the parenthesis represent the number of households sampled in each region.