Title

Essays on Health, Human Capital and Economic Development

Author

Kim, Minki

Publication Date

2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Essays on Health, Human Capital and Economic Development

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Economics

by

Minki Kim

Committee in charge:

Professor David Lagakos, Co-Chair
Professor Tom Vogl, Co-Chair
Professor Titan Alon
Professor Munseob Lee
Professor Craig McIntosh

2023

The Dissertation of Minki Kim is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my deep gratitude and appreciation to my two co-advisors, David Lagakos and Tom Vogl, for their invaluable guidance, support, and encouragement throughout the process of writing this dissertation. Especially during the COVID-19 pandemic, their continuous support and guidance were instrumental in making this work possible.

I would also like to extend my sincere appreciation to the other members of my dissertation committee, Titan Alon, Munseob Lee, and Craig McIntosh for their continuous guidance and support.

I would also like to extend my heartfelt thanks to my family for their unwavering love and support, particularly my parents. Their sacrifices and encouragement have made this journey possible.

Last but not least, I would like to express my deepest appreciation to my wife, June, who is not just my life companion and source of emotional support, but also a fellow researcher who provided me with valuable discussions and comments throughout this process. Her unconditional love, support, and understanding have been a constant source of inspiration and motivation.

VITA

2016	Bachelor of Arts in Economics, Yonsei University
2019	Master of Arts in Economics, University of California San Diego
2023	Doctor of Philosophy in Economics, University of California San Diego

FIELDS OF STUDY

Macroeconomics, Economic Growth and Development

ABSTRACT OF THE DISSERTATION

Essays on Health, Human Capital and Economic Development

by

Minki Kim

Doctor of Philosophy in Economics

University of California San Diego, 2023

Professor David Lagakos, Co-Chair
Professor Tom Vogl, Co-Chair

This dissertation consists of three chapters. In Chapter 1, I study the macroeconomic consequences of eradicating malaria in sub-Saharan Africa. To do so, I combine a general-equilibrium overlapping generations model with reduced-form empirical evidence. I find eliminating malaria in a representative sub-Saharan Africa through vaccination would increase the GDP per capita by 30% in the long run, which is nearly ten times larger than previously estimated. I also find that the gains stem from larger investments in human capital, amplified over multiple generations.

In Chapter 2, in work joint with Titan Alon, Natalie Cox, and Arlene Wong, we study

the welfare and productivity consequences of rising student debts in the United States. We first empirically estimate how student debts affect workers' early career outcomes using NLSY panel data. Then we construct a quantitative life-cycle model calibrated to match the empirical evidence and evaluate the federal student loan policies. We find that student debt forgiveness or repayment elongation policies improve welfare and labor productivity. The model suggests that a big chunk of the productivity gain comes from a small fraction of the workforce, who switch occupations in response to the policies.

In Chapter 3, in work joint with Titan Alon, David Lagakos, and Mitchell VanVuren, we provide a quantitative macroeconomic framework to study why emerging markets fared worse relative to advanced economies and low-income countries during the COVID-19 pandemic. We adopt a workhorse incomplete-markets macro model to include epidemiological dynamics alongside key economic and demographic characteristics that distinguish countries of different income levels. We conclude that emerging markets fared especially poorly due to their high employment share in occupations requiring social interactions and their low level of public transfers. In contrast, low-income countries fared relatively better due mainly to their younger population and larger agricultural sector.

Chapter 1

How Will a New Malaria Vaccine Shape Africa's Economic Future? A Macroeconomic Analysis

1.1 Introduction

Despite preventive technologies and treatment, malaria is still the leading cause of death and a barrier to children's human capital accumulation in sub-Saharan Africa. In 2020 alone, more than 600,000 people died from malaria, mostly African children under five years old (World Health Organization, 2021). Children who survive are also known to suffer from long-lasting cognitive impairments and co-morbidities (Fernando et al., 2010; Chen et al., 2016), which adversely affect their learning outcomes. In recent years, however, scientists have taken a major step towards eliminating malaria, as the newly developed vaccine is reported provide up to 80% protection against infections among young children (Datoo et al., 2022).

How will the new malaria vaccine change the macroeconomic outlook of sub-Saharan African countries? One view in the macroeconomics literature thus far is that eliminating malaria would mainly increase populations but not substantially raise living standards. For example, Acemoglu and Johnson (2007) study the effects of large improvements in life expectancy in the 1940s driven by international health interventions, more effective public health measures, and the introduction of new chemicals. The study concludes that the improvements in life expectancy had little impact on GDP per capita but only increased populations. Ashraf, Lester, and Weil (2008) corroborate this view, arguing that eradicating malaria in a typical sub-Saharan African country would increase GDP per capita only by two percent over a 60-year horizon.

This paper reassesses this conclusion by modeling and quantifying the long-run macroeconomic effects of a successful malaria vaccine and argues that the increase in long-run output per capita from eliminating malaria is much larger than the existing estimates. I focus on several new features that have been absent from previous macroeconomic studies of disease eradication. The first is a quantity-quality tradeoff parents face when making fertility and investment decisions in children. The second is a richer measurement of human capital than just years of schooling, as in Manuelli and Seshadri (2014). Using years of schooling as a measure of human capital implicitly assumes that one year of schooling delivers the same increase in human capital before and after eliminating malaria (Hanushek and Woessmann, 2008). However, eliminating malaria also allows children to learn more from schooling, increasing the amount of human capital gained per year of schooling. The third is the intergenerational dynamics, as healthier children may subsequently adjust their fertility and invest more in their own children's education.

I incorporate these features into a general equilibrium, heterogeneous-agent, overlapping generations model which allows for the interplay of fertility, childhood human capital accumulation, and childhood diseases. In the model, parents endogenously choose how many children to have as well as the educational attainment of their children. Children are born with exogenously heterogeneous learning abilities inherited stochastically from their parents. Parents base their fertility decision on assets, income, and tastes for children. Parents also make a decision on whether to educate their children or send them to work by comparing the higher income and consumption today from child labor against the higher future utility their children will enjoy if they receive more education. Children's subsequent outcomes are then determined by their learning ability, the educational choice made by parents, and the skill-contingent wages determined in the competitive labor market.

I then introduce childhood disease into the model as an idiosyncratic health shock to children. Health shock in the model consists of two dimensions; mortality and morbidity. Mortality reflects the deaths caused by diseases, and morbidity represents the long-term cognitive damage on children, which dampens their human capital accumulation. A key feature of the model is that childhood disease interacts with parents' education and fertility decisions through the quantity-quality tradeoff (Barro and Becker, 1989). Within the quantity-quality framework, eliminating a disease has two opposing effects. On the one hand, lower mortality from the elimination lowers the price of child *quantity*, inducing parents to have more children and reducing per-child educational investment. On the other hand, reduced morbidity lowers the price of child *quality*, inducing parents to have fewer children and educate each child more. In the model, these two effects are summarized by two sufficient statistics which govern the long-run effects of disease elimination: fertility and education elasticities of disease.

To estimate the size of the fertility and education elasticities of malaria, I exploit a recent large-scale antimalarial campaign in sub-Saharan Africa, the Roll Back Malaria (RBM) campaign. Starting from 2003, the RBM campaign significantly reduced malaria prevalence through its aggressive distribution of preventive and treatment technologies, such as insecticide-treated nets (ITNs) and indoor residual spraying (IRS). I focus on Tanzania, one of the few countries where the campaign started at its earliest phase. Exploiting regional variation in the campaign's intensity, I estimate the effects of the reduced malaria risk on women's fertility and children's years of schooling through a difference-in-differences design. Consistent with the previous studies, I find that the reduced prevalence of malaria due to the campaign led to an average of 0.63 more years of schooling for the benefited children.[1] I also find that women in the highest malaria prevalence regions reduced their fertility by 5.9%. However, although fewer children were born, the campaign did not change the number of surviving children because the campaign also reduced child mortality by 10%.

[1]For empirical studies found positive effects of reduced malaria risk on children's educational attainment, see Lucas (2010); Venkataramani (2012); Barofsky et al. (2015); Shih and Lin (2018); Kuecken et al. (2021). For studies focusing on adulthood outcomes such as income or consumption, see Cutler et al. (2010); Bleakley (2010). Unlike this paper, most of these studies used historical episodes of malaria eradication.

I quantitatively solve for the balanced growth path and bring the estimated elasticities to the model by replicating the Roll Back Malaria campaign within the model and matching the campaign's simulated causal effects to their empirical counterparts; a 0.63 year increase in children's schooling and a muted response of fertility. To ensure the model is credible in other dimensions, I jointly estimate parameters to match the empirical elasticities and other relevant aggregate moments of the Tanzanian economy, such as educational attainment and intergenerational mobility measures. Since the causal effects of reduced malaria risk on children's education outcomes and women's fertility are also estimated in Tanzania, all the moments used for calibration are from a single country.

Using the estimated model, I then simulate the long-run general equilibrium effects of a national malaria vaccine policy by solving for a new balanced growth path of the economy under the vaccination policy. The model predicts per capita income would rise by 34% within 60 years. Compared to the short-run, one-generational effects of policy, the long-run increase in per capita output is twice as large as the short-run counterpart, highlighting the importance of the intergenerational amplification channel. While the lowered mortality rate from vaccination is expected to put upward pressure on the population, the population growth rate on the post-vaccine balanced growth path is only modestly higher because households choose to have fewer children in the long run. The model also predicts a significant increase in both primary and secondary school completion rates and an improvement in intergenerational mobility in terms of education. While only 40% and 3% of the children born to uneducated parents complete primary and secondary education before the vaccination, the numbers increase to 52% and 5%, respectively, in the post-vaccination balanced growth path.

The results are surprising because they are nearly ten times larger than the literature's current best estimates from the influential work by Acemoglu and Johnson (2007) and Ashraf, Lester, and Weil (2008). Acemoglu and Johnson (2007) estimate the effects of life expectancy on economic growth, using the major international health improvement in the 1940s (the international epidemiological transition) as a natural experiment. They conclude that while improvement in life expectancy leads to an increase in population, it does not lead to an increase in GDP per capita. A key difference between the international epidemiological transition in the 1940s and malaria vaccines is that while the former mostly lowered mortality without increasing the returns to education, the latter would lower the mortality *and* raise the returns to education. I simulate a counterfactual long-run balanced growth path where only the mortality is lowered. Consistent with Acemoglu and Johnson (2007), the long-run increase in per capita output shrinks from 34% to a mere 2%. The result suggests that the long-run growth effects of a health improvement hinge on whether it facilitates the accumulation of human capital.

Ashraf, Lester, and Weil (2008) specifically focus on malaria in sub-Saharan Africa. Using a standard neoclassical framework, they simulate the effects of eradicating malaria in a typical sub-Saharan African country and conclude that it would raise GDP per capita only by two percent in the long run. The underlying reasons for these different results are the difference in the measures of human capital and the omission of intergenerational dynamics. First, Ashraf et al. (2008) treat malaria eradication as a one-time increase in human capital, rather than considering intergenerational dynamics that amplify the effects in the long run. Second, following contemporary best practice, they employ years of schooling as a measure of human capital. An implicit assumption behind this is that a year of schooling delivers the same increase in human capital before and after eradication (Hanushek and Woessmann, 2008). However, eradicating malaria would also make children healthier and perform better *within* school (Fernando et al., 2010), an important margin that years of schooling cannot address. In the model, the increase in years of schooling only explains 30% of the increase in human capital for the children born after the vaccination, implying that focusing only on years of schooling could underestimate the increase in human capital. In Section 1.5.2, I explore the quantitative implications of omitting the intergenerational dynamics and using years of schooling as the human capital measure and find that without the two channels, the long-run increase in output per capita is reduced from 34% to 5%, much closer to the numbers found in Ashraf et al. (2008).

I conclude that eliminating malaria is not only a life-saving policy, but also a growth policy for sub-Saharan African countries. The estimated model implies that once the long-run intergenerational effects and the better learning in school are considered, the long-run increase in output per capita can be much larger than what is estimated in the literature thus far. This suggests that removing malaria should be a high priority in development strategy for policymakers. The results also imply that improving children's health conditions in the developing world can raise the living standard in the long-run, primarily through the higher human capital accumulation of children. Such implication is consistent with the conclusions of the macro-development literature, which have been emphasizing the role of human capital in explaining the cross-country income differences (Erosa, Koreshkova, and Restuccia, 2010; Schoellman, 2012; Hendricks and Schoellman, 2018; De Philippis and Rossi, 2021).

While this paper specifically focuses on malaria, the model can easily be generalized to other childhood infectious diseases and used as a framework to study the effects of improving health on long-run economic growth. In this vein, this paper builds on a long literature that studied the relationship between health, human capital, and economic growth (Shastry and Weil, 2003; Caselli, 2005). Weil (2007) shows that eliminating health differences across countries has minor effects in narrowing the cross-country gaps in output per capita, while the quantitative results in this paper suggest that such effects can be potentially larger. The theoretical framework in this paper is related to those in Kalemli-Ozcan, Ryder, and Weil (2000); Kalemli-Ozcan (2003); Soares (2005) and Doepke (2005), all of whom focused on the role of mortality decline in human capital investment and growth. Lastly, the quantitative model of this paper is perhaps most closely related to those in Daruich (2020) and Zhou (2022), who study the long-run effects of education subsidy and family policies, respectively, using a general equilibrium, heterogeneous-agent overlapping generations model. However, none of these previous studies considers diseases as a factor disrupting childhood human capital accumulation, which the model suggests to be quantitatively important.

This paper also builds on a growing body of research in macroeconomic development that uses dynamic general equilibrium models to understand the potential long-run general equilibrium effects of development policies, using the short-run, partial equilibrium empirical evidence as ingredients. In this vein, the quantitative exercises of this paper are related to those of Fried and Lagakos (Forthcoming), who use a dynamic general equilibrium model to quantify the effects of power outages on productivity in developing countries. Similarly, Buera, Kaboski, and Shin (2021a) use a dynamic macroeconomic model of credit-constrained firms to study the general equilibrium effects of microfinance, using the partial equilibrium estimates of microfinance to discipline the model. Another example is Lagakos, Mobarak, and Waugh (2023), who quantify the aggregate effects of rural-to-urban migration subsidies compared to the partial equilibrium effects estimated from an RCT. Brooks and Donovan (2020) also use the reduced-form evidence on the effects of rural bridge building to guide the general equilibrium effects of transportation infrastructure.[2]

The rest of the paper is organized as follows. Section 1.2 lays out the quantitative model. Section 1.3 describes the data used to estimate key parameters and the estimation of empirical moments used to discipline the rest of the model parameters. Section 1.4 explains the calibration strategy and model validation results. Section 1.5 investigates the long-run general equilibrium effects of a national malaria vaccination policy and compares the results to the previous estimates. Section 1.6 concludes.

1.2 Model

In this section, I introduce a general equilibrium overlapping generations model with endogenous population dynamics, childhood human capital investment, and health risk. A household consists of parents and cohabiting children. Parents endogenously choose how many children to have as well as the educational level of their children, who are born with exogenously heterogeneous learning abilities. Children's endowment when they become adult are determined by their learning ability, the educational choice made by parents. Workers with different levels of education are imperfect substitutes, and their education-contingent wages are determined in a

[2] See Buera, Kaboski, and Townsend (2021b) for a detailed review of the literature.

competitive labor market.

1.2.1 Environment

Demographics

Time is discrete, and one model period is six years. The economy is populated by a large number of overlapping generations of households who live for 66 years (12 periods in total). Figure 1.1 shows the life cycle and family structure of a household. I use $j \in \{0, 1, \cdots, 12\}$ to denote the period of life.

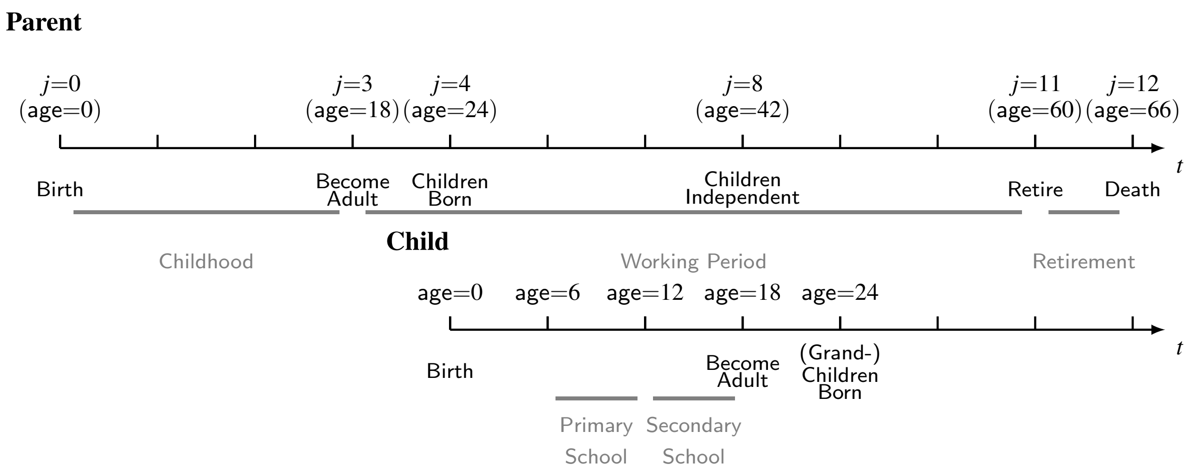


Figure 1.1. Life cycle, family structure, and stages of life

Children live with their parents and do not make any decisions on their own until they reach age 18, when they leave their parents and become independent with zero assets. Throughout their adulthood, individuals choose consumption expenditure and savings. Borrowing is not allowed but they can save through assets with exogenous interest rate r.[3] parents choose how many children to have at age 24, and conditional on having children, how much education their children will receive until they become adults. There are two periods of schooling; parents have to decide whether or not to send children to primary school when the children are six years old,

[3]I abstract from domestic capital market to focus on the main mechanism of the model. Most countries in this context are small open economy with under-developed capital market. I also omit the endogenous labor supply retirement choice, due to high hours worked and short retirement periods in the developing worlds.

and secondary school when they are twelve years old. The initial human capital their children possess is therefore directly influenced by the parents' educational choice. Once the children become independent, there is no further interaction between parents and children and the human capital is fixed throughout the adulthood. All individuals retire at age 60 and die at age 66. During the short retirement periods, individuals live off the assets they have accumulated in the working period.

There are four exogenous sources of heterogeneity in the model. The first one is the standard, idiosyncratic and uninsurable labor productivity shock v_t for the working-age adults, which makes earnings stochastic. I assume that the idiosyncratic labor productivity shock is i.i.d. and drawn every period from a log-normal distribution:

$$\log v_t \overset{\text{iid}}{\sim} N(0, \sigma_v)$$

The second one is the fertility taste, which captures the fertility behaviors not attributed to the model's mechanism. I assume an extreme-value distribution-. The third one is the learning ability that every child is born with, which is imperfectly correlated between parents and children. The last one is the health shock that children face, which represent the childhood diseases, or malaria in our context. Health shock is drawn once at the early childhood when the children are six years old, and lowers the returns from schooling in subsequent childhood periods. I illustrate the way the ability and health shocks are drawn and how they interact with the parents' education decision in the next paragraphs.

Learning Ability

Parents observe their children's learning ability at the beginning of the period $j = 5$, before they decide whether or not to send their children to primary school. The learning ability within a household follows a AR(1) process:

$$\log z_k = \rho_z \log z_p + \varepsilon_z$$

where z_k and z_p are the learning ability of children and parents and ε_z is the idiosyncratic i.i.d. shock. Therefore, learning ability is inherited across generations but only imperfectly.

Diseases in Early Childhood

In addition to the learning abilities, children also draw a idiosyncratic health shock at age six.[4] The health shock has two dimensions: mortality and morbidity. Mortality reflects the deaths caused by the diseases, and morbidity represents the detrimental effects of the diseases on human capital accumulation. Mortality risk is manifested as a survival probability. Specifically, the probability an age-6 child survives to the next period is given as χ^d. Morbidity risk is represented by a proportional reduction in the learning ability, reflecting that children with malaria suffer from worse learning outcomes during and after bouts of malaria (Fernando et al., 2010). I discuss how the lowered learning ability affects children's human capital accumulation in the next paragraph. Specifically, I denote the morbidity shock as m:

$$m = \begin{cases} 1 & \text{w/ probability } 1-\chi^m \\ \underline{m} & \text{w/ probability } \chi^m \end{cases}$$

where $\underline{m} < 1$. Here, $\underline{m}$ denotes the health status of a child hit by the adverse morbidity shock, while the status of a healthy child is normalized to one. The probability that a child is stricken with a negative morbidity shock is given as χ^m.

Schooling

After observing their children's learning ability z_k and the health shock m they have drawn, parents decide whether to send their age-6 children to the primary school. Schooling

[4]The assumption that the health shock is realized at age 6 can be understood as parents learn about their children's health status by the time they enter primary school.

increases human capital deterministically:

$$h_{k,t+1} = \begin{cases} \max\{mz_k \eta_s h_{k,t}, h_k\} & \text{if attend school} \\ h_k & \text{otherwise} \end{cases}$$

where η_s is the deterministic increase in human capital from school s, $s \in \{\text{Primary}, \text{Secondary}\}$. Sending children to school costs money, which are represented by the per-child schooling fee p_P, p_S for primary and secondary school, respectively. The schooling fee encompasses the tuition, uniforms, schooling supplies such as textbooks, etc., representing the goods cost of education. Children can work instead of going to school (child labor). Lastly, schooling decision is sequential; if parents do not send their primary school aged child to school and have them work instead, the child does not have the opportunity to attend the secondary school in the next period.

Production and Aggregation

I assume that there is a profit-maximizing representative firm in the labor market. The representative firm uses skilled (secondary education completed) and unskilled (below secondary education) labor to produce the single consumption good with the following CES aggregate production function:

$$Y = A\left[(H_U + H_P)^{\frac{\lambda-1}{\lambda}} + (H_S)^{\frac{\lambda-1}{\lambda}}\right]^{\frac{\lambda}{\lambda-1}}, \quad \lambda \in (0, \infty)$$

Here, H_s denotes the aggregate efficiency unit of schooling groups U (uneducated), P (primary-completed) and S (secondary-completed), and λ is the elasticity of substitution between the two skill groups. Equilibrium wages for each skill group are then determined as:

$$w_U = A\left[(H_U + H_P)^{\frac{\lambda-1}{\lambda}} + (H_S)^{\frac{\lambda-1}{\lambda}}\right]^{\frac{1}{\lambda-1}} (H_U + H_P)^{-\frac{1}{\lambda}}$$
$$w_S = A\left[(H_U + H_P)^{\frac{\lambda-1}{\lambda}} + (H_S)^{\frac{\lambda-1}{\lambda}}\right]^{\frac{1}{\lambda-1}} H_S^{-\frac{1}{\lambda}}$$

Since the wages w_U and w_S are per efficiency unit, the labor income of an individual with skill level s, human capital h, and idiosyncratic shock v is given as $y(h,v,s) = hvw_s$.

1.2.2 Recursive Formulation of Decision Problems

From independence, an individual's adulthood can be broadly divided into two stages depending on whether they are alone or live with their children. In both stages, individuals solve consumption-savings optimization problem. When with children, however, they make additional decisions on fertility and children's education. In this subsection, I explain the individual optimization problem in each period of life. Borrowing is limited in all periods ($a' \geq 0$), and I suppress the expression in following formulations to simplify the notations. Throughout this section, I will denote all child variables with subscript k, and future variables with primes.

Age 18 $(j = 3)$**: Independence**

Individuals leave their parents and form a new household at age 18 with zero assets. Their initial states are human capital h, schooling level s which is determined in the childhood, and the learning ability z. While the learning ability is not relevant for themselves, it is still included as a state variable because their own children's learning abilities will depend on it. Their income is determined by their human capital h, wage rate w_s, and the idiosyncratic income shock v drawn at the beginning of the period. Since they do not have children yet, they solve a standard consumption-savings problem:

$$
\begin{aligned}
V_3(a,s,h,v) = \max_{c,a'} u(c) + \beta \mathbb{E}\left[V_4\left(a',s,h,v'\right) \right] \\
\text{subject to} \\
c + a' \leq w_s hv + (1+r)a
\end{aligned}
\tag{1.1}
$$

where r is the period interest rate.

Age 24 $(j = 4)$**: Fertility**

At this stage, individuals decide how many children to have, and those who choose to have children become parents. Fertility choice is discrete; individuals choose the number of children n, where $n \in \{0, 1, 2, \cdots, \bar{N}\}$, by choosing n^* that gives them the highest level of utility:

$$V_4 = \max\{V_4^0 + \phi_0, V_4^1 + \phi_1, \cdots, V_4^{\bar{N}} + \phi_{\bar{N}}\}$$

where V_4^n represents the value functions of having n number of children. For each number of children n, I also introduce a taste shock ϕ_n, which are drawn *i.i.d.* from a Gumbel distribution with variance σ_n. The value function corresponding to haveing n children can be written as follows:

$$\begin{aligned} V_4^n(a,s,h,z,v) = \max_{c,a'} u(c) + \beta \mathbb{E}\Big[& V_5\left(a', s, h, v', z_k', m', n'\right)\Big] \\ & \text{subject to} \\ & c + a' \leq w_s h v(1 - t(n)) + (1+r)a \\ & t(n) = \omega_1 n^{\omega_2} \end{aligned} \tag{1.2}$$

where $\boldsymbol{\phi}$ is the vector of fertility taste shock draws. Raising children is costly. Specifically, $t(n)$ amount of parental time is taken away, reducing the available income for consumption and savings. The total time cost is increasing in h and n, reflecting the fact that the sacrificed working hour is more costly to high-income households than low-income ones. Lastly, note that the expectation is taken over the number of *surviving* children in the next period, n'. This is due to the realization of the health shock in the next period.

Age 30 $(j = 5)$**: Ability, Health Shock and Primary Education**

At the beginning of this period, parents observe children's ability z_k and the realization of health shock m, and draw the idiosyncratic income shock v. The number of surviving children n is determined depending on how many children are hit by the mortality shock. Specifically, the probability that n out of N children survive is:

$$f(n;N) = \binom{N}{n} (\chi^d)^{N-n} (1-\chi^d)^n$$

Parents then decide whether to send their children (if any) to primary school ($e=1$) or workplace ($e=0$). If children attend school, a per-child primary schooling fee p_P is deducted from the parents' budget constraint. If the children go to workplace instead (child labor), their income is added to the parents' budget. Since the working children did not receive any education, their have one (initial) level of human capital and receive unskilled wages.[5] A parent's value function with n number of surviving children at this stage is written as follows:

$$\begin{aligned} V_5(a,s,h,v,z_k,m,n) = & \max_{c,a',e\in\{0,1\}} u(c) + \beta\mathbb{E}\left[V_6\left(a',s,h,v',s_k',h_k',z_k,m,n\right)\right] \\ & \text{subject to} \\ & c + a' + enp_P \leq w_s hv(1-t(n)) + nw_U v(1-e) + (1+r)a \\ & h_k' = eh_k\eta_P z_k m + (1-e)h_k \end{aligned} \tag{1.3}$$

Age 36 ($j=6$): Secondary Education and Dynastic Altruism

Parents who had sent their children to the primary school in the previous period decide whether to continue sending children to secondary school. Because of the sequential nature of the schooling system, parents who did not send their children to the primary school do not have option to send them to secondary school ($e=0$ for them). The value function of the parents with

[5] I assume that the children's income is also subject to the same idiosyncratic income shock v that their parents have drawn; the income shock is common across the household members.

secondary school age children is:

$$V_6(a,s,h,v,s_k,h_k,z_k,m,n) = \max_{c,a',e} u(c) + \beta \mathbb{E}\left[V_7\left(a',s,h,v'\right)\right] + \beta b(n) \mathbb{E}\left[V_{4,k}\left(s_k', h_k', z_k, v_k'\right)\right]$$

$$\begin{aligned} &\text{subject to} \\ &c + a' + np_S e \leq w_s hv(1-t(n)) + nw_{s_k} h_k v(1-e) + (1+r)a \\ &h_k' = eh_k \eta_S z_k m + (1-e)h_k \end{aligned} \tag{1.4}$$

Note that the value function at this stage now includes the continuation value of the children, $V_{4,k}$, discounted by the altruism function $b(n)$. Since the problem is written recursively, this continuation value captures the parental altruism toward their children; they take into account the utility value of all of their descendants and make decisions accordingly.

Age 42 to 66 ($j = 7 - 12$): Mature Parents with Grown-up Children

At the beginning of the period 7, children become independent and leave their parents. Once the children become independent, there is no further interaction between parents and children and parents solve a simple consumption-savings problem. Value of working-age individuals after the child-rearing periods is the same as (1.1). At age 60, households retire and are assumed to provide no work. The value function after retirement is given as:

$$V_j(a) = \max_{c,a'} u(c) + \beta V_{j+1}\left(a'\right)$$

$$c + a' \leq (1+r)a$$

1.2.3 Competitive Equilibrium and Balanced Growth Path

In this economy, population is endogenous due to the endogenous fertility. Therefore, I focus on a balanced growth path of the economy where the population growth rate remains constant over time, as does the distribution of households' state variables. I introduce the concepts of the recursive competitive equilibrium and the balanced growth path below.

Recursive Competitive Equilibrium

To save notations, let the vector of age-j individual's state variables $(a, s, v, s_k, h_k, z_k, m, n)$ as $\mathbf{X}_j$ and the distribution of the age-j state variables as $\mu(\mathbf{X}_j)$. A recursive competitive equilibrium consists of

(a) Household value functions $V_j(\mathbf{X})$ and policy functions $c_j(\mathbf{X}), a'_j(\mathbf{X}), n_4(\mathbf{X}), e_5(\mathbf{X}), e_6(\mathbf{X})$

(b) Prices for each skill group w_U and w_S

such that

(i) V, a', c, n_4, e_5, e_6 solve the individual's optimization problem conditional on w_U and w_S

(ii) The representative firm maximizes its profit:

$$w_U = A\left[(H_U + H_P)^{\frac{\lambda-1}{\lambda}} + (H_S)^{\frac{\lambda-1}{\lambda}}\right]^{\frac{1}{\lambda-1}} (H_U + H_P)^{-\frac{1}{\lambda}}$$
$$w_S = A\left[(H_U + H_P)^{\frac{\lambda-1}{\lambda}} + (H_S)^{\frac{\lambda-1}{\lambda}}\right]^{\frac{1}{\lambda-1}} H_S^{-\frac{1}{\lambda}}$$

(iii) Prices clear the labor market

Balanced Growth Path

A balanced growth path is a particular case of recursive competitive equilibrium which satisfies further conditions and is defined below. Let P the aggregate population. A balanced growth path is a recursive competitive equilibrium that satisfies:

(a) Aggregate population grows at a constant rate: $\frac{P'}{P} = \nu$ for some constant ν

(b) The distribution of households is stationary: $\mu'(\mathbf{X}_j) = \mu(\mathbf{X}_j)\ \forall j$

(c) Decision rules (a) are stationary and do not depend on P

1.3 Empirical Analysis of the Effects of Malaria on Fertility and Children's Human Capital

The model has a rich set of interactions between disease, fertility, and children's human capital, which is encapsulated by the quantity-quality tradeoff. Within this framework, eliminating a disease has two opposing effects. On the one hand, lower mortality from the elimination lowers the price of child *quantity*, inducing parents to have more children and reducing per-child educational investment. On the other hand, reduced morbidity lowers the price of child *quality*, inducing parents to have fewer children and educate each child more. In the model, these two effects can be summarized by sufficient statistics: fertility and education elasticities of disease. Since these two elasticities govern how parents adjust their fertility and education decisions in response to a better health environment, the sizes of these elasticities are essential to understand the long-run effects of disease elimination.

To ensure that the model's implied elasticities are indeed credible, I empirically estimate the elasticities and ask the model to replicate them. To this end, I use the Roll Back Malaria campaign, a recent large-scale anti-malaria campaign that took place in many sub-Saharan African countries. I focus on one country, Tanzania, and employ a difference-in-differences design to identify the causal effect of the reduction of malaria burden on fertility and children's human capital, exploiting the spatial variations in pre-campaign malaria prevalence as the identifying variations.

1.3.1 Background: The Roll-Back Malaria Campaign in sub-Saharan Africa

It's worth looking at the malaria situation in Tanzania and how the campaign was implemented. Before the Roll Back Malaria (RBM henceforth) campaign started, more than 90% of Tanzania's population was at risk of malaria, putting the nation in a high malaria burden category. Malaria was also a huge contributor to childhood deaths in Tanzania; there were, on average, around 11 million clinical malaria cases per year prior to the campaign (National Malaria Control Programme, 2010), contributing to about 36% of all deaths in Tanzania in children under five. Despite such a high burden, little effort was taken to reduce malaria transmission. Prior to 2003, the coverage of insecticide-treated nets (ITNs) was nearly zero everywhere in the country.

The Roll Back Malaria Partnership was launched jointly by the WHO, the World Bank, and the United Nations in 1998, aiming to halve the malaria burden between 2000 and 2010. A major difference between the RBM from previous anti-malarial movements was its unprecedented level of external funding - approximately $4.6 billion between 2003–2009. Between 2003 and 2009, 81 of the 108 malaria-endemic received financial support from the global community for their malaria-control work (Johansson et al., 2010). As sub-Saharan Africa accounted for around 85% of the global malaria burden, a large fraction of the financial aid was concentrated in Africa. The strategy the RBM adopted was massive distribution of insecticide-treated nets (ITN) and indoor residual spraying (IRS), both proven methods of reducing malaria transmission. With coordinated actions across countries for a decade, worldwide malaria death were cut in half in 2014. Among the recipient countries, Tanzania provides an ideal setting to study the effects of the RBM campaign as a representative sub-Saharan African country with high malaria burden. It is one of the twelve malaria-endemic countries which received financial support from as early as 2003 (Johansson et al., 2010), and where the RBM campaign was highly successful in reducing the disease burden.

The financial support from the RBM enabled the Tanzanian government to scale up the malaria interventions to the entire country. For example, ITNs started to be distributed to the most vulnerable groups in 2004, and free long-lasting insecticidal nets (LLINs) were delivered to children under five years old starting in 2009. Indoor Residual Spraying (IRS) was introduced to epidemic-prone areas in 2009 (National Malaria Control Programme, 2010). As a result of the coordinated efforts, Tanzania's malaria prevalence had been reduced significantly by 2012, about a decade after the onset of the RBM campaign. Figure 1.2 shows the reduction in malaria prevalence during this period. Malaria prevalence is measured by PfPR_{2-10}, which represents the proportion of children between the age of 2 and 10 who are found to carry *P.falciparum* parasites in their blood. PfPR_{2-10} is a commonly used index to measure malaria prevalence and its transmission intensity, which I continue to use as a proxy of malaria prevalence in the subsequent sections.

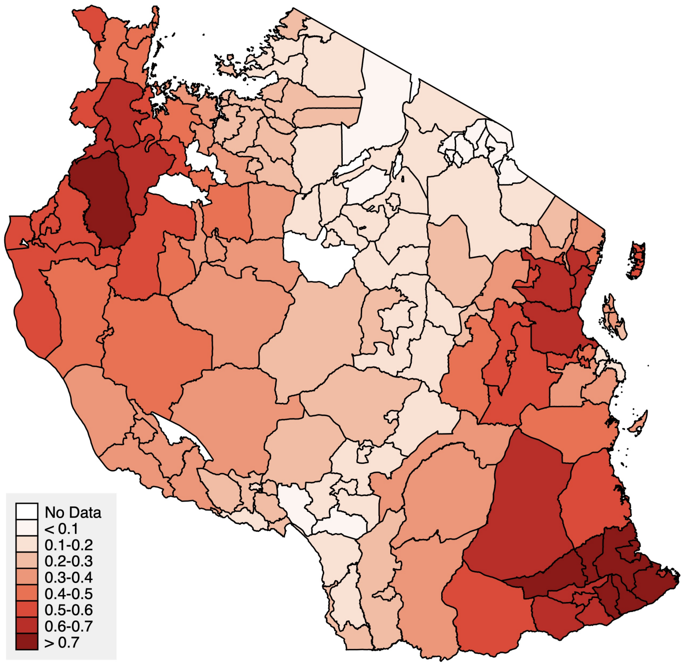
Tanzania is an exemplary country where the RBM significantly reduced the malaria burden, and available microdata also allows us to estimate the effects of the campaign on fertility and child human capital outcomes. The main dataset I am leveraging is the Population Census, which has three waves: 1988, 2002, and 2012. Unlike the commonly used Demographic and Health Survey (DHS) data, the Census data provides a broader range of variables and hence enables us to better identify the effects of the RBM on our outcome variables of interest. In the next subsection, I explain the dataset's structure and advantages over other commonly used datasets.

1.3.2 Data

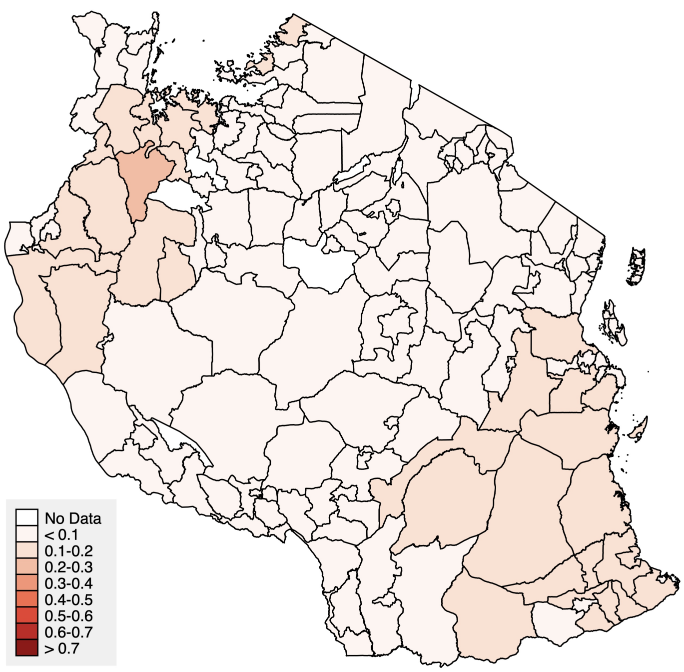
There are two main datasets I use for the empirical analysis. First, the information on malaria prevalence is derived from the Malaria Atlas Project (MAP). Second, household and individual level information on socioeconomic characteristics, mortality, fertility, and parental investment in children's human capital are from the three waves of the Tanzania National Population Census. I describe both datasets separately below.

Malaria Atlas Project (MAP)

I obtain information on malaria prevalence prior to the intervention from the Malaria Atlas Project (MAP). The MAP provides annual estimates of malaria prevalence for a number of sub-Saharan African countries. Within each country, regional estimates are up to second-level administrative units (corresponding to the GIS-2 level). Using this data, I can recover the spatial distribution of malaria risk before and after the RBM campaign. I use the region-level average of the PfPR as a measure of malaria prevalence. Figure A.1 shows the changes in PfPR from 2001



(a) PfPR in 2001



(b) PfPR in 2012

Figure 1.2. Spatial distribution of malaria prevalence rate, pre- and post- campaign

Notes: Geographic boundary is at the level of districts, which are the second level administrative units in Tanzania. Boundaries are harmonized between 1988 and 2012, to account for political boundary changes across census years. Data downloaded from the IPUMS-International (Minnesota Population Center, 2020). PfPR data are taken from Malaria Atlas Project (MAP).

(pre-campaign) and 2012 (post-campaign). As shown in the figure, regions with high malaria prevalence in 2001 experienced a larger reduction in malaria burden after the campaign.

I merge the regional malaria prevalence data from MAP to the Tanzania National Census using the region of residence of the Census households. MAP follows the administrative boundaries announced by the Tanzania National Bureau of Statistics in 2012. In cases where administrative boundaries have changed over time, I harmonized them using the spatially harmonized geographic boundaries between 1988 and 2012 provided by the IPUMS-International.

Tanzania Population Census

Household- and individual-level data on fertility and educational attainment are obtained from the Tanzania National Census, which has three waves: 1988, 2002, and 2012. Three outcome variables are of interest: mortality, fertility, and children's schooling (as a proxy for their human capital). The Census asked each female survey respondent for a complete birth history: the timing and number of every child they had given birth to, including those deceased at the time of the survey. I use the number of children who ever died to measure mortality. For fertility, I use the number of children ever born and the number of surviving children. The former is the number of all births a woman had given, including children who died. Hence, this variable captures the gross fertility. On the other hand, the number of surviving children represents the net fertility. For mortality and fertility, I restrict the sample to women between the age of 35 and 49. Lastly, I used years of schooling, which is the total number of years a child had attended by the time of the survey for the children's human capital.

It is worth noting the advantages of using the Census data over the commonly used Demographic and Health Survey (DHS) data in estimating the effects of anti-malaria campaigns on our outcomes of interest. First, unlike the DHS data, the Census collects not only the regions of residence but also the region of birth. By restricting the sample to those born and residing in the same place, I can control for internal migration, which could be a potential confounding factor in estimation. Second, although the DHS covers many different sub-Saharan African countries, focusing on one specific country alleviates the problems coming from different timing of intervention across countries. Recent econometrics literature has pointed out that when there is heterogeneity in the timing of the treatment across groups, the two-way fixed effects (TWFE) difference-in-differences estimates can be biased (Callaway and Sant'Anna, 2021). Since the RBM campaign began to roll out at the national level in 2003, my estimates are free from such concerns.

1.3.3 Empirical Specification and Identification

The estimation strategy employed here is similar to those used in Wilde et al. (2019) and Kuecken et al. (2021). Since the RBM campaign was targeted toward the area with high malaria prevalence, I exploit the pre-campaign malaria prevalence as a proxy for the campaign intensity. I employ a difference-in-differences model with discrete treatment intensity, where each region falls into one of the four categories representing a different level of pre-campaign level of malaria prevalence. Specifically, I use the PfPR in the year 2001 as a criterion to allocate regions into the treatment category. The four campaign intensity categories are as follows: low prevalence regions (regions with PfPR $< 20\%$ in 2001), medium-low prevalence regions (PfPR between $20-50\%$ in 2001), medium-high prevalence regions (PfPR between $50-75\%$ in 2001), and high prevalence regions with PfPR $> 75\%$ in 2001.[6]

For mortality and fertility, where the dependent variables are count variables (number of children ever born or died), I estimate a Poisson regression model, in which the dependent variable is the logarithm of the expected number of child death/birth experienced by the woman.[7]

[6]The thresholds used for the categorization are consistent with the standard cutoffs used in epidemiological studies of malaria, except that the malariometry literature label a region as low prevalence (hypoendemic) when the prevalence rate is less than 10%. However, the purpose of such classification is to facilitate easier communication between epidemiologist. For the purpose of the empirical analysis, I use the 20% cutoff as a baseline and conduct robustness analysis with different cutoff values.

[7]Poisson regression is a commonly used method in the analysis of survival data. See Lu and Vogl (2023) for an application to the analysis of child mortality.

Specifically, I estimate the following mortality and fertility equations:

$$M_{irt}^{c} = \beta_1^m \text{Post}_t + \sum_{j=2}^{4} \beta_j^m \text{Post}_t \times \text{Prev}_{r,2001}^{j} + \boldsymbol{X}_{ijct}{}' \cdot \boldsymbol{\Gamma} + \eta_r \tag{1.5}$$

$$F_{irt}^{c} = \beta_1^f \text{Post}_t + \sum_{j=2}^{4} \beta_j^f \text{Post}_t \times \text{Prev}_{r,2001}^{j} + \boldsymbol{X}_{ijct}{}' \cdot \boldsymbol{\Gamma} + \eta_r \tag{1.6}$$

where M_{irt}^{c} and F_{irt}^{c} is the logarithm of the expected number of child death or birth experienced by the woman i in age group c at time t, who was born and surveyed in the region r. The variable Post_t is the dummy indicating the pre-and post-treatment. It has a value of one in 2012 and zero in 2002. $\text{Prev}_{r,2001}^{j}$ is the indicator of whether a region r is in the prevalence group j, where $j \in \{2,3,4\}$ with 4 being the highest prevalence regions where PfPR exceeds 75%. Lastly, $\boldsymbol{X}_{ijct}$ is the vector of control variables, which are the age and years of schooling of the respondents and urban-rural residential status. I also include fixed effects for regions in all specifications.

For education outcomes, I estimate the following OLS regression:

$$E_{irt}^{c} = \beta_1^e \text{Post}_t + \sum_{j=2}^{4} \beta_j^e \text{Post}_t \times \text{Prev}_{r,2001}^{j} + \boldsymbol{X}_{ijct}{}' \cdot \boldsymbol{\Gamma} + \eta_r + \varepsilon_{ijt} \tag{1.7}$$

where E_{irt}^{c} is the years of schooling of a child i in age group c at time t, who were born and surveyed in the region r. The rest of the variables are the same as the fertility and mortality models, except that the years of schooling is now excluded from the control variables. Unlike the fertility and mortality regressions, the sample now includes both male and female respondents.

In the empirical analysis, I primarily focus on the the low- and high-prevalence regions as control and treatment groups. As seen in Figure A.2, malaria prevalence was consistently low throughout the 2001-2012 period in the low prevalence regions, and there is a clear reduction of prevalence in the high prevalence regions at the onset of the RBM program in 2003. Although the other regions in the medium-low and medium-high prevalence categories also experienced a reduction in malaria prevalence, it is unclear whether such a reduction was due to the RBM campaign because there is a pre-existing secular decline in the PfPR.[8] Appendix Table A.1 reports the descriptive statistics for the entire sample as well as for the high- and low-prevalence regions.

The main parameter of interest in all specifications is β_4, which is on the interaction between Post_t and $\text{Prev}^4_{r,2001}$. For instance, if the RBM campaign was effective in reducing malaria-related mortality among children and women's fertility in the high malaria prevalence regions, we would expect $\beta_4^{f,m} < 0$. Similarly, if the RBM campaign's effect on children's educational attainment was positive, we would expect $\beta_4^e > 0$.

1.3.4 Results and Interpretation

Child quality: RBM's effects on children's educational attainment

Figure A.5 and A.6 illustrates the parallel trends in years of schooling between the high- and low-prevalence regions from 1998 and 2012, and Table 1.1 presents the results of the regression (1.7). Each column reports the estimate of β_4^e coefficients for different age groups. Columns 1 to 3 correspond to the cohorts of children likely to be affected by the RBM campaign, with descending intensity. For instance, the children between 10 to 15 years old in 2012 (column 1) represent the children born after the RBM campaign, hence likely to be fully benefited from the campaign. On the other hand, the last column represents an older cohort who are not likely to be benefited from the campaign, since these individuals were already beyond the school age and likely to have finished education. In that sense, the column can be interpreted as a placebo group. I only report the coefficients on the interaction of the post-treatment dummy and the indicator for the high-prevalence regions, since I am comparing the lowest and highest prevalence regions as control and treatment groups.

The positive and significant coefficients indicate that the children born in the regions with the highest malaria prevalence in 2001, and hence benefited the most from the RBM campaign,

[8] Results are robust to the exclusion of middle two prevalence groups and available upon request.

Table 1.1. Effects of the RBM on Years of Schooling

Notes: This table reports the estimation restuls from OLS regression (1.7). Brackets contain standard errors clustered at the region level. PfPR_{2-10} (75%+) $\times$ Post indicates the interaction between the indicator of high-prevalence regions (PfPR in 2001 exceeding 0.75) and the post-treatment indicator. Samples are restricted to the individuals who were born and residing (surveyed) in the same region in 2012. Control variables included are age and urban-rural residential status. All columns include region fixed effects. Full table containing the estimates for other prevalence groups can be found in Table A.2. *, **, and *** indicate significance at the 10, 5, 1% levels.

	Age group in 2012			
	Age 10-15	**Age 15-20**	**Age 20-25**	**Age 25-30**
Dependent variable mean in 2012	4.28	6.85	6.68	5.78
PfPR_{2-10} (75%+) $\times$ Post	0.633***	0.974***	0.473***	-0.172
	(0.098)	(0.122)	(0.096)	(0.181)
Observations	1,096,274	856,753	674,743	607,976

experienced an increase in years of schooling.[9] Compared to the low-prevalence regions, children in the highest prevalence regions who were between age 1 and 6 when the campaign began had additional 0.63 years of schooling. Children who were already in school when the campaign started had also benefited from it; those who were between 6 to 11 years old also experienced additional 0.97 years of schooling. However, the positive effects of the RBM dissipate for older age group (age 25-30 in 2012), as those individuals were likely to have completed schooling already when the campaign started in 2003.

The magnitude of the effects is in line with the existing literature's estimates. Lucas (2010) estimate that malaria eradication increased female educational attainment by as much as two years in the most heavily infected region, using Sri Lanka's national malaria eradication campaign in 1945. Bleakley (2010) examine historical episodes of malaria eradication in six Latin American countries and estimate positive effects on children's educational attainment in many countries. More recently, Kuecken et al. (2021) estimate 0.4 years increase in educational attainment among children who were exposed to the RBM campaign.[10]

[9] Results are similar between boys and girls. Appendix Table A.6 contains the results from the same regression run separately by gender.

[10] For a comprehensive review of the literature, see Currie and Vogl (2013)

Child quantity: RBM's effects on mortality and fertility

Figure A.3 and A.4 plot the predicted values of the outcome variables in OLS version of equation (1.5) and (1.6), demonstrating the parallel trends for mortality and fertility conditional on observable control variables. As seen in the figures, both mortality and fertility rates were declining in all regions, but we observe a steeper decline in mortality in the high-prevalence regions upon the launch of the RBM campaign.

Table 1.2. Effects of the RBM on Fertility

Notes: This table reports the estimation results for the Poisson regression (1.5) and (1.6). $PfPR_{2-10}$ (75%+) × Post indicates the interaction between the indicator of high-prevalence regions (PfPR in 2001 exceeding 0.75) and the post-treatment indicator. Samples are restricted to women between age 30 and 49 in 2012, and those who were born and residing (surveyed) in the same region in 2012. Control variables included are age and years of schooling of the respondents and urban-rural residential status. All columns include region fixed effects. Full table containing the estimates for other prevalence groups can be found in Table A.3. *, **, and *** indicate significance at the 10, 5, 1% levels.

Dependent variable	Gross Fertility Children ever born	Mortality Children ever dead	Net Fertility Surviving children
Dependent variable mean in 2012	5.30	0.71	4.57
$PfPR_{2-10}$ (>75%) × Post	–0.0596*** (0.008)	–0.107*** (0.035)	–0.0255 (0.017)
Observations	586,836	586,836	586,836

Table 1.2 reports the coefficients from regression (1.5) and (1.6). Each cell reports the coefficients from the Poisson regression, so the interpretation of the coefficient is the expected change in the log of the mean number of the dependent variable between the high-prevalence regions and the low-prevalence regions. In other words, being in the high-prevalence regions multiplies the mean of the dependent variable by a factor of $\exp(\beta_4)$. For example, when the dependent variable is the number of children ever born, a negative coefficient β_4^f is interpreted that women in the high-prevalence regions reduced fertility in response to the RBM campaign.

Gross Fertility

The first column summarizes the results for gross fertility, where the number of children ever born was used as a dependent variable. The coefficient β_4^f shows that childbearing-age women in the high-prevalence regions reduced their fertility by 5.96%. The negative response of fertility is consistent with Kuecken et al. (2021), which also found a reduced probability of childbirths in response to the RBM campaign in multiple sub-Saharan African countries. Within the framework of the quantity-quality tradeoff, the negative fertility response indicates that the drop in the price of child *quality* induced by the RBM campaign was larger than the drop in the price of child *quantity*.

Child Mortality

Although the RBM campaign led the women in the high-prevalence regions to decrease their fertility, it does not immediately imply that the *net* fertility rate declined because the RBM campaign also lowered child mortality. If the reduction in gross fertility is smaller than the drop in child mortality, net fertility can still increase in response to the campaign. The second column of Table 1.2 shows to what extent the RBM campaign reduced the number of deaths among children in the high-prevalence regions. The coefficient β_4^m shows the RBM campaign reduced the number of child deaths per woman by 10.7%, indicating that the RBM campaign was indeed effective for reducing child mortality.[11]

Net Fertility

Combining the results from the gross fertility and child mortality regressions together, the last column shows the effects of the RBM campaign on *net* fertility. Instead of the number of children ever born, the number of surviving children was used as the dependent variable, which accounts for the immature deaths of children. Although the coefficient is negative, it is not statistically significant, indicating that although the RBM campaign induced the women have

[11] The results are consistent with Wilde et al. (2019) and Kuecken et al. (2021), which studied the effects of the RBM campaign on child mortality as well and found that the campaign reduced all-cause child mortality.

fewer children, it did not reduce the number of *surviving* children, at least after the ten years since the campaign started.

One interpretation for the muted response of net fertility is that the RBM campaign did not change the number of overall births women intended to have. In this interpretation, women reduced their fertility in response to the lowered mortality so that the number of surviving children remained the same. This interpretation can also explain the increase in children's educational attainment since having fewer children would have allowed parents to increase educational input in each child.

Another interpretation is a possibility that the RBM induced women to delay their fertility. Although possible, such an interpretation is not likely to change the conclusion that the net fertility did not change in response to the RBM campaign. The chance of pregnancy declines with age, so even if these women are delaying their fertility and planning to have more children at older ages, it is unlikely that it will lead to an increase in net fertility. Therefore, I conclude that the RBM campaign reduced gross fertility but did not change net fertility, although further investigation with extended data is needed.

Summary of Empirical Analysis

In summary, the RBM campaign is found to reduced women's fertility and child mortality by 5.9% and 10.7%, respectively. The campaign did not change the number of surviving children, suggesting that women reduced their fertility so that the number of surviving children remain constant.[12] The campaign also increased the years of schooling of the benefited children by 0.63 years. In the next section, I use these empirical estimates to discipline the model's quantitative mechanism.

[12]The muted response of net fertility is also consistent with the macroeconomic literature of demographic transition, which has been arguing that while a decline in child mortality is key for a decline in *gross* fertility, but is not sufficient to drive a decline in *net* fertility (Doepke, 2005).

1.4 Model Parameterization and Quantification

Guided by the empirical evidence in the previous section, I discipline the magnitude of the model's key mechanism by replicating the RBM campaign within the model and matching the empirical moments obtained in the previous section. The empirical analysis suggests that the reduction in malaria burden from the RBM campaign has induced the parents to increase educational investment to their children by sending them more to school, but the parents adjusted the fertility so that the net fertility does not change. I use these two moments which summarizes the effects of malaria on fertility behavior and children's human capital: the muted response of net fertility in (1.6) and increased educational attainment of children (β_4^e) in (1.7). To ensure the model is credible in other dimensions, I jointly estimate parameters to match the empirical elasticities and other relevant aggregate moments of the Tanzanian economy, such as educational attainment and intergenerational mobility measures. These moments are calculated from the microdata when possible, or taken from the existing literature otherwise.

1.4.1 Exogenously chosen parameters

A small set of parameters are chosen exogenously. Such parameters are summarized in Table 1.3. The exogenously chosen parameters can be classified into two broad categories. The first one is a set of parameters that are standard in the macroeconomics literature. Others are the parameters related to malaria, which are taken from epidemiological/health studies.

Standard Parameters

The first panel of Table 1.3 shows the standard parameter values. The discount factor is chosen to be 0.96^6, consistent with the typical values in the literature, adjusted for the fact that the model period corresponds to six years. The exogenous gross interest rate is set to be 2 percent per year, representing the low financial access and low savings rate among households in low-income economies (Donovan, 2021).

It is worth mentioning the role of parameter γ, whose interpretation is the inverse of

Table 1.3. Exogenously Chosen Parameters

Var	Description	Value	Source / Target
Economic Parameters			
β	Discount factor	0.96^6	Standard value
r	Exogenous gross interest rate	1.02^6	Deposit interest rate
γ	Inverse of IES	0.5	Daruich and Kozlowski (2020)
ω_2	Curvature of the time cost function	0.68	Folbre (2008)
$\bar{N}$	Maximum number of children	6	Tanzania Census 2002
λ	Elasticity of substitution between skill groups	4.0	Bils et al. (2022)
Epidemiological Parameters			
χ_U^d	Mortality rate, child from uneducated parents	0.109	Ogbo et al. (2019)
χ_P^d	Mortality rate, child from primary-educated parents	0.109	Ogbo et al. (2019)
χ_S^d	Mortality rate, child from secondary-educated parents	0.078	Ogbo et al. (2019)
χ_U^m	Prob. catching malaria, child from uneducated parents	0.79	Gonçalves et al. (2014)
χ_P^m	Prob. catching malaria, child from primary-educated parents	0.79	Gonçalves et al. (2014)
χ_S^m	Prob. catching malaria, child from secondary-educated parents	0.56	Gonçalves et al. (2014)

the intergenerational elasticity of substitution. This parameter is commonly interpreted as *intertemporal* elasticity of substitution in the macroeconomics literature, which governs the consumption tradeoffs between today's and tomorrow's consumption. In the life-cycle model with inter-generational linkage, γ also governs the degree of *intergenerational* elasticity of substitution because this parameter also affects how parents weigh their children's utility to their own utility. For instance, higher γ means that parents' marginal utility of consumption decreases faster as they become richer, which makes children relatively more valuable. I set the value of γ to 0.5, following Daruich and Kozlowski (2020). The fact that γ is less than one also ensures the utility is positive everywhere. That is, parents always "enjoy" having children, and the implicit value of being childless or having a child die is zero.[13]

Fertility and Childcare Cost

The maximum number of children parents can choose in the model is capped at $\bar{N} = 6$. Since a parent in the model corresponds to a household of two parents, a parent choosing

[13]When the utility is allowed to be negative, we need extra assumptions on the value of being childless and having children die. See Jones and Schoonbroodt (2010) for discussion on this matter.

$n = \bar{N}$ corresponds to a household having 12 children in reality. I choose $\bar{N} = 6$ because according to the Tanzania Census in 2002, 95% of the households have less than 12 children. The time cost function, $t(n) = \omega_1 n^{\omega_2}$, is an increasing and concave in the number of children. Typically, the literature estimates both the level and curvature parameters using the time-use data[14]. Unfortunately, detailed time-use data is not available in most low-income countries, and Tanzania is not an exception. Absent the available data, I follow the literature and set the value of the curvature parameter, ω_2 at 0.68, based on Table 6.4 in Folbre (2008).

Malaria

Parameters related to child mortality and malaria are taken directly from epidemiological studies of malaria in the Sub-saharan Africa context. Two sets of parameters are needed: mortality (probabilities of deaths) and morbidity (probability of being sick from malaria). The mortality parameter, χ_s^d, corresponds to the under-age five mortality rate. I take this parameter from Figure 1 of Ogbo et al. (2019), which estimated the under-5 mortality rate in Tanzania between 2004 and 2016 using the Demographic and Health Survey. I use the estimates from the 2004-2005 DHS wave, which is the closest to the pre-RBM periods. Ogbo et al. (2019) also reports that children under five years whose mothers had primary or no education were 38% more likely to die before their fifth birthday compared to those whose mothers had a secondary or higher level of education (Table 2). Following their estimates, I set the mortality rate of children born from secondary-educated parents to be 38% lower than its lower-educated parents counterparts.

The parameter χ_s^d encompasses the probability of dying from all causes, including causes other than malaria. To calculate the fraction of the mortality rate attributed to malaria, I use the four waves of the Tanzania National Panel Survey (2008, 2010, 2012, and 2014) to calculate the fraction of malaria-caused deaths among all deaths. The survey asks the diagnosed cause of death for all deaths within the households, and malaria consistently accounts for about 51%

[14]See Lee and Seshadri (2019); Daruich and Kozlowski (2020), and Yum (2023) for the estimates of ω_2 from the US time use data.

of childhood deaths[15]. However, this number is likely to be an upper bound since only 5.9% of all deaths are identified, meaning that a lot of parents did not know which disease caused their child's death. If malaria is relatively easier to identify as a cause of death, then such a high fraction might be an overestimation. Instead, the Institute for Health Metrics and Evaluation (IHME) reports that about 17% of under-5 deaths were attributable to malaria in 2002 in Tanzania. I take the average of the two numbers and assume that 35% of all under-5 deaths are caused by malaria.

Parameters for morbidity probabilities, χ_d^m, are taken from epidemiological studies in Tanzania. From a study conducted between 2002 and 2005 among children in areas with intense malaria transmission in Tanzania, Gonçalves et al. (2014) concludes that the unconditional probability of experiencing malaria (either mild or severe) is 79%. Following this, I set 79.0% as the baseline probability of being hit by a morbidity shock in the model. I adjust the probability to 56% for children born from secondary-educated parents, as I did for mortality.

1.4.2 The RBM Campaign within the Context of the Model

I use the causal effects of the RBM campaign on fertility and children's educational attainment estimated in Section 1.3 to discipline the model's implied elasticities. To do so, I simulate the RBM campaign within the model and compare the model-generated moments to the estimated data moments. In this subsection, I describe how I interpret the RBM campaign from the perspective of the model and the procedure of the simulated method of moments.

I calibrate the baseline economy before the RBM campaign to the high-prevalence regions. When the region-specific parameters of moments are not available due to the data availability, I use the country-wide counterparts.[16] I then feed the RBM campaign to the model as an unexpected, universal reduction in the probability of negative health shocks. Compared to the pre-campaign periods, the malaria prevalence level was reduced by 70% (Figure A.2).

[15]Exact numbers, along with the numbers for other causes of deaths, are reported in Table A.4.

[16]The high-prevalence regions are relatively poorer than the rest of Tanzania and have smaller family sizes. However, the relationship between fertility, income, and parental investment is similar to the rest of the country.

Guided by this fact, I lower the morbidity probability χ_s^m and the malaria-attributed part of the mortality probability χ_s^d by 70% for all schooling groups s. Since the RBM was a nationwide campaign, all households in the model face the same reduction in malaria risk.

I also assure that the households in the economy do not expect a reduction in malaria risk and behave in anticipation. This approach prevents behaviors like accumulating assets in advance to increase educational investment when the campaign starts. To do so, I first simulate the balanced growth path of the economy with the pre-RBM level of malaria risks and then reduce the malaria risk unexpectedly. Households then adjust their fertility and education decisions according to the new decision rules under the changed disease environment. The simulated moments and their data counterparts are summarized in the last panel of Table 1.4, and the estimated parameter values are summarized in Table 1.5.

1.4.3 Parameters Estimated from the RBM Campaign

Aside from the exogenously calibrated parameters, the remaining 12 parameters of the model are calibrated to match the 14 moments, either from the RBM campaign or from the aggregate data. The two main outcomes of the RBM campaign are its treatment effects on fertility and children's education, measured by the number of children and years of schooling. Although all moments in the model are jointly determined, the parameter $\underline{m}$, which represents the negative effects of morbidity shock on returns from education, is most closely related to the two moments. The lower the $\underline{m}$ is, the more detrimental malaria is for a child's human capital formation. Hence, reducing the probability of malaria will have greater effects with lower $\underline{m}$. The estimated value of $\underline{m}$ is 0.76, suggesting that children with malaria receive 24% lower returns from schooling. This is broadly consistent with the epidemiological literature on malaria's effects on children's cognitive ability and school performance. For instance, in a study carried out in Sri Lanka, Fernando et al. (2003) find that children who experienced less than three attacks of malaria scored at least 15% more in both the special and school examinations than children who experienced more than five attacks of malaria during the same period. In another study that took place in Yemen, Al Serouri et al. (2000) conclude that having at least one attack of malaria was significantly associated with poor (below the 50th percentile for class/grade) school performance.

Table 1.4. Targeted Moments

Moments	Source	Data	Model
Education			
Primary completion rate (%)	Tanzania Census 2012	45.6	44.8
Secondary completion rate (%)	Tanzania Census 2012	11.5	10.3
Primary ed. wage premium (%)	Leyaro et al. (2014)	59.9	59.4
Secondary ed. wage premium (%)	Leyaro et al. (2014)	115.2	121.8
Intergenerational Mobility			
Uneducated-Primary intergenerational upward mobility	Alesina et al. (2021)	46.9	39.9
Uneducated-Secondary intergenerational upward mobility	Alesina et al. (2021)	4.7	3.3
Primary-Secondary intergenerational upward mobility	Alesina et al. (2021)	14.0	15.8
Differential Fertility			
Total fertility rate, uneducated parents	Tanzania Census 2012	5.00	4.94
Total fertility rate, primary completed parents	Tanzania Census 2012	4.67	4.56
Total fertility rate, secondary completed parents	Tanzania Census 2012	3.40	3.88
% of secondary-educated parents with 6+ children	Tanzania Census 2012	23.35	22.76
Inequality			
Gini coefficient	Younger et al. (2016)	0.38	0.42
Roll Back Malaria			
Treatment effect of the RBM on schooling (years)	Section 1.3	0.63	0.66
Treatment effect of the RBM on net fertility (%)	Section 1.3	0.00	0.01

1.4.4 Parameters Estimated from Aggregate Data

The remaining parameters jointly determine the targeted moments from aggregate data, as well as the moments from the RBM campaign. Primary and secondary school completion rates are calculated from the 2012 Census among adults younger than 30 years old. The low completion rates for both primary and secondary schools inform that child labor is quite prevalent. The two schooling fee parameters, p_P and p_S, pin down these two moments. The estimated parameter value of the primary school fee parameter p_P is -0.36. The negative sign of this parameter suggests it is unlikely that the credit constraints are preventing the parents from sending their children to primary school because a negative p_P means parents receive money by sending kids to school. Instead, it means that child labor is an attractive alternative to educating children. On the other hand, the schooling fee for the secondary school, p_S, is positive at 1.18, implying a monetary cost of sending children to secondary school.

Table 1.5. Parameters and Estimated Values

Parameter	Value	Description
η_P	1.33	Human capital gain from primary education
η_S	1.64	Human capital gain from secondary education
p_P	-0.36	Goods cost of primary education
p_S	1.18	Goods cost of secondary education
σ_v	0.16	Standard deviation of idiosyncratic income shock
σ_z	0.16	Standard deviation of the learning ability draw
ρ_z	0.92	Intergenerational persistence of learning ability
θ	0.39	Gumbel scale parameter of the fertility taste shock
ω_1	0.17	Level of time cost of childcare
δ	0.75	Level of intergenerational altruism
λ_n	0.72	Curvature of the altruism function
$\underline{m}$	0.76	Severity of malaria morbidity shock

The earnings premiums that primary- and secondary-educated workers enjoy (relative to the uneducated group), are taken from Table 3 of Leyaro et al. (2014), which estimated the education premiums with two the 2001/2006 Tanzania Integrated Labour Force Survey.[17] The two parameters related to these moments are η_P and η_S, which governs how much human capital grows from attending primary and secondary school, respectively. The estimated parameters are 1.33 and 1.64. Along with the low secondary school completion rate, the large η_S shape the sizable earnings premium the secondary-educated workers enjoy.

There are two parameters that govern intergenerational mobility in the model. The first is the AR(1) persistence parameters of the learning ability draw, ρ_z, and the other is the variance

[17]There are multiple sources for wage premium, all reporting similar estimates. Leyaro et al. (2014) use both IFS and the urban worker sample. Joseph (2020) used the Integrated Labor Force Survey, using primary educated workers as a reference group. Mlacha and Ndanshau (2018) also use the integrated labor force survey and get similar estimates.

of the shock in the AR(1) process, σ_z. Intuitively, a high level of persistence would reduce intergenerational mobility, while a high level of variance of the shock increases intergenerational mobility by making the learning ability more random. Given this intuition, I target three measures of intergenerational upward mobility, estimated in Alesina et al. (2021). Each moment represents the likelihood that children born to parents that have a certain level of education manage to attain a higher level of education. In the model, these moments are constructed from parent-children pairs from the balanced growth path of the model.

Several parameters jointly affect the fertility behaviors in the model. First, there are two parameters in the intergenerational discount function, $b(n) = \delta n^{\lambda_n}$, which govern the level of intergenerational altruism and its curvature to the number of children. These parameters are closely related to fertility decisions, although they also affect education decisions through intergenerational altruism. The estimated parameter values for δ, and λ_n are 0.75 and 0.72, matching the overall fertility rate. Another parameter that governs fertility behavior is ω_1, which represents the time cost of child-rearing. If the time cost is high, then high-skilled parents would be more reluctant to have children, generating a stronger negative relationship between income and fertility. I hence use ω_1 to match the differential fertility across the three skill groups. The estimated value of ω_1 is 0.17, and the fertility declines with the parental school level. Lastly, I use the Gumbel scale parameter of the fertility taste shock θ to match the distribution of fertility at the tail.

The remaining parameter σ_v governs the standard deviation of the idiosyncratic shock to adult human capital. Given that this is also related to the variance of adult income, I discipline this parameter by matching the Gini coefficient in the year 2010, measured in Younger et al. (2016) using the Tanzania Household Budget Survey.

1.5 Long-Run General Equilibrium Effects of Malaria Vaccine

Using the estimated model, I simulate the long-run effects of a nationwide distribution of a successful malaria vaccine. A recent epidemiological study reports that the new malaria vaccine is up to 80% effective at preventing the disease in young children. Through the lens of the model, this is interpreted as an 80% reduction in malaria mortality (χ_s^d) and morbidity shock probability (χ_s^m). Since malaria accounts for 35% of all under five mortality in the model, the resulting reduction in mortality rate is sizable. To see the long-run effects, I simulate the new balanced growth path of the economy with the reduced malaria parameters and compare the aggregate moments between the pre-and post-vaccine balanced growth paths.

1.5.1 Quantitative Results

Table 1.6 summarizes the long-run, general equilibrium effects of a nationwide distribution of a successful malaria vaccine. The primary school completion rate increases significantly, from 45 percent in the pre-vaccine balanced growth path to almost 60 percent in the post-vaccine balanced growth path. The increase in the secondary school completion rate is relatively smaller. The smaller increase in secondary completion rate stems from the higher returns from child labor from secondary education-aged children. The higher school completion rates mean that the parents on the post-vaccine balanced growth path possess higher human capital. As a result, the negative income-fertility relationship becomes stronger. As seen in the second panel, fertility (the number of children ever born) falls in all skill groups. What is notable here is that despite the falls in gross fertility, the population growth rate *increases* in the post-vaccine balanced growth path, although the change is modest. The higher population growth implies that the decline in gross fertility rates is insufficient to compensate for the drop in mortality rate, as argued in Acemoglu and Johnson (2007).

The second panel of Table 1.6 demonstrates the changes in the skilled and unskilled wages

Table 1.6. Long-Run General Equilibrium Effects of National Malaria Vaccine Policy

	Pre-Vaccine BGP	**Post-Vaccine BGP**	**Change from Pre-Vaccine BGP**
Education			
Primary completion rate (%)	44.8	58.3	+ 12.4 p.p.
Secondary completion rate (%)	10.3	14.4	+ 4.1 p.p.
Output per Capita and Prices			
Unskilled wage			+ 4.5%
Skilled wage			− 8.9%
Output per capita			+ 34.3%
Fertility			
Total fertility rate, uneducated parents	4.94	4.91	− 0.7%
Total fertility rate, primary completed parents	4.56	4.49	− 1.7%
Total fertility rate, secondary completed parents	3.88	3.77	− 2.8%
Population growth rate (%)	3.13	3.15	+ 2.0p.p.
Intergenerational Mobility			
Uneducated-Primary intergenerational upward mobility	39.9	51.6	+ 13.5 p.p.
Uneducated-Secondary intergenerational upward mobility	3.3	4.9	+ 1.6 p.p.
Primary-Secondary intergenerational upward mobility	15.8	18.1	+ 2.3 p.p.

Notes: This table reports the long-run, general equilibrium effects of nationwide malaria vaccination. The efficacy of the vaccine is assumed to be 80%.

and the increase in GDP per capita. GDP per capita increases significantly, up to 34% in the long run. Given that the population growth rate is also higher in the post-vaccine balanced growth path, such a high increase in GDP per capita suggests that the increase in aggregate production is more than enough to compensate for the larger population size. In the next subsection, I conduct several decomposition exercises to break down the sources of the GDP per capita gain. Lastly, the relative wage of unskilled workers rises by 4.4 percent while the skilled wage falls by 8.8 percent, reflecting the larger supply of skilled workers in the post-vaccine world.[18]

The last panel of Table 1.6 reports the changes in intergenerational upward mobility in terms of educational attainment. The fraction of children born to uneducated parents who manage to receive primary education increases by 12 percentage points. Upward mobility measures for

[18]The rise in unskilled wage and fall in skilled wage is consistent with Khanna (Forthcoming), who finds a large decline in the relative wages of skilled workers following an expansion in schooling in India.

higher education levels also increase, but with smaller magnitudes. The model implies that the improvement in intergenerational upward mobility is larger in households with less educated parents, even though the vaccine was universally distributed. Mass vaccination is pro-poor because when a child has malaria, unskilled parents are less likely to send the child to schools than skilled parents. Figure 1.3 illustrates the probability that a healthy (no malaria shock) and an unhealthy (malaria shock) attend primary school, depending on their parent's education level, after controlling for the number of children within the household. Two patterns are observed here. First, in all three groups of education level, parents are more likely to send their children to school. In other words, parents tend to *reinforce* the adverse effects of the malaria shock on their child's human capital by not sending them to school. The second pattern is the heterogeneity of the reinforcing behaviors across different groups of education levels. Although parents from all educational backgrounds are less likely to send their children to school when they are sick, it is more so for the less educated parents. In other words, more educated parents are likely to *compensate* for the negative effects of malaria on children's human capital by sending them to school more, compared to the less educated parents.[19] As a consequence, the model predicts that vaccination is pro-poor by benefiting the children from less educated parents more.

1.5.2 Sources of the Long Run Gains

The 34% increase in long-run GDP per capita is much larger than what previous literature has found. In this subsection, I formally compare this number to the numbers found in the previous literature and investigate the sources of the model's predicted large increase in per capita GDP.

Comparison to Acemoglu and Johnson (2007)

Acemoglu and Johnson (2007) examined the effect of longer life expectancy on economic growth by exploiting large the improvements in life expectancy driven by international health

[19] The terms ”reinforce” and ”compensate” are commonly used in the literature, but more commonly in a context of intra-household allocation of resources between children with different abilities or health conditions.

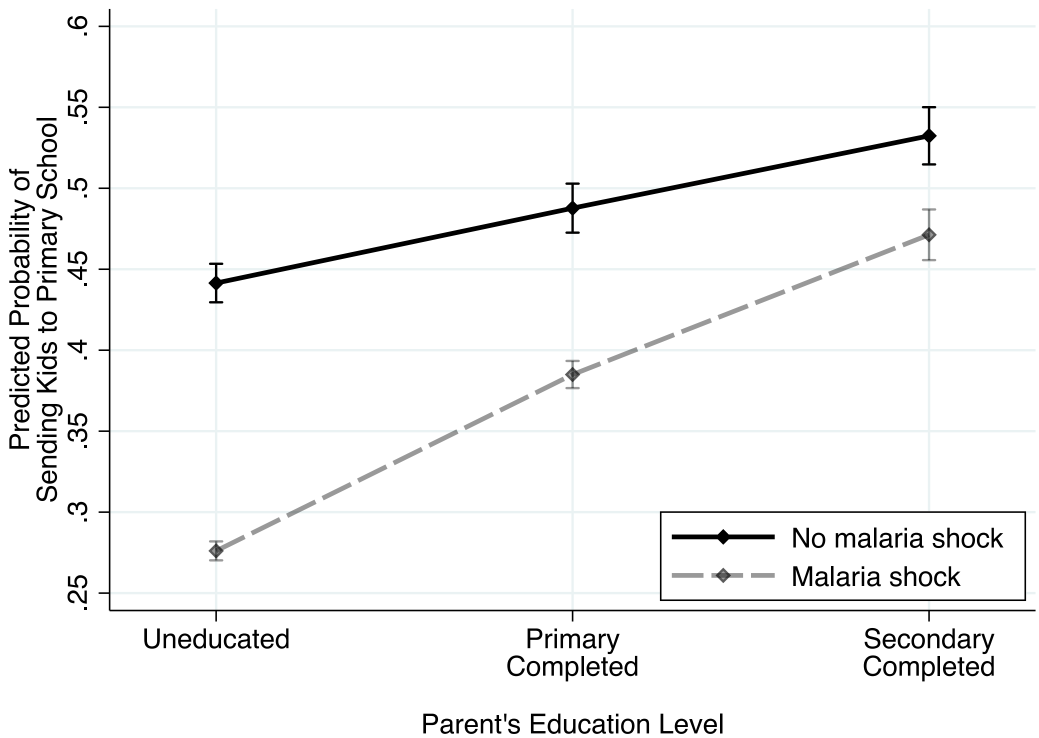


Figure 1.3. Child's Probability of Attending Primary School

Notes: This figure illustrates the likelihood that children born to parents with different levels of educational attainment complete primary schooling, depending on their health status at age six on the pre-vaccination balanced growth path. I simulated 50,000 parent-children pairs from the balanced growth path distribution and ran probit regression where the dependent variable is a binary variable of whether a parent sends children to primary school. Control variables include the idiosyncratic productivity shock, the number of children within the household, and children's learning abilities. Regression coefficients are plotted with 95% confidence intervals.

interventions in the 1940s, namely the international epidemiological transition. They found that a 1% improvement in life expectancy leads to a 1.7–2% increase in population but found no evidence of growth in per capita income following a substantial increase in life expectancy. They even found relative *decline* in GDP per capita in countries that experienced large increases in life expectancy, suggesting that longer life expectancy contributes to population growth rather than improvement in economic growth.

The small or even negative changes in per capita output in Acemoglu and Johnson (2007) are seemingly contradictory to the model's large long-run increase in per capita output.

However, their findings are perfectly consistent with the model's predictions once we interpret the international epidemiological transition through the lens of the model. Generally speaking, the model says whether eradicating diseases can decrease the long-run net fertility depends on if the eradication increases the children's human capital. In this sense, if the international epidemiological transition only lengthened the life expectancy and did not facilitate children's human capital accumulation, then simulating the transition within the model would generate the similar results as in Acemoglu and Johnson (2007).[20]

Table 1.7. Decompositon of the Long-Run Effects

Notes: This table shows the long-run changes in educational attainment and per-capita output when mortality and/or morbidity are lowered. The second column contains the results from a simulation with an 80% reduction in malaria mortality (χ_s^d) while the morbidity shock probability (χ_s^m) is unchanged. The third column contains the results from a simulation with no change in malaria mortality while the morbidity shock probability is reduced by 80%. The first column is the baseline long-run simulation, where both mortality and morbidity probabilities are lowered.

	Lower Both	**Lower Mortality**	**Lower Morbidity**
Population Growth Rate (%)	3.15	3.26	3.02
Primary Completion Rate (%)	58.3	46.0	58.5
Secondary Completion Rate (%)	14.4	11.0	14.7
Δ Per capita GDP	+ 34.3%	+ 1.6%	+ 31.7%

To confirm this idea, I simulate a counterfactual long-run balanced growth path where I only lower either the mortality or morbidity to the post-vaccine level while keeping the other dimension of the disease at the pre-vaccine level. Table 1.7 contains the results from this counterfactual exercise. Although this is not the perfect way to replicate the epidemiological transition within the model, it is an informative exercise to check whether the model can nest the empirical observations in the previous literature. The second column is the closest to the long-run effects of the epidemiological transition; fewer children die, but their returns from schooling are

[20]There are several reasons to believe that the epidemiological transition did little in increasing children's human capital. The most obvious observation is that schooling was not universal back in the 1940s, and many children were not able to take advantage of better health conditions by attending schools. Child labor was also more prevalent due to the lack of legal restrictions against it.

still low because morbidity shock is unchanged. The per capita output gain shrinks to a mere 1.6 percent. Moreover, the increase in primary school completion rate becomes very modest, compared to 45% in the pre-vaccine balanced growth path. Most importantly, a large increase in population growth rate (+ 0.13 p.p.compared to the benchmark post-vaccine balanced growth path) suggests that any individual-level income gains are dwarfed by the larger population size, resulting in a negative per capita income gain.

The key for a public health intervention to have a growth impact is whether it can facilitate education. Overall, the calibrated model suggests that even a significant reduction in mortality is not enough to generate a transformative increase in education when it's not accompanied by a reduction in morbidity. A key takeaway here is that uni-dimensional, mortality-focused modeling of disease could lead to biased results when calculating the economic effects.

Comparison to Ashraf, Lester, and Weil (2008)

Ashraf et al. (2008) simulate the effects of eradicating malaria using a standard neoclassical framework. In their framework, eradicating malaria raises life expectancy and years of schooling. Hence, they considered both the mortality and morbidity aspects of malaria. Nevertheless, they conclude that the effect of eradicating malaria in a typical sub-Saharan African country would be to raise GDP per capita by only about 2% in the long run.

There are two main differences between the simulation in this paper and that of Ashraf et al. (2008). The first one is the channels through which disease eradication (or reduction of disease prevalence in the current context) affects the human capital. Both assume that schooling increases human capital, but Ashraf et al. (2008) assume that eradicating malaria increases human capital only through increased years of schooling. This approach does not capture the channels other than years of schooling through which malaria eradication increase children's human capital, such as better learning in school due to higher cognitive ability, or better physical human capital.[21] The measurement of human capital in the model capture overcomes this

[21] Bleakley (2010), for example, uses the malaria eradication in six Latin American countries circa 1955 as a

weakness because, in the model, the malaria vaccine increases human capital not only through more schooling (more children sent to primary and secondary school) but also through improved learning in school.

The second difference of Ashraf et al. (2008) from the quantitative exercise of this paper is the lack of intergenerational dynamics. Ashraf et al. (2008) treat malaria eradication as a one-time increase in human capital, which does not affect future generations. However, it is debatable that eradicating malaria only affects one generation, without inducing the subsequent generations to change their education decisions for their children. If the children who benefited from eradication further increase educational investment for their own children, the increase in per capita output will be even larger in the long run. The overlapping generational structure and the intergenerational linkage embedded in the model allow us to capture this channel and investigate its quantitative importance.

Table 1.8. Decomposition of the Short-Run Effects of the Vaccination

Notes: This table shows the short-run, one-generational effects of the malaria vaccine on children's educational attainment and earnings in the first period of adulthood. The numbers in the second column are calculated from a simulation where I do not allow parents to make endogenous fertility choices and assign the number of children that corresponds to the pre-vaccine balanced growth path. Parents still make education investments choice in this case. For the numbers in the third column, I do not allow parents to change their educational investment choice in response to the vaccination while letting them make fertility choices.

	Benchmark Model	**Exogenous Fertility**	**Exogenous Schooling**
Children born right after the vaccination			
Primary completion rate (%)	55.2	55.1	44.8
Secondary completion rate (%)	11.0	10.9	9.4
Earnings at age 18 vs. parents	+ 18.0%	+ 17.6%	+ 12.8%

To understand the quantitative implications of the two departures mentioned above, I calculate the short-run effects of the malaria vaccine from two alternative scenarios where parents

natural experiment and studies how much adulthood earnings increased for the children who were born after the eradication. The study finds that 100% eradication of malarial infections in those countries increased subsequent adult income by >40%, and an increase in years of schooling only accounts for 25% of the earnings increase. This suggests that malaria eradication might have partially affected people's income through channels other than years of schooling.

are not adjusting fertility or schooling decisions. Specifically, in each scenario, I do not allow the parents to have fewer children (exogenous fertility) or send children to school more (exogenous schooling). The second and third columns of Table 1.8 summarize the results of this exercise.

First, comparing the benchmark model where both fertility and schooling decisions are endogenous (column 1) manifests the importance of the amplifying effects of the intergenerational linkage. Comparing the 18% short-run partial equilibrium increase in per capita output with 34% of the long-run general-equilibrium counterpart reveals that the long-run general equilibrium effects are about 1.9 times larger. This implies that ignoring the intergenerational compounding effects could potentially underestimate the long-run gains by almost 50%.

Second, comparing the benchmark simulation (column 1) and the exogenous schooling scenario (column 3) uncovers the quantitative importance of using years of schooling as the sole measure of human capital. The model predicts that even if we did not observe any increase in years of schooling among the children who receive the vaccines, their earnings in adulthood would have increased by 12.8 percent. In other words, an increase in years of schooling only accounts for 29% of the total increase in human capital. This is consistent with Bleakley (2010), who found that cohorts born after malaria eradication had higher adult income in six American countries. Of all six countries, years of schooling accounted for only less than a quarter of the effect of income. The model implies that measuring human capital through years of schooling would significantly underestimate the effects of malaria on human capital.

Combined together, we can do a back-of-the-envelope calculation of how much the long-run gains in output per capita will be reduced when we mimic the analysis in Ashraf et al. (2008). Since years of schooling accounts for only 29% of the human capital gain and intergenerational channel amplifies the short-run effects by 1.75 times, the long-run increase in per capita output in this scenario would be $34 \times 0.29 \times (1/1.9) \approx 5.2$ percent. This number is much closer to the 2 percent long-run gain that Ashraf et al. (2008) found.

1.5.3 Policy Evaluation: Vaccine Efficacy and the Cost of Vaccination

Although the increase in output per capita is large in the long run, producing and administering vaccines to the mass population can be quite costly. For instance, Sicuri et al. (2019) estimate the per-child cost of administering malaria vaccine to range from \$25 to \$37 in a typical sub-Saharan African country, including all associated costs and assuming a vaccine price of \$5 per dose.[22] Moreover, the vaccine's efficacy might turn out to be lower than the current 80%, and if so, the costs of vaccinating people might exceed the benefits.

To see whether the long-run benefits are large enough to compensate for the costs of the vaccination policy, I solve the post-vaccine balanced growth path with a lower level of efficacy and compare the per-capita output gain to the cost of per-child vaccination. I take a conservative stance on the cost of vaccination and assume that per-child cost is \$50, which roughly corresponds to a vaccine price of \$10 per dose. Figure 1.4 shows the cost-benefit comparison of various efficacy scenarios. The vertical line at the 75% efficacy level represents the WHO-specified efficacy threshold for funding. The horizontal line indicates the \$50 per-child cost of vaccination as a fraction of GDP per capita, using the GDP per capita in 2005 as a reference.

As seen in the figure, the long-run increase in output per capita clearly dominates the cost of vaccination, even though the costs used were conservatively high. Considering that the long-run gains are about twice as large as the gains for the first generation who would receive the vaccines, the results imply that any vaccines with at least 40% efficacy would be cost-efficient, which is far below the WHO-specified efficacy level for financial support.

1.6 Conclusion

High mortality and poor health conditions have been widely thought of as major development obstacles by policymakers in the developing world. Yet the macroeconomics literature thus

[22] According to Datoo et al. (2022), four doses (three initial doses and a booster dose one year later) are needed to fully vaccinate one child.

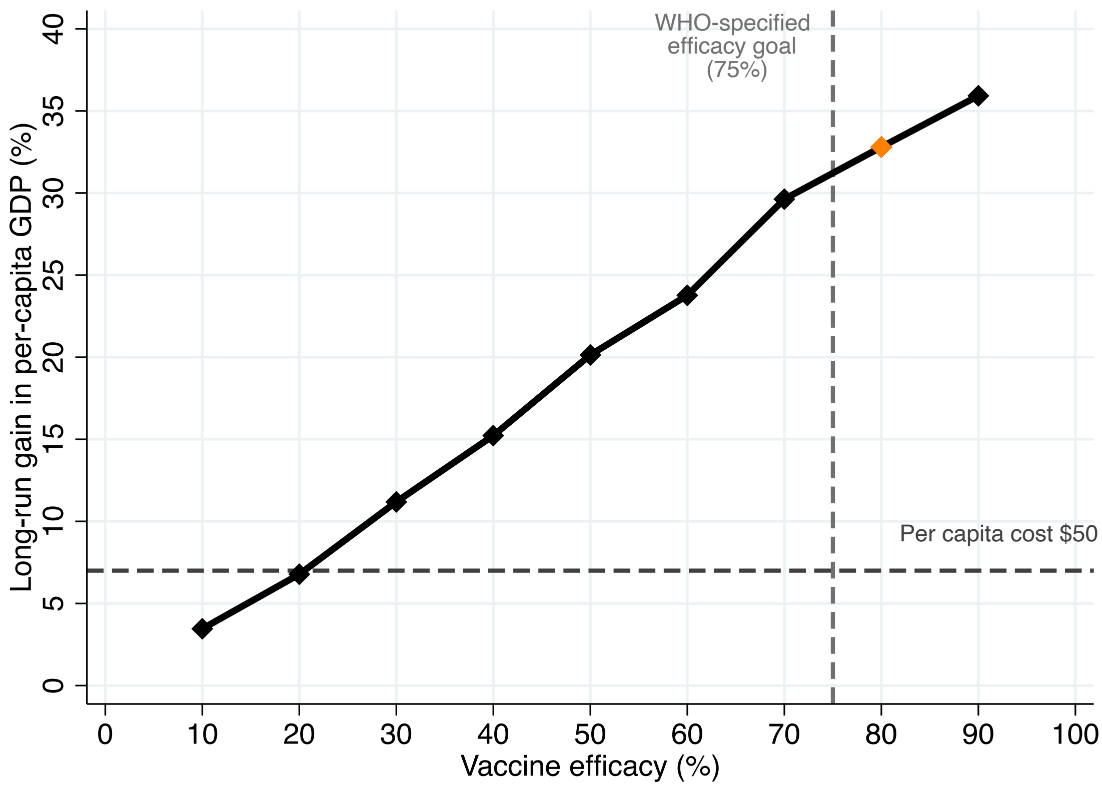


Figure 1.4. Vaccine Efficacy and the Cost of Vaccination

Note: Y-axis is the long-run percentage change in per-capita output between the pre-and post-vaccination balanced growth paths. Orange dot at the 80% efficacy denotes the reported efficacy of the current vaccine.

far has found only small growth effects of health improvement. Using a quantitative dynamic macroeconomic model informed by reduced-form empirical evidence, this paper analyzes the long-run macroeconomic impacts of eliminating malaria, one of the deadliest diseases with a long-lasting cognitive damage for children in sub-Saharan Africa. In contrast to the previous literature, this paper argues that eliminating malaria in a typical sub-Saharan African country would generate large increases in output per capita, almost ten times as large as previously estimated. Through the lens of the model, eliminating malaria cause parents and children to undertake higher investments in human capital, which drives substantial increase in living standards in the long run.

Developing countries, especially sub-Saharan African countries under the current context, are characterized by low fiscal and state capacity. While the vaccine will undoubtedly save many lives, whether the distribution of vaccines should be prioritized over other imperative development objectives depends on the size of economic benefits it generates. The results in this paper support prioritizing removing malaria as a core development objective. Moreover, since the long-run gains are materialized through the better education of healthier children, eliminating malaria can be complementary to policies aiming to improve the education system. Aside from the policy relevance, this paper also suggests that children's poor health conditions are one of the main reasons income per capita remains low in the developing world.

The model has several simplifying assumptions. Most prominent is perhaps the absence of physical capital. Improvements in health conditions may lead to a higher stock of physical capital since healthier workers can save more and increase the capital's marginal product. In this sense, including physical capital might further increase the long-run gains. However, it is not straightforward to expect that the physical capital would amplify the long-run gains because financial frictions might dampen the accumulation of capital, especially in developing countries where disease eradication is most needed (Banerjee and Duflo, 2005). Considering other channels, including the physical capital, through which health improvement affects long-run growth remains a topic for future research.

1.7 Acknowledgements

This chapter is currently being prepared for submission for publication of the material. Kim, M. The dissertation author was the primary investigator and author of this material. This project was funded through the generous support of the Structural Transformation and Economic Growth (STEG) programme.

Chapter 2

Debt, Human Capital, and the Allocation of Talent

2.1 Introduction

Total outstanding student loan debt reached \$1.57 *trillion* in 2022, surpassing auto loans and credit card debt to become the second largest household financial liability after home mortgages.[1] On the one hand, the increase in student debt represents a policy success in which subsidized federal loans alleviated credit frictions to help broaden access to higher education. On the other hand, the rise in student debt may prevent many indebted students from fully realizing the benefits of a college education by exacerbating subsequent credit constraints after graduation. Surveys of non-delinquent borrowers from federal loan programs suggest such concerns are well founded, with 34% of borrowers reporting their student loans resulted in more hardship than anticipated; 54% reported they would borrow less if they could repeat college; and nearly one-fifth reported "significantly changing career plans because of student loan burdens".[2]

This paper provides a theoretical and empirical analysis of the welfare and productivity consequences of rising student debt. It focuses in particular on student debt's effect on aggregate labor markets and the early career outcomes of college graduates. We show that students graduating with more student debt have higher initial earnings, but lower returns to experience in

[1] See (Federal Reserve Bank of New York, 2022) for details on the composition of household debt.
[2] See (Baum and O'Malley, 2003) for details on the complete questionnaire and survey methodology.

the first ten years of their career. Nearly half the effect is mediated by initial occupation choice after graduation, with more indebted students selecting into occupations with more front loaded compensation schemes. We develop a quantitative heterogeneous-agent, incomplete markets model to rationalize the findings and study their implications for federal student loan policies. The results show that budget-neutral reforms which alter the principal or duration on outstanding federal student loans can increase aggregate welfare and labor productivity. The gains arise from reducing credit constraints that inhibit human capital accumulation on the job and give rise to a misallocation of talent by reducing the efficiency of occupational sorting.

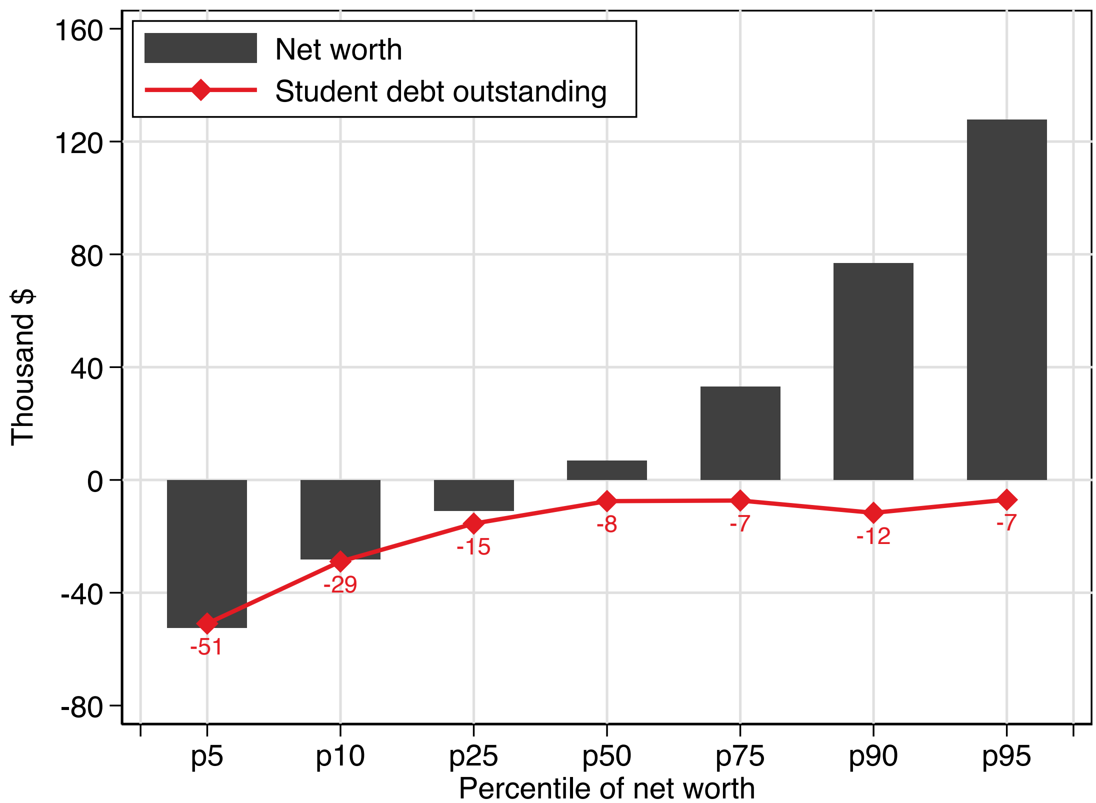


Figure 2.1. Net Worth and Student Debt of Young Workers, Ages 22-25.

Note: Source data from 2018 Survey of Consumer Finance. Student debt defined as total value of aggregate loan balance of education-purpose expenses. Sample is restricted to BA degree holders between age 22 and 25.

While our findings apply more broadly to the interaction of credit market frictions and labor markets, there are several advantages to focusing in particular on the impact of student debt. First, student loan debt is among the largest and fastest growing forms of household credit. Second, it is primarily incurred early in life, before individual labor market experiences diverge, making its effects on lifecycle outcomes easier to measure and isolate. Figure 2.1 shows that student debt accounts for virtually all debt held by young workers and a substantial portion of the the variation in their net worth. Third, student debt is largely non-dischargeable, allowing us to largely abstract from strategic delinquency and bankruptcy considerations. Empirically, households are much less likely to become seriously delinquent (90-days or more) on their student debt than on other forms of credit, such as credit card debt or auto loans.[3] Finally, the federal government plays an out-sized role in student debt markets, accounting for 92.7% of the total outstanding student loan debt.[4] Changes in federal loan policies can therefore generate aggregate variation in household debt, providing a natural laboratory to study the causal effects on labor market outcomes. The dominant role of the federal government also means that there is broad scope for policy improvements to deliver substantial aggregate welfare and productivity gains.

Employing panel microdata from the National Longitudinal Survey of Youth (NLSY97), we provide empirical evidence on the effect of student debt on early career labor market outcomes. We instrument student debt levels using variation in the share of grant funding within a college and across cohorts. Constructing the instrument requires accessing restricted-use NLSY97 identifiers for the participants' educational institutions to merge in annual data on college-level loans and grants from the National Center of Education Statics (NCES). The empirical result show that an additional \$1000 of student debt increases initial earnings by 3.14%, but *reduces* the returns to experience by 1.37%. The effect on initial earnings corresponds to an additional \$508 in annual earnings for every \$1000 in student debt, in line with similar estimates in the

[3]In 2022Q3, only 1.04% of outstanding student debt became seriously delinquent, the lowest rate among all major categories of household credit except mortgages and home equity. See (Federal Reserve Bank of New York, 2022).

[4]Student loan debt held by the federal government is composed overwhelmingly of Direct Loans, which account for \$1.4 trillion of the total. The remaining balance is made up mostly of Title IV loans issued through the FFEL Program and federal Perkins loans. See (Consumer Financial Protection Bureau, 2022) for further details.

literature.[5] The effects of student debt on the returns to experience are statistically significant and also sizable given that the average annual earnings growth of individuals between the ages of 25 to 30 are estimated to be 7.75% (Guvenen et al., 2021). Finally, nearly half the estimated effect on earnings is mediated by initial occupation choice after graduation, as more indebted students select into occupations with more front loaded compensation schemes.

To understand the data, we develop a dynamic model of lifecycle human capital accumulation and occupation choice in the presence of credit market frictions. When credit constraints bind, household discounting of future income streams is greater than prevailing market interest rates. The resulting intertemporal distortions lead households to dis-invest in human capital accumulation as an alternative form of consumption smoothing, reducing lifetime earnings and aggregate labor productivity. A novel feature of the model is that these adjustments can occur not only through reduced investment on the job, but also through changes in occupation choice. Indebted households disproportionately select into occupations with more front loaded compensation schemes, even when their abilities make them a better match for other types of work. The result is a misallocation of talent that can compound the effects of student loan debt on aggregate labor productivity.

To study the implications for federal policies, we embed these mechanisms into a quantitative heterogeneous agent, incomplete markets model that can be taken to the data. Individuals are born with heterogeneous family assets and occupation-specific abilities. They endogenously incur student debt when deciding whether to attend college, accounting for any selection effects that may be in the data. After graduation, households choose an occupation based on their innate abilities and financial assets, taking prevailing wages as given. Earnings evolve endogenously over their lifecycle as a consequence of costly investments in human capital and idiosyncratic labor market shocks. Households are also subject to progressive taxation, have access to public unemployment insurance and retirement benefits which depend on their earnings,

[5]For instance, see (Rothstein and Rouse, 2011), (Chapman, 2015), and (Luo and Mongey, 2019) for comparable estimates of the impact of student debt on initial earnings.

and face realistic student debt repayment provisions and borrowing constraints. The calibrated model matches the aggregate earnings profile and the joint distribution of family assets, student debt, and college matriculation. It also replicates the employment shares, initial earnings, and returns to experience in the 18 detailed occupation groups households choose between.

Using the quantitative model, we measure the aggregate welfare and productivity consequences of budget-neutral changes to the repayment duration or principal of federal student loans. The first exercise computes the effect of 2 and 5 year extensions to the standard federal repayment program. In each case, the interest rate on outstanding federal loans is recomputed to ensure that the net present value of individual student debt does not change. The model predicts these extensions would meaningfully boost welfare and labor productivity, with the gains increasing in the program's duration. For instance, a 5 year budget-neutral extension to the standard federal repayment program would increase aggregate labor productivity by 0.57% and consumption equivalent welfare by 0.45%, with benefits concentrated among low wealth households. The results also show that induced occupation switchers contribute negligibly to aggregate welfare gains, but account for one-third of the rise in aggregate labor productivity despite representing only 0.41% of the treated population. Decompositions of the aggregate productivity increase reveals that the out-sized contribution of induced-switchers is the result of a systematic pattern in the direction of occupation switching; the policy induces workers to flow predominantly from low-skill service jobs into more human capital intensive occupations, mainly Sales, Engineering, Math and Computer Science.

The second set of exercises examines the effect of reducing the principal of outstanding federal student loan debt via student debt forgiveness. It computes the aggregate and distributional effects of student debt forgiveness capped at the 10k, 50k, and 100k thresholds. Given the empirical distribution of student debt, the 100k-cap program effectively amounts to complete debt forgiveness. To ensure the program is budget-neutral, the debt forgiveness is funded by income taxes which are distributed across households with the same proportionality as the overall U.S. tax system.

The results show that the impact of debt forgiveness policies is non-monotonic in the size of the program, with middle-sized programs performing the best. The 50k-capped program yields a modest 0.02% rise in consumption equivalent welfare and a 0.20% increase in aggregate labor productivity. In contrast, the 100k-capped program yielded only a 0.09% increase in labor productivity and a decline in consumption equivalent welfare, while the small 10k program lead to aggregate declines in both welfare and productivity. The non-monotonicity arises from the countervailing effects of debt forgiveness and the distortionary taxation which funds them. The small 10k program is insufficient to substantially alleviate credit constraints, despite still requiring a substantial rise in tax revenue. The 100k program virtually eliminates student debt, but requires large increases in distortionary taxation that reduce the benefits by discouraging human capital accumulation on their own. The model suggests that middle-sized programs which are large enough to alleviate credit constraints, but no so large as to substantially increase distortionary taxes, deliver the best returns. Overall, the results suggest that extended repayment duration policies are a better tool than student debt forgiveness. While both alleviate credit constraints, the former does not require distortionary taxes to support transfers across households.

Related Literature

Our findings contribute to the literature examining how credit market frictions effect labor market outcomes. Recent contributions have shown that access to consumer credit can effect household job search behavior with aggregate implications for the efficiency of worker sorting and business cycle volatility (Herkenhoff et al., 2016; Herkenhoff, 2019). This paper focuses in particular on student debt and its effect on the early career outcomes of college graduates. Our work complements research on credit frictions in the financing of higher education (Lochner and Monge-Naranjo, 2012; Lochner et al., 2021) by investigating how student debt subsequently effects labor market outcomes after graduation.

Empirically, this paper presents new evidence on how student debt effects household earnings profiles and occupation choice. It contributes to a growing literature documenting the effect of student debt on household lifecycle outcomes, such as homeownership, marriage, fertility, and attending graduate school (Goodman et al., 2018; Chakrabarti et al., 2020). In particular, we provide additional evidence of how debt effects lifecycle earnings through distortions to occupation choice (Rothstein and Rouse, 2011; Luo and Mongey, 2019; Herkenhoff et al., 2021).

Our analysis employs a dynamic stochastic heterogeneous agent model of lifecycle earnings with incomplete markets. Following the literature, we analyze the effect of policy reforms by calibrating the model with microdata on household assets, student debt, education, and labor market outcomes (Ionescu, 2009; Huggett et al., 2011; Abbott et al., 2019; Fu et al., 2021). A novel feature of our framework is the inclusion of 18 distinct occupations which, based in part on the endogenous sorting of workers, exhibit different earnings and returns to experience. In this sense, our paper is most similar to (Luo and Mongey, 2019) who develop a quantitative model of how student debt effects household earnings and occupation choice. We build on their contribution by showing how student debt reduces the returns to experience by incentivizing households to sort into occupations with front loaded compensation schemes. As a result, while (Luo and Mongey, 2019) find that reducing student debt leads workers to sort into occupations with greater amenity value, we find that workers flow into occupations with greater scope for earnings growth. Taken together, our results suggest that greater scope for human capital accumulation on the job may be one particular non-wage amenity driving the sorting patterns they uncover.

The remainder of the paper is organized as follows. Section 2.2 presents a simple model to show how credit constraints lead to intertemporal distortions that can reduce human capital investment and give rise to a misallocation of talent. Section 2.3 presents empirical evidence on the effect of student debt lifecycle earnings profiles and occupation choice. Section 2.4 describes the quantitative model, calibration strategy, and reports the model fit. Section 2.5 presents the computational results. Section 2.6 concludes.

2.2 An Illustrative Model

This section introduces a simple model of human capital accumulation and occupation choice in the presence of credit constraints. When constraints bind, households dis-invest in human capital accumulation as an alternative mode of consumption smoothing. Section 2.2.1 describes the intertemporal distortions to human capital accumulation on the job. Section 2.2.2 shows how credit constraints also effect aggregate human capital by distorting occupation choice, pushing households towards occupations with front loaded compensation schemes over those which are best matched to their abilities.

Households live for two periods, young (y) and old (o), and choose consumption, savings, an occupation, and how much to invest in human capital accumulation. Each household is endowed with one unit of time in each period and begins life with initial assets a_y. The problem of a household in occupation k is given by

$$V_k = \max_{c_y, c_o, a_o, s} u(c_y) + \beta u(c_o)$$

subject to

$$c_y = w_k(1-s) + a_y - a_o$$

$$c_o = w_k h(\theta_{ik}, s) + (1+r)a_o$$

$$a_o \geq -\bar{a} \ , \ \ s \in [0,1]$$

where $h(s,\theta)$ is the human capital in adulthood of a household with talent θ_i who invested s time in human capital accumulation. Human capital during youth is normalized to one. The human capital technology $h(s,\theta)$ is an increasing and concave function of time invested,

$$\frac{\partial h}{\partial s} > 0, \qquad \frac{\partial^2 h}{\partial s^2} < 0$$

To interpret θ_{ik} as an index of individual talent which makes it easier for individual i to accumulation human capital in occupation k, we impose that

$$\frac{\partial h}{\partial \theta} > 0, \qquad \frac{\partial}{\partial \theta}\frac{\partial h}{\partial s} > 0, \qquad \frac{\partial}{\partial \theta}\frac{\partial^2 h}{\partial s^2} < 0$$

Most widely used models of human capital accumulation satisfy these criteria, such as (Ben-Porath, 1967) and (Mincer, 1974). Households realize their occupation specific abilities $\Theta = \{\theta_k\}$ as soon as they enter the labor market and, taking wages w_k as given, choose the occupation which maximizes their discounted lifetime utility. Formally, the optimal occupation choice is

$$k^* = \text{argmax}\ \{\ V_1, V_2, ..., V_K\ \}$$

The credit constraints appear in the parameter $\bar{a}$ which limits the amount households can borrow against future income. The key economic friction is that households cannot collateralize their human capital in order to loosen present day borrowing constraints.

2.2.1 Intertemporal Distortions to Human Capital

In the absence of frictions, households invest in human capital accumulation until the marginal return of further investment equals the return on physical capital. The optimal investment in human capital s^* is given by,

$$\frac{\partial h(\theta_{ik}, s^*)}{\partial s} = 1 + r \tag{2.1}$$

When borrowing constraints bind, households discount future income streams relative to the present at a rate that is greater than the market interest rate $1+r$. As a result, households dis-invest in human capital as an alternative form of consumption smoothing. When credit constraints bind, households invest in human capital accumulation until the marginal return equals the shadow interest rate, $1+r^c$, so that

$$\frac{\partial h(\theta_{ik}, s^c)}{\partial s} > 1+r$$

Consequently, credit constraints lead households to invest less in human capital accumulation, $s^c < s^*$, leading to higher initial earnings

$$w_k(1-s^c) > w_k(1-s^*)$$

but lower returns to experience

$$\frac{h(\,\cdot\,, s^c)}{(1-s^c)} < \frac{h(\,\cdot\,, s^*)}{(1-s^*)}$$

Moreover, total lifecycle human capital accumulation is reduced for workers facing credit constraints, leading to a reduction in aggregate labor productivity alongside changes in the structure of lifecycle earnings.

2.2.2 The Misallocation of Talent

In addition to reducing investment on the job, households can also respond to credit constraints by switching occupations. Limited ability to borrow against future income leads some households to switch away from occupations that offer opportunities for human capital accumulation on the job to those with more front-loaded compensation schemes. As a result, occupation choice will depend on initial household assets, giving rise to a misallocation of talent as some workers select into occupations for which their abilities are not optimally matched.

A key condition for credit constraints to give rise to a misallocation of talent is heterogeneity in occupational wages w_k. The variation in wages provides a trade-off in the level and steepness of the lifecycle earnings profile when moving across occupations in a manner similar to the *within* occupation trade-offs offered by training on the job, s, reviewed in section 2.2.1. In particular, with heterogeneity in w_k, each occupation will have its own cutoff $\bar{\theta}_k$ such that all workers with talent $\theta_k > \bar{\theta}_k$ will be constrained. In other words, it is again the high ability individuals that will be most effected by the credit constraints. Intuitively, the cutoff $\bar{\theta}_k$ corresponds to the ability level at which individuals expect sufficiently rapid earnings growth during their career that they would want to borrow beyond the limit $\bar{a}$. Formally, the cutoff can be expressed implicitly by,

$$\frac{\beta}{1+\beta}\left[a_y + w_k(1-s^*)\right] - \frac{1}{(1+\beta)(1+r)} w_k h(s^*, \bar{\theta}_k) = -\bar{a} \tag{2.2}$$

where s^* is optimal human capital investment on the job, as defined in equation (2.1), and the whole left hand side corresponds to optimal asset holdings a^* in the absence of constraints. The expression shows how the cutoffs depend on both initial household assets, a_y, and vary across occupations with w_k. Consistent with economic intuition, the cutoffs in all occupations are increasing in initial household assets a_y, so that it is highly talented individuals from poor households that are most effected by the constraints.

To illustrate the misallocation of talent, it is useful to consider a simple parametric case with log utility and human capital technology $h(s,\theta) = \theta^{1-\alpha}s^{\alpha}$, which satisfies the assumptions made above. Suppose households can choose between two occupations where, without loss of generality, occupation $k = 2$ offers higher wages, so $w_2 > w_1$. In the absence of binding credit constraints, households select the occupation which offers them the highest present discounted value of lifetime earnings. Given a realization of occupation specific abilities, the condition reduces to choosing occupation 1 if

$$\theta_1 > \frac{w_2 - w_1}{\kappa w_1} + \frac{w_2}{w_1} \cdot \theta_2 \tag{2.3}$$

where $\kappa = \left(\frac{1}{1+r}\right)^{\frac{1}{1-\alpha}} \left[\alpha^{\frac{\alpha}{1-\alpha}} - \alpha^{\frac{1}{1-\alpha}}\right] > 0$. The condition in (2.3) is economically intuitive: it is optimal to choose the occupation offering a lower wage provided one has sufficiently high ability to generate greater overall lifetime earnings.

In contrast, we can consider an individual with high level of debt (or sufficiently low assets) such that they are constrained in both occupations. In this case, they value not only discounted lifetime earnings, but also the timing of that income. Given a realization of occupation specific abilities, an individual chooses occupation 1 if

$$\theta_1 > \left(\frac{w_2 + a_y}{w_1 + a_y}\right)^{\frac{1+\alpha\beta}{(1-\alpha)\beta}} \cdot \frac{w_2}{w_1} \cdot \theta_2 \tag{2.4}$$

Comparing the occupational sorting rules in equation (2.3) to (2.4), one can see that there are individuals $\blacksquare = \{\theta_1, \theta_2\}$ who would choose occupation 1 in the absence of credit constraints, but choose occupation 2 if they were financially constrained by debt or insufficient assets.[6] The extent to which these distortions to occupation choice effect aggregate outcomes depends on the joint distribution of assets, a_y, and talent, Θ, in the population. The more highly talented individuals are encumbered with large debts, the greater will be the effect on aggregate labor productivity.

Figure 2.2 illustrates the misallocation of talent by plotting how occupational sorting depends on the presence of credit constraints. For a given initial assets a_y, it shows how individuals sort into each occupation as a function of their talents. The cutoffs $\bar{\theta}_1$ and $\bar{\theta}_2$ are defined as in (2.2), indicating the regions where credit constraints bind in each occupation.[7] The shaded region represents the population of workers who switch from occupation 1 to occupation 2 in the presence of credit constraints. The lower border of the region corresponds to the optimal occupation sorting rule in the absence of credit constraints expressed in (2.3). The low ability population, those with $\theta_1 < \bar{\theta}_1$ and $\theta_2 < \bar{\theta}_2$, continue to choose occupations according to this

[6]To see this more directly, note that the occupation sorting condition in (2.4) is a line passing through the origin with slope greater than w_2/w_1. This implies that that the slope of the sorting rule in Θ space is steeper when individuals are constrained than when they are not constrained.

[7]Given the parametric example being considered, equation (2.2) can be solved explicitly so that

$$\bar{\theta}_k = \frac{\beta(1+r)^{\frac{1}{1-\alpha}}}{\alpha^{\frac{1}{1-\alpha}} + \alpha^{\frac{\alpha}{1-\alpha}}}\left(1 + \frac{a_y}{w_k}\right)$$

rule, since they are sufficiently untalented that credit constraints do not effect them. The high ability population, those with $\theta_1 > \bar{\theta}_1$ and $\theta_2 > \bar{\theta}_2$, are most affected by the constraints and sort according to equation (2.4). Since $w_2 > w_1$, there is an additional population whose occupation choices are distorted because they become constrained in occupation 1 $\theta_1 > \bar{\theta}_1$ but not occupation 2, $\theta_2 < \bar{\theta}_2$, and hence switch to the latter.

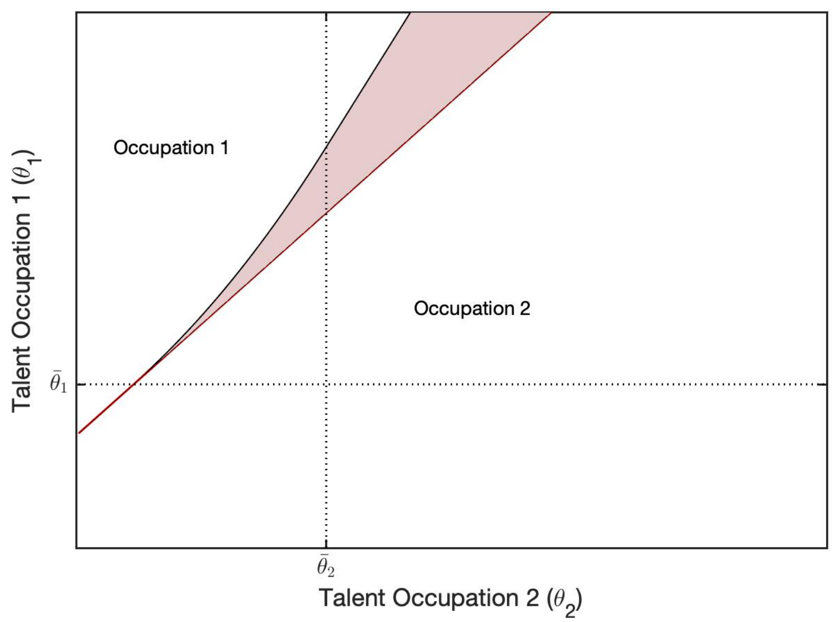


Figure 2.2. Misallocation of Talent

Note: This figure illustrates how the misallocation of talent depends on individual's abilities, for a given initial asset a_y. Occupation 2 is assumed to offer higher wages, $w_2 > w_1$.

Moving to populations with higher assets, the cutoffs $\bar{\theta}$ increase and the upper border of the misallocation region shifts inward toward the optimal sorting rule, shrinking the region where credit constraints distort occupation choices. Conversely, moving to households with lower assets leads the $\bar{\theta}$ cutoffs to shrink and the misallocation region to expand, resulting in greater distortions to occupation choice. As a result, the aggregate effect of the credit frictions will depend on the joint distribution of assets a and talents Θ in the population.

This joint distribution will also determine the marginal effect of policies like student debt forgiveness, since it will determine on the mass of workers near the threshold of misallocation

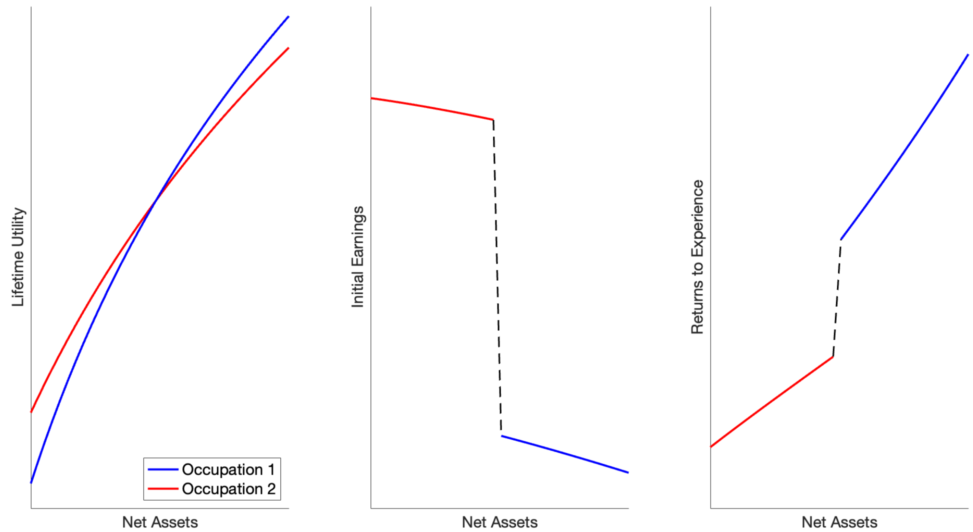


Figure 2.3. Household Assets and Occupation Switching

Note: This figure illustrates the effects of reducing an individual household's debt level on lifetime utility, initial earnings, and returns to experience. The simulation corresponds to an individual who is has higher ability in occupation 1, $\theta_1 > \theta_2$, while occupation 2 offers the higher wages, $w_2 > w_1$.

region. As discussed in section 2.2.1, policies which alleviate credit constraints will boost human capital accumulation and alter the lifecycle earnings profile by reducing the shadow interest rate. For the population near the misallocation region, alleviating credit constraints can also induce occupation switching that leads to discrete adjustments in both productivity and lifecycle earnings. Figure 2.3 illustrates the effect, showing how initial earnings, returns to experience, and occupation choice of a constrained household responds to reductions in their debt. Initial reductions in household debt lead to increased human capital accumulation, reduce initial earnings and boosting returns to experience on the job. Eventually, household debt falls sufficiently far that the household passes out of the misallocaiton region and switches from occupation 2 to occupation 1, eschewing the higher initial earnings for greater lifetime earnings. The occupation switch leads to a large, discrete jump in earnings and returns to experience. After the switch, further debt reductions continues to effect human capital accumulation within the occupation, albeit with slightly larger effects in the more human capital intensive occupation.

The illustration shows how endogenous occupation choice can mediate part of the impact of credit constraints on aggregate labor productivity and household earnings. Assessing the true aggregate contribution of these channels therefore requires measuring the joint distribution of assets and talent in the relevant population which will together determine the within and between occupation effects of credit frictions, as well as how they respond to policy interventions. We turn to addressing these challenges of measurement in the following sections, both empirically by exploiting plausibly exogenous variations in student debt as well as structurally through the calibration of a larger scale quantitative model which embeds these core mechanisms.

2.3 The Empirical Evidence

In this section, we provide reduced-form empirical evidence that individuals' earnings trajectories change in the presence of student debt, and that at least part of this effect is explained by occupational choice. We use panel data from the NLSY 1997 and an instrumental variables design to first estimate the impact of student debt on both initial earnings and returns to experience.

Our findings suggest that individuals with high student debt are forced to trade off between current and future income early in life – initial earnings are higher for those with student debt, while returns to experience are lower. Moreover, this trade-off appears to depend significantly on their *first* occupation and industry choice upon graduation. When we include occupation and industry fixed effects in the IV regression, they explain almost half of our estimated effects.

To further understand why occupational choice matters so much for our estimates, we conduct a supplementary analysis to test whether earnings trajectories vary considerably across occupations. We use the Current Population Survey to construct age-earnings profiles for each occupation in order to investigate how differential sorting into occupations potentially drives our earnings results. In the raw data, we indeed find a strong negative correlation between initial earnings and returns to experience across occupations. These empirical findings support the set-up of our theoretical model of occupational choice, student debt and human capital accumulation, and motivate our quantitative analysis.

2.3.1 Data

Our empirical analysis draws from several data sources. The primary dataset is the NLSY 1997, an individual-level panel dataset that contains information on higher education, student debt, and labor outcomes. It follows individuals from 1997 through 2015. Summary statistics are provided in Table B.1 in the appendix. We restrict our analysis to individuals whose highest level of education is a bachelors degree.[8]

Using the NLSY, we instrument for student debt using variation in the share of grant funding within college and across cohorts, and measure how incremental debt impacts labor market decisions and lifetime earnings trajectories. To construct our instrument, we have accessed restricted-use data that identifies NLSY participants' educational institution. Using the college identifier, we then merge in information from the National Center for Education Statistics (NCES) on the amount of loans and amount of grants used at that given college in a given year.

Supplementary analysis of occupation-specific age-earnings profiles uses the Current Population Survey. While the CPS does not contain information on student debt, it does contain comprehensive earnings information across the spectrum of occupations, industries, and ages.

[8]We have re-run our analysis while also including individuals with an Associates degree. While the first and second stage results are largely consistent with those we find on the BA-only population, we refrain from making this our primary sample since BA and AA degree recipients make very different human capital investment decisions in college and have markedly different observed occupational choices post-degree.

2.3.2 Instrumental Variable Design

To estimate the effect of student debt on initial earnings and returns to experience, we consider the following equation:

$$\begin{aligned} y_{it} = \; & \underbrace{\alpha_0 + X_{it}\beta}_{\text{initial (log) earnings if no student debt}} + \underbrace{\alpha_1 \mathrm{Exp}_{it}}_{\text{returns to experience if no student debt}} + \\ & \underbrace{\alpha_2 \mathrm{SD}_{it}}_{\text{effect of student debt on initial (log) earnings}} + \underbrace{\alpha_3 \mathrm{SD}_{it} \times \mathrm{Exp}_{it}}_{\text{effect of student debt on returns to experience}} + \varepsilon_{it} \end{aligned} \tag{2.5}$$

where y_{it} is an outcome measure of individual i in year t for annual log earnings. The variable *SD* denotes the level of student debt. The variable *Exp* denotes the years of experience. The variable X_{it} includes gender, race and cohort fixed effects.

We seek an unbiased and consistent estimate of α_2 and α_3. The effect of student debt on initial wages is measured by α_2. The effect of student debt on the returns to experience is measured by α_3. There are potential challenges to estimating equation 2.5 using OLS. For instance, we may be concerned that there is a correlation between the level of debt an individual takes on and the unobservable quality or ability of an individual. The bias can go either way. Individuals with high ability may expect to have higher future wage growth. They may decide to borrow more today to smooth consumption over time, leading to an upward bias in α_2. On the other hand, debt may be positively selected. For instance, low ability individuals may come from low income households, who are unable to provide parental support for their child's education. This shows up as higher borrowing for the low ability individual, leading to a downward bias in α_2 and α_3.[9]

To address these identification challenges, we estimate the causal impact of student debt on earnings using an school-cohort-level instrumental variable. Our instrument follows that used

[9] These identification challenges to identifying a causal impact of student debt on earnings are also highlighted by (Field, 2009; Rothstein and Rouse, 2011; Luo and Mongey, 2019). These papers uses variation in forgiveness of debt (Field, 2009) and variation in grants (Luo and Mongey, 2019) within a school across cohorts to instrument for student debt. Their identification comes from comparing outcomes of cohorts within the same school, when cohorts within the school differ in terms of grants received.

in (Luo and Mongey, 2019) — it is defined as the share of grant funding, out of all grant and federal student loan funding, issued by a college in a given year. Specifically, our instrument is defined as:

$$Z_{c(i),j} = \frac{\text{total grants}_{c(i),j}}{\text{total grants}_{c(i),j} + \text{total loans}_{c(i),j}} \tag{2.6}$$

The instrument utilizes the fact that students must fund their college tuition costs through a combination of parental funding, grants, work study aid, and student loans. While parental funding is specific and fixed at the student level, grant funding can vary significantly at the college-year level. As shown in (Luo and Mongey, 2019), variation in grant funding is substantial both across and within institutions and years.

Intuitively, the instrument captures the fact that when colleges have less to award to students in the form of grants, students must make up the remaining "gap" in funding using student loans. To meet the exogeneity assumption of a valid instrument, we argue that yearly variation in the *total* amount of grant funding available at a college is unrelated to the ability (or other unobserved characteristics) of any given student at that college. However, to meet the relevance assumption, this variation in grant funding must also create a meaningful change in amount of student debt that students take out. Table B.2 in the appendix shows a strong first stage effect of shifts in the college-year grant share on individuals' student debt. Importantly, the table also shows that changes in grant funding are compensated for *entirely* and *exclusively* by changes in student debt, not other sources of funding. Total funding for college remains constant in response to one standard deviation increase in the college grant share. And while the level of student debt decreases almost one-for-one with the increase in grant funding, family and work study aid remain constant. This precise, isolated substitution is important, because it allows us to study the impact of an increase in student debt on future earnings, absent of other confounding factors like more parental aid or increased work study while in college.

We also check whether variation in the college grant share changes other important educational outcomes which could confound our results, like the probability of college completion or the ability and wealth distribution of enrolled students. These results are shown in Table B.3. Similar to (Luo and Mongey, 2019), we do not find a significant evidence of our instrument impacting enrollment or student selection on observables.

In our second stage, in which we regress our instrumented student debt variable on earnings, we include fixed effects for college type – for example, private, public, for-profit, etc. While it would be ideal to include fixed effects for each individual college, our small sample size does not allow this – we have very few instances in which more than one student attended the same institution.

2.3.3 Estimated Impact of Student Debt on Initial Earnings and Returns to Experience

Using our instrument, we next investigate whether those who have subsequently higher debt choose jobs with significantly different earnings profiles. Table 2.1 shows the instrumented regression coefficients for α_2 and α_3 for log earnings. The coefficient on α_2 is positive, while the coefficient on α_3 is negative. These coefficients imply that an individual with more student debt has higher initial earnings upon graduation, but subsequently lower returns to experience.

To interpret the magnitudes, the coefficients in column (I) imply that a additional \$1K of student debt increases initial earnings by 3.14%. For our data sample, this equates to an additional \$508 annual earnings upon graduation, for every \$1K of additional student debt. While we use a different sample from the existing literature, we arrive at estimates on the effects for initial earnings that are consistent with existing estimates.[10]

Our new empirical evidence is shown in columns (II). Specifically, we find that total earnings grow by 1.37 ppts *slower* per year of experience respectively, for every \$1K of additional student debt. These effects are statistically significant. The magnitudes are also sizable, given the earnings of $25-30$ year olds are estimated to grow at a rate of 7.75% on average each year

[10] For instance, (Luo and Mongey, 2019), (Rothstein and Rouse, 2011), (Chapman, 2015).

(Guvenen et al., 2021).

Table 2.1. IV coefficients of student debt on initial earnings and returns to experience, with and without first occupation fixed effects

Notes: The Table reports the instrumented estimates from regression 2.5 using the NLSY data. Our IV utilizes changes in the college-year grant share, which in turn impacts the amount of student debt taken out by individual students. Our dependent variable is log yearly earnings. See text for more details.

Effect of student debt ($000s) on:	Grants-based IV	
	(I)	(II)
(i) Log initial earnings	3.14%	1.56%
(pvalue)	0.08	0.13
(ii) Mean returns to experience	-1.37%	-0.78%
(pvalue)	0.08	0.01
Controls:		
Ability, age, college-type, race, gender	Yes	Yes
Occupation FE & Occupation x Exp		Yes
Source of variation:	Across-cohort, within college-type (among all students)	

Occupation FE & Occupation x Exp	No	Yes
Average initial earnings		50%
Average returns to experience explained		43%

We explore what factors explain the wage gap between those with and without student debt by including different controls. The evidence suggests that a large part of the wage gap between individuals that graduated with and without student debt is due to the selection into different occupations upon graduation. The first occupation choice accounts for almost half of the gap in wage profiles between the two groups.

To see this, Column (II) includes fixed effects for the first occupation that an individual chooses upon graduation. We also include the interaction of the fixed effects with the years of experience. The inclusion of these controls reduces the initial earnings gap by 50% (α_2 declines from 3.14 to 1.56). The marginal effect of student debt on returns to experience declines by 43% (α_3 declines from -1.37 to -0.78). These changes imply that initial occupational choices can explain much of the difference in the subsequent earnings profiles between those with and without student debt. These results imply student debt impacts the selection into different occupations upon graduation.

2.3.4 The Role of Occupational Choice and Other Mechanisms

Our instrumental variables analysis finds that additional student debt generates age-earnings profiles with initially higher earnings and lower returns to experience. About half of this effect is explained by occupational fixed effects, meaning that individuals with student debt sort into professions that have a predictably flatter, front-loaded income trajectory. We provide some suggestive evidence in the data on potential sources for the heterogeneity in occupational earnings trajectories, and then test for sorting of individuals with student debt into occupations that are characterized by a flatter profile.

For this analysis, we use the Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS) from 2010-2018 to construct occupation and industry specific age earnings profiles. These are cross-sectional profiles, meaning that the trajectory is constructed *across* individuals of different ages, rather than *within* person over time. While this technique is sensitive to changes in cohort composition over time, it allows us to look at longer trajectories and at more specific occupation groups. We pool multiple cross-sections from the CPS in order to control for some cohort effects.

We use a quadratic specification to estimate occupation-specific earning trajectories. In each 2-digit occupation category, we fit the regression specification $earnings_i = \alpha + \beta_1 age_i + \beta_2 age_i^2 + X_i + \varepsilon_i$ to the cross-section. We restrict the sample to individuals who are between 21-60 years old and have exactly a bachelors degree; we control for household size, race, gender, state, and CPS cohort.

While we do not claim that these profiles are exogenous (e.g. unaffected by the sorting of individuals with different debt or ability levels into certain categories), they do help explain why occupational fixed effects have such a large impact on our regression results. We find that there is considerable heterogeneity across occupations in the level of earnings received upon graduation and returns to experience. If we summarize the age-earnings trajectories by their intercept and slope, the trade-off that individuals face between front-loaded wages and lower growth becomes more apparent. As shown in Figure 2.4 there is a statistically significant negative relationship between occupations' initial log earnings after graduation (y-axis), and their average yearly growth rate in the first 15 years (x-axis).

An individual who is constrained at age 23 by student debt may choose a career in Health Practice rather than in Law, since it provides higher initial wages that can be used to service the debt (see Figure 2.4). The debt constraint would thus induce them to choose a career with potentially lower expected net present value. For constrained individuals, the earnings trajectory becomes an imperfect, costly means of transferring consumption from the future to the present.

It is worth noting that these are equilibrium earnings trajectories, which encompass both endogenous labor supply decisions and occupation-specific technological differences. However, these differences in occupational earnings trajectories may stem from inherent technological differences. Performance, and consequently earnings, in some occupations may benefit more from on the job training. This will lead to a larger impact of human capital investment on returns to experience, and thus steeper earnings profiles, in particular occupations.

We next test whether individuals with higher student debt sort into occupations with a flatter, front-loaded earnings profile. We characterize occupation-specific earnings profiles using the annualized earnings growth rate between ages 23 and 48 from the CPS. We then use our instrumental variables regression from the NLSY to test whether individuals with more debt choose occupations with lower growth rates. There is a statistically significant negative relationship between instrumented student debt and occupation specific growth rate of earnings. This confirms our earlier IV results, that found instrumented student debt had a positive effect on earnings and a negative effect on returns to experience. A similar negative relationship exists for the OLS regression.

Using the OLS coefficients, we then predict the average debt level for individuals by occupation. Figure 2.4 plots the predicted student debt levels by occupation using differently sized bubbles. Large bubbles have relatively high predicted student debt levels compared to small bubbles. In line with our model predictions, high debt occupations also have higher initial earnings and flatter profiles – i.e. they are concentrated in the upper left hand portion of the plot. This graph not only documents the slope-intercept earnings trade-off across occupations, but also the sorting of constrained individuals into front-loaded options.

Figure 2.4. Initial earnings, earnings growth, and student debt by occupation

Notes: This figure plots the predicted student debt levels by occupation using differently sized bubbles. Large bubbles have relatively high predicted student debt levels compared to small bubbles. In line with our model predictions, high debt occupations also have higher initial earnings and flatter profiles – i.e. they are concentrated in the upper left hand portion of the plot.

2.3.5 Summary of Empirical Evidence

In summary, our empirical estimates find that exogenous increases in student debt leads to higher initial earnings and lower returns to experience. Our evidence suggests that part of the effect comes from *within* occupation changes in earnings. The other part of the effect comes from changes in earnings resulting from the first-occupation choice upon graduation. These empirical findings support the intuition behind our illustrative model, which explicitly models the human capital investment and occupational choices of credit constrained individuals.

2.4 The Quantitative Model

In this section, we construct and calibrate a quantitative heterogeneous agent, incomplete markets model with occupation choice and on-the-job human capital investment.

Each period corresponds to one year. Life begins at age 18 when individuals are endowed with initial assets (a), realize their occupation specific talents (θ_k), and decide whether or not to go to college subject to a secondary school taste shock. Individuals who decide to go to college may endogenously incur student debt d if they have insufficient assets to cover tuition. At age 22, graduates enter the labor market and choose an occupation $k \in \{1,2,\dots,K\}$ to maximize their expected lifetime income.

$$k^* = \text{argmax}\ \{\ V_1, V_2, ..., V_K\ \}$$

Working life continues until retirement at age 63; all households die at age 80. Households can be identified by their assets (a), human capital (h), student debt (d), occupation (k), employment status (z), and age (t). The problem of an employed working age household can be expressed recursively,

$$V_k(a,h,d,e,t) = \max_{c,s,a'} u(c)\ +\ \beta\mathbb{E}\left[V_k(a',h',d',z',t+1)\right]$$

subject to

$$
\begin{aligned}
c + a' &= \mathbb{T}\left(w_k(1-s)h_k\right) + (1+r)a - \phi(a,h,d,z,t) \\
h_k' &= \theta_{ik}(sh_k)^{\alpha} + (1-\delta)h_k \\
d' &= (1+r_d)d - \psi(a,h,d,z,t) \\
a' &\geq -\bar{a} \;\;,\;\; s \in [0,1]
\end{aligned}
$$

where households have CRRA preferences $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$. The functions $\mathbb{T}\left(w_k(1-s)h_k\right)$ and $\phi(a,h,d,z,t)$ represent the tax system and student debt repayment rule, respectively, which we explain further below.

Lifecycle Employment Risk.

Households face idiosyncratic risk of unemployment which varies over the lifecycle. Introducing unemployment risk means all households have some probability of being constrained or unconstrained–smoothing out the stark contrast between these types in our illustrative example. We calibrate the probability of job loss to match the job separation rates by age in (Michelacci and Ruffo, 2015). Figure 2.5 plots how the risk of unemployment varies over the lifecycle. Capturing the age structure of unemployment risk is important quantitatively because the data show these risks are concentrated early in life, precisely the same period when credit constraints from student burdens would bind most acutely.

When households are unemployed, they receive unemployment benefits and have their student debt payments deferred with accrued interest at rate r_d. Households also experience skill depreciation (”scarring”) while unemployed by losing access to on the job training. Formally, an age t unemployed ($z = u$) household solves

$$
V_k(a,h,d,u,t) = \max_{c,a'} u(c) + \beta \mathbb{E}\left[V_k(a',h',d',z',t+1)\right]
$$

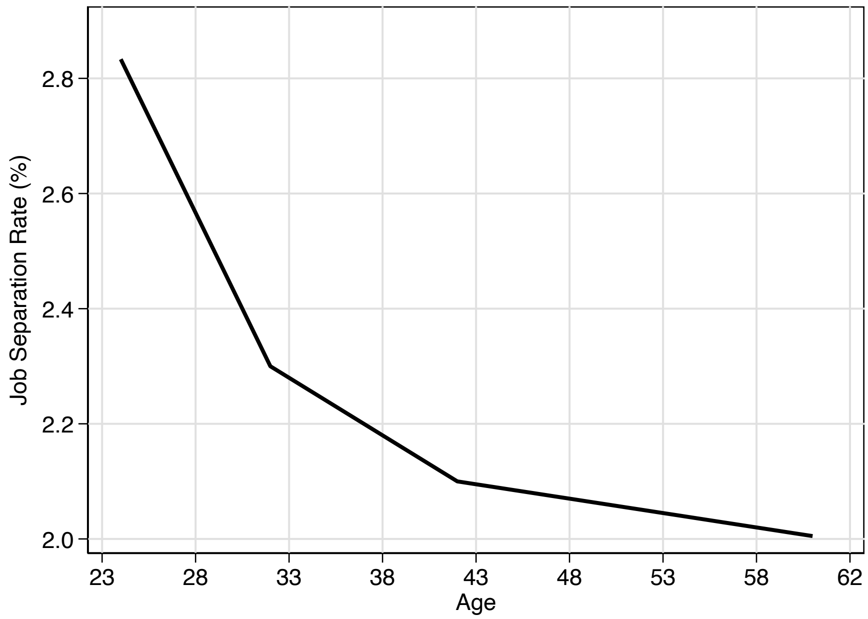


Figure 2.5. Lifecycle Risk of Unemployment [$P(z_{t+1} = u | z_t = e)$]

Notes: This figure plots the risk of unemployment over the life cycle. First (at age 23) and last period (at age 62) job separation rate are assumed to be zero.

subject to

$$c + a' = b(y) + (1 + r)a$$

$$h'_k = (1 - \delta)h_k$$

$$d' = (1 + r_d)d$$

$$a' \geq -\bar{a}$$

The function $b(y)$ captures government unemployment benefits, which are a function of household income at the time of job loss y, e.g. $y = w_k(1 - s_{t-1})h_{t-1}$ for a newly unemployed household. The functional form of $b(\cdot)$ is given by

$$b(y) = \min\{ \underbrace{\$9,600}_{\text{annual benefit cap}} , \underbrace{0.45 \times y}_{\text{income replacement rate}} \}$$

Student Debt and College Matriculation

Households endogenously incur student debt when making the decision whether or not to go to college. Individuals make their matriculation decision at age 18–the beginning of life in the model. We abstract from earnings heterogeneity among high school graduates and model high school as a common outside option, as in (Hsieh et al., 2019). High school graduates receive lifetime utility commiserate with the high school wage and an individual college taste shock $\zeta \sim \text{Frechet}(\varepsilon)$, so that $V_{hs} = \zeta \times \sum u(c_t^*)$. The purpose of the taste shock is to smooth matriculation decisions by family background to better reflect the data.

If individuals decide to attend college (e.g. $V_{k^*} > V_{hs}$), then they will incur student debt if their initial assets are insufficient to cover the cost of college tuition net of any grants or family assistance they receive. Formally, letting τ denote tuition net of any grants or family assistance, household debt is given by

$$d = \begin{cases} 0 & \text{if } V_{k^*} < V_{hs} \\ \min\{\,0\,,\, a_0 - x \cdot \tau\,\} & \text{if } V_{k^*} > V_{hs} \end{cases}$$

where we note that the value of college $V_{k^*}(a,h,d,z,t)$, and hence the optimal college matriculation decision, implicitly depends on the amount of college debt households would take on. To capture the variety of individual circumstances determining access to college grants and family assistance, we assume that the out-of-pocket net tuition individuals must pay to attend college depends stochastically on family background. Parameter x captures extensive probability of having student debt; with probability $1-x$, an individual is able to attend college without incurring any out-of-pocket expenses. With probability x, individuals receive an out-of-pocket net-tuition cost τ of attending college. In calibrating the model, we allow probability x to depend on family assets and, conditional on facing out-of-pocket expenses, assume net tuition costs are stochastic and jointly log-normal in the population such that,

$$\begin{pmatrix} a_0 \\ \tau \\ x \end{pmatrix} \sim \ln N \left[\begin{pmatrix} \mu_a \\ \mu_\tau \\ \mu_x \end{pmatrix} , \begin{pmatrix} \sigma_a^2 & \rho_{a\tau} & \rho_{ax} \\ \rho_{a\tau} & \sigma_\tau^2 & 0 \\ \rho_{ax} & 0 & \sigma_x^2 \end{pmatrix} \right]$$

We choose the parameters x, τ which determine the net-tuition offers households receive to match the *post-matriculation* realizations of student debt by family background. The flexible parameterization allows us to capture the endogeneity of student debt and patterns of selection into college that we observe in the data. Finally, we note that even after receiving grants and financial aid, some households may still have insufficient assets to pay for college and so are unable to matriculate even if very talented. These households highlight the key credit market friction that motivates government education loans and grants in the first place: individuals cannot borrow against their future human capital.

Student Debt Repayment

Households begin repaying student loans after they graduate college. The payments a household makes to service its student loan debt depend on its outstanding balance d, time to maturity $\bar{T} - t$, household assets, potential earnings, and employment status z. The repayment function $\psi(a,h,d,z,t)$ summarizes how repayments depend on individual circumstances.

In normal circumstances, an employed household with sufficient financial resources will make payments $\rho(d,t)$ to amortize its student loan over a repayment period $t < \bar{T}$, as in (Luo and Mongey, 2019), given by

$$\rho(d,t) = \left[\frac{r_d}{1-(1+r_d)^{-(\bar{T}-t+1)}} \right] d$$

Due to the stochastic unemployment risk, it is possible that households find themselves unable to make their student loan payments. This may happen if a household becomes unemployed, has insufficient assets, or has experienced long unemployment spells which resulted in skill depreciation (e.g. low income due to ”scarring”). Consistent with “undue hardship” provisions of student debt repayment programs, we capture the effect of these circumstances on debt repayment through the function $\xi(z)$ such that the repayment function is given by,

$$\phi(a,h,d,z,t) = \min\left\{ \rho(d,t)\ ,\ \xi(z)(a+w(1-s)h) \right\}$$

so that households are never forced to make student debt payments in excess of $\xi(z)$ of their households net worth (e.g. income and assets). When employed, households become delinquent on their student loan debt if their annual payments are in excess of $\xi(e)$ fraction of their total net worth. Similarly, as discussed in the previous section, households can defer their debt repayments when unemployed, so that $\xi(u) = 0$.

Finally, while economic hardship can allow households to defer debt repayments and accrue interest on their outstanding balance, we do not allow debt deferral to continue beyond the maturity ceiling $\bar{T}$. In the final period of debt maturity households must pay off their total outstanding balance or default. Households do not make any student debt payments beyond the maturity period, e.g. $\phi(a,h,d,z,t) = 0$ for all $t > \bar{T}$.

The Tax System

$\mathbb{T}(y)$ is a function which represents the prevailing tax system and transforms gross household income y into after-tax income. In particular, $\mathbb{T}(y)$ takes the form of a step function

$$\mathbb{T}(y) = (1-\tau(y)) \cdot y$$

where $\tau(y)$ represent the effective marginal tax rates for the tax bracket of individuals with income y. The brackets and marginal rates $\tau(y)$ are chosen to match the effective tax rates estimated by the Congressional Budget Office (CBO) displayed in figure 2.6. Accurately modelling the effective marginal tax rates is quantitatively important since these will influence household

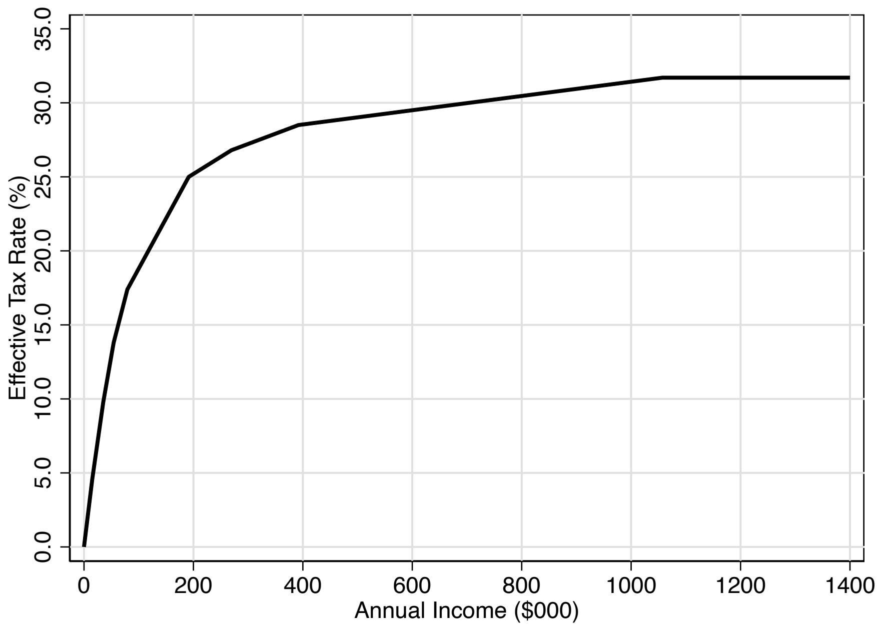


Figure 2.6. CBO Effective Tax Rates

Notes: This figure plots effective tax rates by income categories. Effective tax rate includes individual income taxes, social security taxes, corporate income taxes, and exercise taxes. Source: Congressional Budget Office, *Effective Federal Tax Rates, 1979–2004* (December 2006), Table 1.

incentives to raise their income by attending college and investing in human capital (Saez et al., 2012; Jones, 2022). Furthermore, several of our policy counterfactuals are redistributive (e.g. debt forgiveness) and we use tax system $\mathbb{T}(y)$ to quantitatively match the progressivity of tax incidence required to finance these programs.

Retirement and Retirement Benefits.

Households retire deterministically at age $t = 63$ and continue to make consumption and savings decisions until mortality at age $t = 80$. Retired households fund consumption out of savings and retirement benefits $\pi(\hat{y})$, which they begin receiving after retiring. Formally, retired households solve

$$V_R(a, \pi) = \max_{c, a'} u(c) + \beta V_R(a', \pi)$$

subject to

$$c + a' = \pi(\hat{y}) + (1+r)a$$
$$a' \geq -\bar{a}$$

Household retirement benefits $\pi(\hat{y})$ depend on individual earnings at the end of working life, just before retirement, e.g. $\hat{y} = w_k h_k$. The dependence of retirement benefits on household earnings creates another incentive for households to accumulate human capital over the course of working life, beyond those proxied by the assets held at retirement. On the one hand, these benefits increase the incentive to accumulate human capital, exacerbating the effect of credit constraints which hinder investment early in life. On the other hand, they provide insurance in old age, reducing the need to save and accumulate assets.

To quantitatively account for the importance of these incentives, we parameterize the retirement benefits function following the approach in (Daruich, 2020) to capture the benefits formula used by the U.S. Social Security System. Specifically,

$$\pi(\widehat{y}) = \begin{cases} 0.9\hat{y} & \text{if } \hat{y} \leq 0.3\bar{y} \\ 0.9(0.3\bar{y}) + 0.32(\widehat{y} - 0.3\bar{y}) & \text{if } 0.3\bar{y} \leq \hat{y} \leq 2\bar{y} \\ 0.9(0.3\bar{y}) + 0.32(2-0.3)\bar{y} + 0.15(\widehat{y} - 2\bar{y}) & \text{if } 2\bar{y} \leq \hat{y} \leq 4.1\bar{y} \\ 0.9(0.3\bar{y}) + 0.32(2-0.3)\bar{y} + 0.15(4.1-2)\bar{y} & \text{if } 4.1\bar{y} \leq \widehat{y} \end{cases} \tag{2.7}$$

where $\bar{y}$ is economy-wide average income, approximately \$71,700. Figure 2.7 depicts the calibrate retirement function and compares it to a proportional benefit program and the flat benefits program employed by (Huggett et al., 2011).

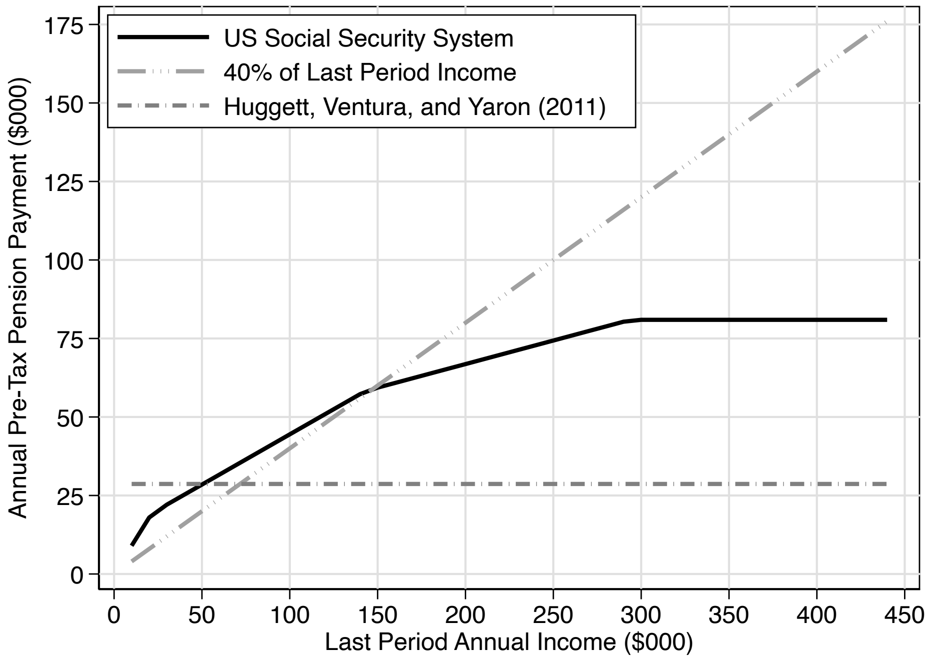


Figure 2.7. Calibrated Retirement Benefits Function

2.4.1 Internal Calibration and Model Fit

The household risk preference parameter, the structural parameters of the human capital technology, and the interest rates on debt are set exogenously. The CRRA preference parameters ρ is set equal to 2, consistent with standard values employed in the literature. While the population distribution of occupation specific talents θ_{ik} are calibrated internally, the structural parameters governing the human capital technology are set exogenously to be consistent with the literature. Following (Huggett et al., 2011), α is set to 0.7 and the depreciation rate δ is set to 0.05 to match the in-sample end of working life decline in earnings. The exogenous interest rate on household savings r equals 0.04 and the annual interest rate on student debt r_d is set to 0.042, consistent with the average federal student loan rate observed in the data.

The remaining parameters are simultaneously internally calibrated so that the model endogenously generates a joint distribution of assets and student debt (as in Figure 2.1), college matriculation patterns, and occupational heterogeneity in earnings, returns to experience, and

Table 2.2. Internal Calibration Targets: Assets, Debt, and College Matriculation

	Data	Model
Mean of log initial assets (population)	10.50	10.22
Variance of log initial assets (population)	2.47	2.16
Mean of log initial assets (college grads only)	11.10	10.61
Variance of log initial assets (college grads only)	2.19	2.18
Mean student debts ($)	17,784	17,468
Standard deviation of student debts ($)	24,889	24,173
Corr(Asset, Student Debt)	-0.16	-0.16
Fraction of sample with student debts (%)	67.0	62.1
Corr(Asset, Having Student Debt)	-0.25	-0.22
College completion rate by family background (%)		
- first asset quintile (Q1)	18.2	24.4
- second asset quintile (Q2)	22.8	26.2
- third asset quintile (Q3)	26.7	29.5
- fourth asset quintile (Q4)	37.7	35.4
- fifth asset quintile (Q5)	50.7	51.2
Variance of log earnings in the first 10 years	0.39	0.39
Ratio of assets, old-to-young workers	5.76	5.79

worker sorting (as in Figure 2.4) that are consistent with the data.

The parameters $\{\beta, \mu_a, \sigma_a, \mu_\tau, \sigma_\tau, \rho_{a\tau}, \mu_x, \rho_{ax}, \varepsilon, \sigma_\theta\}$ primarily determine the joint distribution of household assets, student debt, college matriculation, and earnings inequality. Table 2.2 summarizes the 16 targets most closely associated with these 9 parameters. These parameters are fit internally since student debt will be incurred endogenously based on selection into college, which itself will depend on the parameters governing expected payoffs in the labor market (see Table 2.3). Parameters μ_a and σ_a pin down the distribution of initial assets in the population. Together with $\mu_\tau, \sigma_\tau, \rho_{a\tau}$, these parameters determine the joint distribution of assets and student debt among the indebted, conditional on selection into college. Similarly, parameters μ_x and ρ_{ax} will determine the extensive margin of college debt among graduates. Combined with the distribution of assets and student debt, parameter ε governing the high school taste shock will determine college matriculation patterns by family background. The common variance of student

Table 2.3. Internal Calibration Targets: Occupational Heterogeneity

Notes: This table summarizes the 54 calibrated parameters $\{w_k, \mu_k, \nu_k\}_{k=1,\ldots,18}$ and their data counterparts. Employment share is calculated among college graduates, so they add up to 100%.

Occupation Group	Mean Log Earnings ($)		Returns to Experience (%)		Employment Share (%)	
	Data	Model	Data	Model	Data	Model
Executive Admin	9.95	9.93	12.06	11.83	7.35	7.14
Management	9.93	9.91	13.28	12.85	8.75	8.88
Math and Computer Science	10.14	10.12	11.34	11.17	4.56	4.26
Architects and Engineers	10.09	10.07	11.20	11.07	6.12	5.93
Counslers	9.69	9.66	10.44	10.65	5.69	5.95
Teachers	9.83	9.82	8.41	8.47	14.17	14.16
Education	9.55	9.56	4.23	3.89	1.61	1.73
Sports	9.37	9.35	18.09	18.49	3.44	2.99
Media	9.60	9.56	11.60	11.68	2.90	3.03
Health Practitioners	10.29	10.27	4.28	4.49	4.03	4.06
Health Support	9.74	9.74	8.23	8.09	2.20	2.20
Food Services	9.76	9.74	2.81	2.87	5.31	5.34
Cleaning	9.44	9.46	10.85	9.84	1.29	1.29
Service Workers	9.38	9.39	7.80	7.43	3.17	3.10
Sales	9.73	9.75	13.44	12.77	11.11	12.29
Office & Admin	9.69	9.69	9.49	9.43	15.30	14.71
Maintenance	10.32	10.33	3.49	3.09	1.07	1.06
Transportation	9.20	9.16	18.25	18.83	1.93	1.88

ability, σ_θ, is set to match the variance of log earnings we observe in the data. Finally, given life-cycle unemployment risk and the human capital technologies, the discount fact β determines the ratio of assets held by old-to-young workers.[11] Table B.4 summarizes the calibrated parameters.

The parameters $\{w_k, \mu_k, \nu_k\}_{k=1,\ldots,18}$ primarily determine that occupational heterogeneity in mean earnings, returns to experience, and worker sorting across jobs. Table 2.3 displays the 54 data targets primarily associated with these parameters and the model fit. The occupation-specific wage rates w_k mainly control occupation k's average earnings; the mean population talent μ_k

[11] We used the ratio of liquid assets (excluding real estate) between under age 40 workers and age 55-69 workers in 2022 to calculate the data and model moments. Source data from the Distributional Financial Accounts, Federal Reserve Board Governors.

in occupation k most effects the returns to experience of workers who select into occupation k; and the occupation-specific amenity v_k are chosen to match the job choice probabilities, given the equilibrium occupation earnings profiles. These parameters must be fit simultaneously with those in Table 2.2 since the model targets are interdependent. Labor market returns to a college education will determine the willingness of households matriculate and take on student debt; at the same time, the distribution of debt among graduates will influence their subsequent investments in on-the-job training and occupational choice.

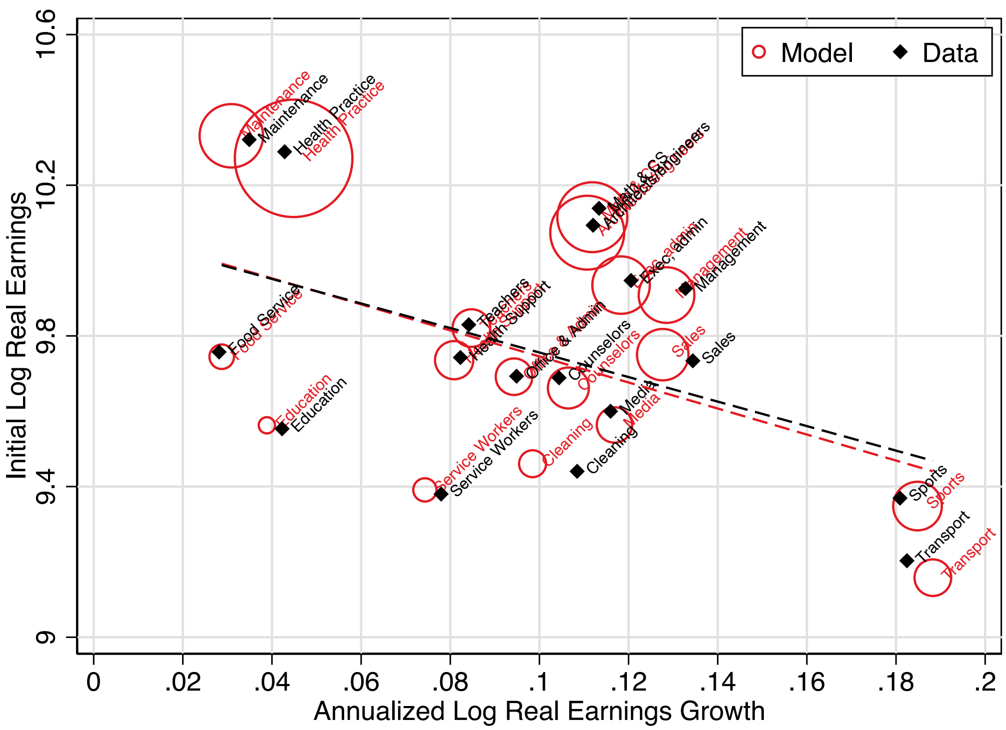


Figure 2.8. Model Fit: Occupational Heterogeneity and Student Debt

Notes: This figure plots the mean log earnings and returns to experience for each occupation from the calibrated model along with their data counterparts. The size of the bubbles represent the average amount of student debts among individual in that occupation.

Figure 2.8 plots the calibrated model's fit of the relationship between student debt and occupational heterogeneity displayed originally in Figure 2.4. While the occupational heterogeneity in mean earnings and returns to experience are targets of the calibration process, the sorting of students with different levels of student debt across the occupations is not. Nevertheless,

Figure 2.8 shows the model does a good job at replicating these patterns of sorting, consistent with the mechanism we study. The calibrated model also does relatively well at replicating the response of the aggregate earnings profile to changes in student debt, as summarized by the IV results in section 2.3. In particular, simulating the effect of an exogenous increase in student debt among the population of college graduates within the model predicts that a thousand dollar increase in debt leads to a 2.48 percentage point rise of initial earnings and -0.28 percentage point drop in returns to experience. Though untargetted, the model's predictions are close to our IV results and within range presented in the literature.

2.5 Computational Results

This section uses the calibrated model to compute the aggregate welfare and productivity consequences of federal policies which aim to alleviate student debt burdens. The analysis focuses in particular on two classes of policies: extended repayment programs and student debt forgiveness.

2.5.1 Extended Repayment Programs

Table 2.4 reports the aggregate welfare and productivity effects of extending student loan repayment periods. The current Federal Standard Repayment Plan requires students repay loans in fixed monthly installments over the course of ten years.[12] The results in Table 2.4 report the impact of extending the standard repayment period on federal loans by 2 and 5 years. In each case, the interest rate on outstanding loans is adjusted so that the net present value of individual liabilities do not change. In other words, the program has no net cost to the government and does not require any cross-household redistribution; rather, the program redistributes costs *within* each individual's lifecycle. The model captures the extensive participation margin by allowing households to select out of extended repayment programs and stay with the standard repayment

[12]While 10-years is standard repayment period on federal loans, in practice repayment periods often vary with individual circumstances. For instance, students with Direct Consolidation Loans can face repayment periods between 10 to 30 years.

plan.

Table 2.4. Welfare and Productivity Effects of Extended Repayment Programs

Notes: This table report the impacts of extending the standard 10-year repayment period to 12 and 15 years on welfare and aggregate TFPR. % of job switchers report the fraction of job switchers induced by the policy, and % choosing longer repayment is the fraction of college graduates who prefer/choose the longer repayment duration. Change in TFPR is defined as average of change in discounted lifetime income across households.

	ΔWelfare			**Δ TFPR**
	Total	**Income**	**Amenity**	
2 Year Elongated Repayment				
Everyone	0.25%	0.45%	-0.20%	0.41%
Switchers	1.80%	45.72%	-43.92%	58.55%
Stayers	0.25%	0.25%		0.22%
% of job switchers	0.38%			
% choosing longer repayment	73.49%			
5 Year Elongated Repayment				
Everyone	0.45%	0.66%	-0.21%	0.57%
Switchers	1.98%	43.35%	-41.37%	60.87%
Stayers	0.45%	0.45%		0.37%
% of job switchers	0.41%			
% choosing longer repayment	84.68%			

The results show that the extended repayment programs modestly increase welfare and labor productivity, and that the effect is monotonically increasing in the duration of the extensions analyzed. Furthermore, while the number of workers induced to switch occupations was small, ranging from 0.38% - 0.41% of the population, they experience by far the largest welfare and productivity gains from the program. While welfare increased by 0.25 - 0.45% for those who did not switch jobs, it grew by 1.80 - 1.98% for occupation switchers. Despite these large effects, the aggregate welfare effects of the program still closely mirror those of job stayers. In part, this is due to the fact that policy induced job-switchers remain a small fraction of the overall population. It is also due to the *direction* of job switches induced by policy. As the welfare decompositions in Table 2.4 show, the pattern of occupation switching was predominantly individuals leaving high amenity jobs for occupations which offered more scope for human capital accumulation and hence higher lifetime income. Due to the offsetting effects of income and amenities, the net welfare consequences of this reallocation of workers remains modest despite large changes in the underlying sources of welfare.

While the occupation switching induced by the policy has a minimal effect on *aggregate* household welfare, it is responsible for nearly half of the gains in aggregate labor productivity. By extending repayment periods, the policy change reduces the shadow cost of borrowing which stimulates investment in human capital accumulation. Households switch from high amenity jobs to those with more scope for investment on the job and hence greater returns to experience and higher lifetime earnings. Consistent with theory in the preceding sections, high talented individuals will be most effected by the constraints, making the productivity gains among induced switchers especially large. Despite only representing a small share of the population, the productivity gains among switchers nearly doubles the *aggregate* productivity gains from the policy relative to those experienced by job stayers, from 0.22% - 0.37% to 0.41% -.57%. The results suggest that the impact of student debt and credit constraints on the occupation choice of college graduates can have substantial economic effects on aggregate labor productivity.

Figure 2.9 accounts for the sources of aggregate gains in labor productivity across the occupations using a shift-share decomposition. It reports the total contribution of each occupation to aggregate productivity, as well as its within-occupation and between-occupation components. Nearly one-third of the aggregate rise in productivity is due to the reallocation of workers across occupations; the total between-occupation contribution is 0.17% while the within-occupation contribution is 0.40%.[13] The decomposition also shows that nearly all of the productivity gains are due to a small number of occupations: Sales, Management, Exec. Admin, Architects, Engineering, Math and Computer Science. As repayment elongation policies alleviate credit constraints, many workers switch to these occupations because they offer greater scope for

[13]In the case of the two-year debt elongation policy, the total within-occupation contribution is 0.24% and the between-occupation contribution is 0.17%.

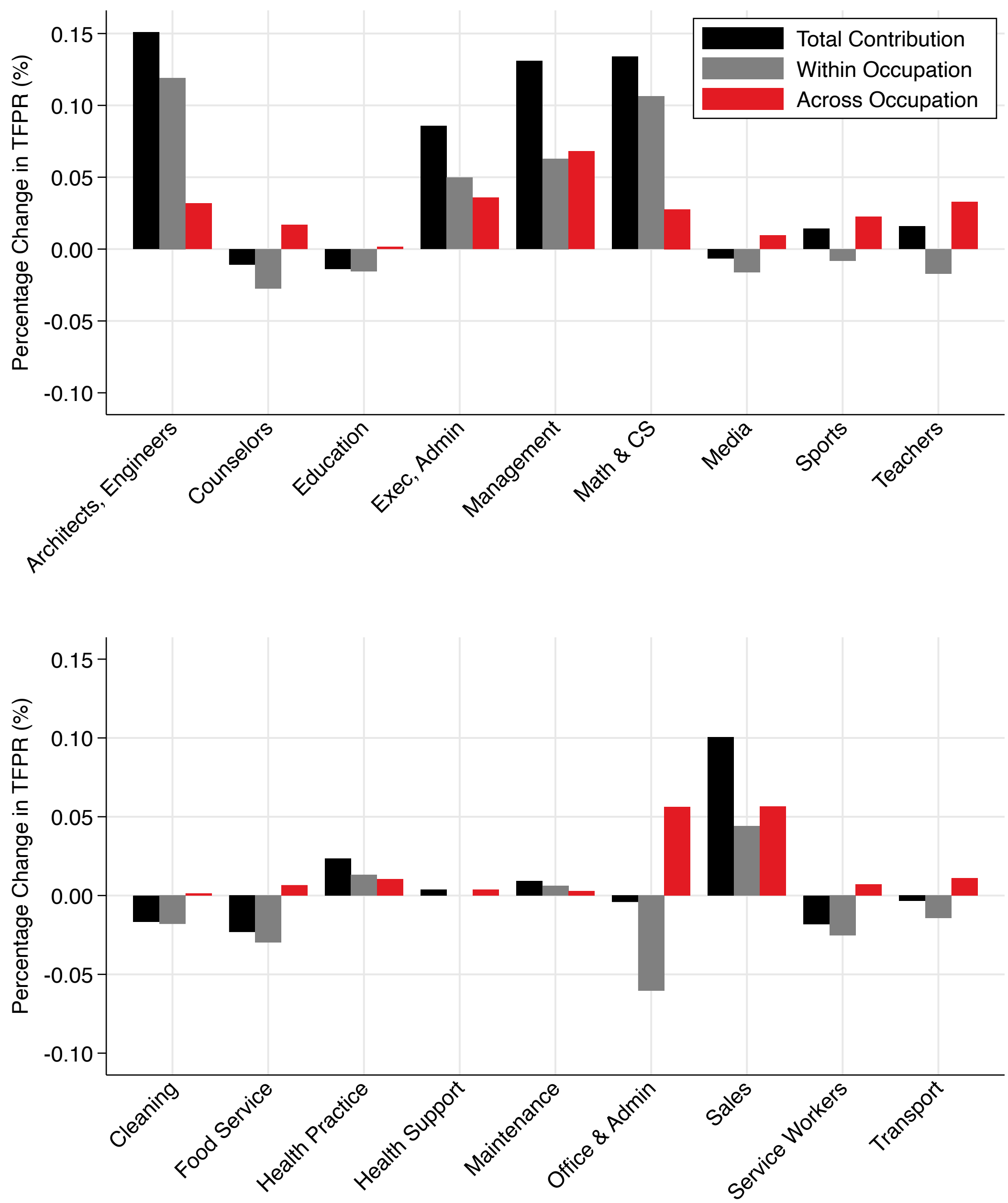


Figure 2.9. Repayment Elongation (5yr): Shift-Share Decomposition of Aggregate TFPR

Notes: This figures shows the shift-share decomposition of aggregate gains in labor productivity across occupations, for the five-year debt elongation policy. Black bars represent the total gains in the labor productivity. Gray and red bars represent the across- and within-occupation shares of the labor productivity gains, respectively.

human capital accumulation and hence higher lifetime earnings. The effect is evident in both the large between and within contributions of these occupations to aggregate productivity. Not only do repayment elongation policies lead more people to choose these occupations, they also increase the intensity of human capital investments on the job within these professions.

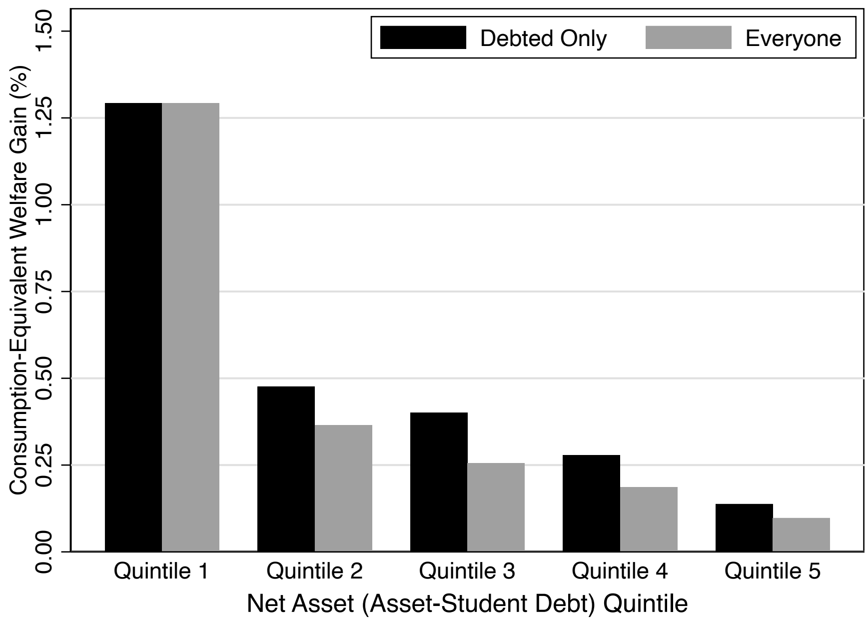


Figure 2.10. Five Year Extended Repayment

Notes: This figures shows the distribution of welfare gains under the five-year debt elongation policy across the household net asset distribution. Net asset is defined as initial asset minus total amount of student debt at the initial period. Black bars represent the welfare gains among individual who had positive student debts. Gray bars represent the welfare gains among everyone, including individuals without student debts.

Finally, even though the repayment extension policies are budget neutral and do not require cross-household redistribution, the policies nevertheless have distributions consequences. The distributional consequences are the result of the uneven way that student debt is distributed across households (recall Figure 2.1). Figure 2.10 plots the distribution of welfare gains under the five year extended repayment program across the household wealth distribution. It shows that the welfare gains at the bottom of the wealth distribution are 3 to 4 times as large as the average. Consequently, alongside welfare and productivity gains, extended repayment policies can also reduce economic inequality even without explicit redistribution.

2.5.2 Debt Forgiveness Programs

Table 2.5 reports the aggregate welfare and productivity consequences of student debt forgiveness policies. It reports the effect of forgiving outstanding student debt up to a cap of 10k, 50k, and 100k; the final cap of 100k is essentially equivalent to complete student debt forgiveness. In contrast to extended repayment policies like those analyzed in section 2.5.1, student debt forgiveness requires requires new net expenditures by the government and hence redistribution across households. To be consistent with the prevailing tax system, we assume that the costs associated with any debt forgiveness are distributed across households as lump-sum tax liabilities with the same progressivity as the existing income system.

The results in Table 2.5 show that the effect of student debt forgiveness on household welfare and productivity is non-linear in the size of the program. The smaller (10k) and larger (100k) programs actually reduce aggregate household welfare, while the smaller program also leads to a reduction in aggregate labor productivity. In contrast, the middle sized 50k program delivers modest welfare and labor productivity gains. The mixed effects reflect the distortionary consequences of redistribution, since those who bear the costs of taxation are not the same as those who benefit from student debt forgiveness. In particular, higher income households bear most of the cost while the benefits accrue to lower income households (see Appendix Figures B.2 and B.3).

The net impact of each program is determined by the countervailing effects of relaxing credit constraints for those who receive debt forgiveness and the added distortions from taxation required to fund the programs. The small 10k program leads to both modest welfare and productivity losses because the amount of debt forgiven is too small to have a substantial effect on the human capital investments of the heavily indebted population, but it still results in substantial additional tax burdens for higher income earners. In contrast, the median sized 50k program boosts aggregate welfare and productivity since it is sufficiently large to alleviate credit

Table 2.5. Welfare and Productivity Effects of Student Debt Forgiveness

Notes: This table report the effects of student debt forgiveness on welfare and aggregate TFPR. Fraction of job switchers induced by the policy reported in parentheses. Change in TFPR is defined as average of change in discounted lifetime income across households within each group.

		ΔWelfare			**Δ TFPR**
		Total	**Income**	**Amenity**	
10K cap student debt forgiveness					
Everyone		-1.05%	-0.70%	-0.35%	-0.17%
Switchers	(4.36% of treated)	-3.30%	34.00%	-37.30%	50.65%
Stayers		-1.03%	-1.03%		-0.49%
50K cap student debt forgiveness					
Everyone		0.02%	0.20%	-0.18%	0.20%
Switchers	(6.61% of treated)	0.43%	10.08%	-9.65%	8.89%
Stayers		0.01%	0.01%		0.07%
100K cap student debt forgiveness					
Everyone		-0.18%	0.09%	-0.27%	0.20%
Switchers.	(7.48% of treated)	1.47%	13.84%	-12.37%	14.15%
Stayers		-0.21%	-0.21%		-0.04%

constraints for much of the indebted population. The large 100k debt forgiveness programs again yields negative household welfare. The larger program helps a smaller population of constrained graduates at the margin, and leads to far greater tax burdens for higher income households, discouraging the pursuit of higher income through human capital accumulation.

As in the case of elongated repayment policies, the population of switchers play an important role in driving aggregate outcomes, especially in terms of productivity. The share of the population switching occupations following debt forgiveness is also much larger than under repayment elongation policies. In part, this reflects the fact that forgiving debt provides a more drastic alleviation of credit constraints compared with extending the duration of repayments. In addition, debt forgiveness induces some occupation switching among higher income households as the increases in taxation necessary to fund debt forgiveness discourages human capital investments at the top of the income distribution (Saez et al., 2012; Jones, 2022). Figure 2.11 shows how the incidence of taxation to fund debt forgiveness effects the composition of occupation switchers. Without any new taxes, the population of occupation switchers would decline monotonically with household net debt, as in the case of debt elongation policies. Using flat taxes to fund the debt forgiveness would lead to many more occupations from the middle of the income distribution. Funding debt forgiveness through progressive taxation, as in the benchmark results of Table 2.5, discourages human capital accumulation at the top of the income distribution and leads to occupational downgrading among wealthier households.

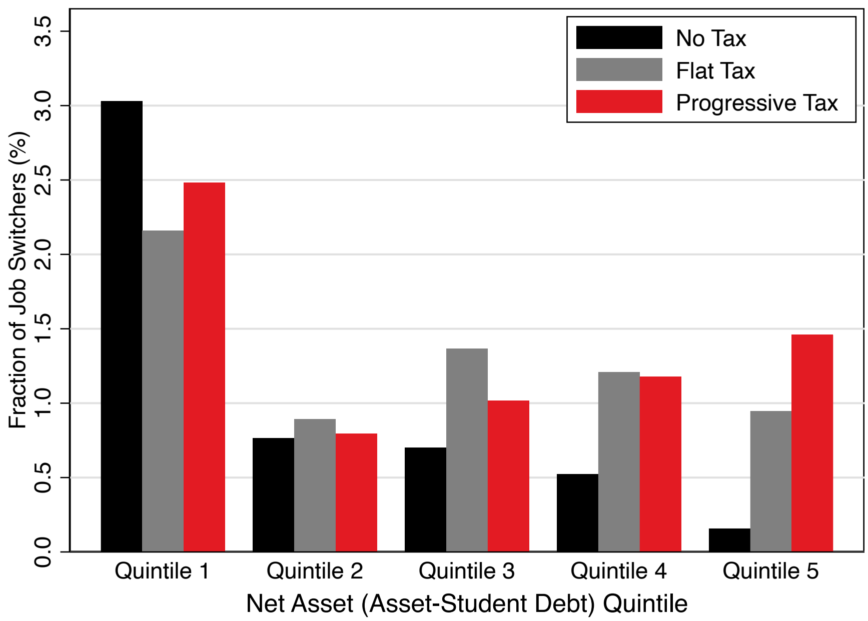


Figure 2.11. Occupation Switching by Net Assets under 30K Debt Forgiveness

Notes: This figure plots the fraction of job switchers within each of the five net asset quintiles, under 30K-cap student debt forgiveness policy with different ways of financing the policy.

Figure 2.12 displays a shift-share decomposition of the contribution of each occupation to aggregate labor productivity following a 50k-cap student debt forgiveness policy. Noteworthy is the fact that the relative contribution of each occupation has changed relative to the decomposition of the debt elongation policies in Figure 2.9. Whereas under the repayment elongation programs most switchers moved to higher earning occupations with greater scope for human capital accumulation such as Management, Engineering, Math, and Computer Science (recall Figure 2.9), most switching following debt forgiveness is toward middle income occupations, such as Teachers, Office Administration, Media, and Counseling. As discussed above, these differences in the patterns of occupation switching reflect the countervailing forces of debt forgiveness and increased taxation, which together predominantly effect households in the tails of the aggregate wealth distribution and incentives them to move to the middle.

Partly as a consequence of these differential switching patterns, the composition of within-occupation and between-occupation contributions to aggregate productivity are also meaningfully different under the two policies. Under the 50k-capped program, within-occupation effects raise aggregate labor productivity by 0.21% while between-occupation effects lead to a -0.01% decline. Moreover, the composition of these effects varies non-linearly across program sizes as the relative impact of debt forgiveness and increased taxation varies, leading to different patterns of occupation switching. Under the small 10k-capped forgiveness program, within-occupation effects reduce aggregate productivity by -0.43% while between occupation effects raise it by 0.26%. Under the 100k-capped program, which effectively amounts to forgiving nearly all outstanding student debt, the within-occupation contribution to aggregate productivity is 0.12% while the between-occupation contribution is 0.08%. While the source of productivity gains varies student debt forgiveness programs, their cumulative contribution to aggregate labor productivity always appears less than repayment elongation policies which similarly alleviate credit constraints without incurring the distortionary effects of redistributive taxation.

2.6 Conclusion

We empirically and theoretically examine how household assets affect individual career development and aggregate labor market outcomes. Using panel microdata on the early career development of recent college graduates, we document the relationship between assets, debt, occupation choice, and the earnings lifecycle. Exploiting exogenous variation in student debt burdens following changes in the generosity of university tuition grants, we find that those with more initial debt chose careers with higher initial earnings but lower returns to experience over the next 10-15 years. Initial occupation choice mediates a substantial part of the measured effect of debt on the earnings lifecycle.

To understand the data and its implications, we develop a model in which credit constraints interact with human capital decisions. High debt burdens lead workers to distort labor market choices toward careers which offer more front-loaded compensation. The adjustment process occurs both on the intensive margin, by reducing on-the-job investment, and through an extensive margin adjustment in occupation choice. Calibrating the model to replicate key features of the microdata, we investigate the consequences of extended repayment and student debt forgiveness programs on lifecycle earnings, occupation choice, welfare, and aggregate productivity.

Our counterfactual analysis suggests that both policies increase welfare and labor productivity by allowing households with large amounts of student debt to choose occupations better matched to their abilities. We find that although the fraction of household who are induced to switch occupation by the policies is small, they drive the majority of welfare and productivity gains. This means that the occupational choice channel is important to account for. While the repayment extension policy is universally welfare improving, the distributional effects of debt forgiveness policies is only welfare improving under certain parameters. This is because debt forgiveness requires the government to introduce distortionary taxation to fund the program.

2.7 Acknowledgements

This chapter is currently being prepared for submission for publication of the material. Kim, Minki; Alon, Titan; Cox, Natalie; Wong, Arlene. The dissertation author was the primary author of this chapter.

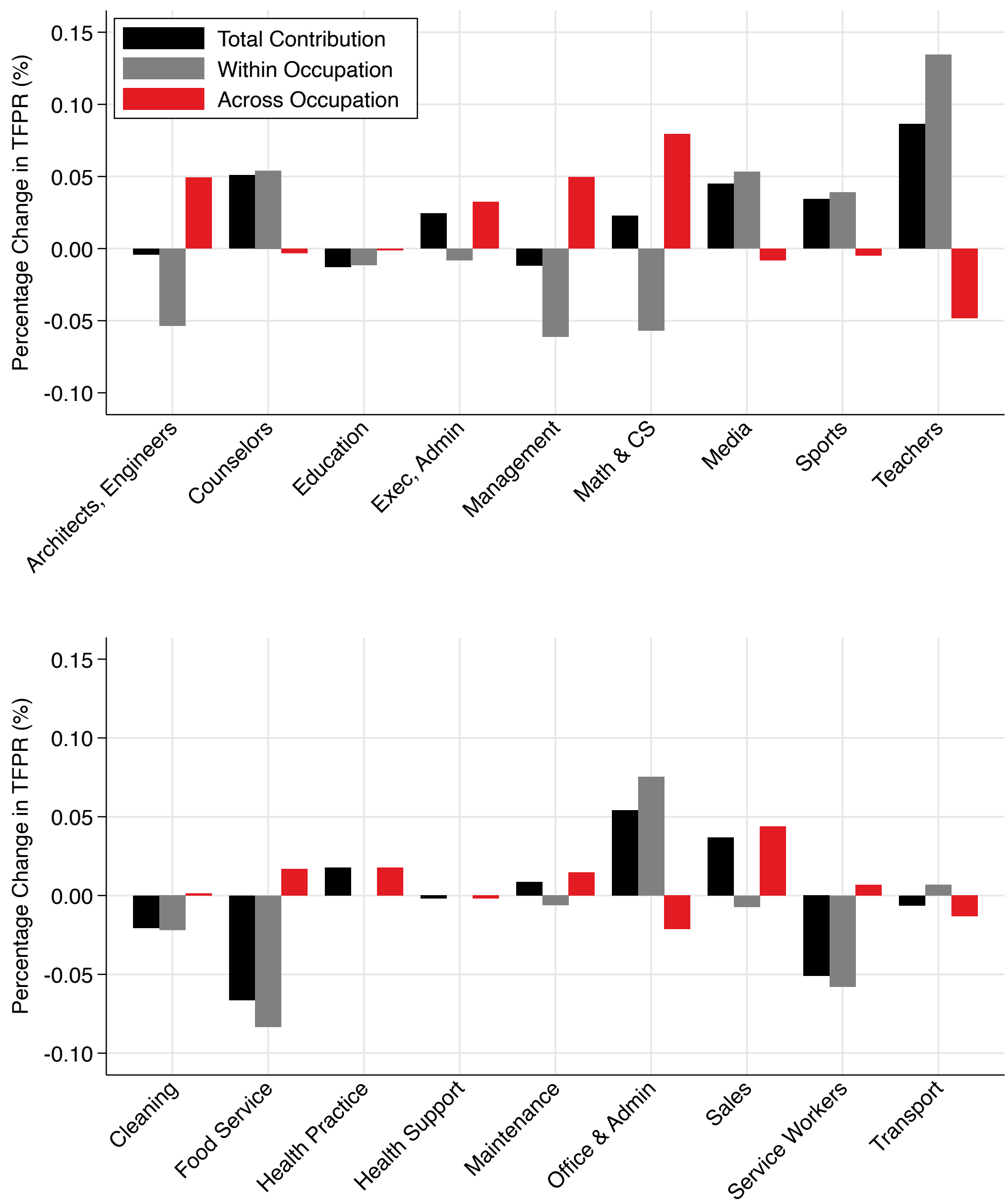


Figure 2.12. Debt Forgiveness (50k cap): Shift-Share Decomposition of Aggregate TFPR

Notes: This figures shows the shift-share decomposition of aggregate gains in labor productivity across occupations, for the 50K-cap debt forgiveness policy. Black bars represent the total gains in the labor productivity. Gray and red bars represent the across- and within-occupation shares of the labor productivity gains, respectively.

Chapter 3

Macroeconomic Effects of COVID-19 Across the World Income Distribution

3.1 Introduction

While every country has been adversely affected by the coronavirus pandemic, the damage it has wrought varied widely around the world. In this paper, we investigate how and why the pandemic's macroeconomic consequences have differed (so far) across the world income distribution. We focus in particular on variation in output and excess mortality across three broad groups of countries: low-income economies, emerging markets, and advanced economies, as classified by the International Monetary Fund (IMF). As we detail below, data from a variety of sources reveals that the pandemic's cost in terms of lives and livelihoods was roughly U-shaped in national income, with emerging markets experiencing the worst public health and macroeconomic consequences. On average, GDP per capita in emerging markets declined by 6.7 percent from 2019 to 2020, compared to 2.4 percent in advanced economies and 3.6 percent in low income countries. Excess mortality exhibits a similar pattern. According to estimates by *The Economist*, excess mortality was 75 percent higher in emerging markets than in advanced economies. While credible excess mortality data for low-income countries are still largely unavailable, the few existing estimates point to lower mortality rates.

We assess the extent to which variation in policy or preexisting economic and demographic characteristics can explain the cross-country GDP and mortality outcomes in the data.

In part, these outcomes could stem from differences in government policy responses to combat the coronavirus pandemic. While most countries enacted similar "lockdown style" policies and expanded social insurance programs, the scope of such efforts varied substantially. According to the Oxford Coronavirus Government Response Tracker, the stringency of lockdown policies aiming to restrict individual behavior (such as school and workplace closures) were somewhat stricter in emerging markets. The generosity of social insurance programs, in contrast, were substantially higher in richer countries. Accounting for these differences in policy is important because they can directly affect both fatalities and growth during the pandemic.

The cross-country variation may also arise from stark underlying differences in economic and demographic characteristics that predate the pandemic. For instance, low-income countries may face very different public health risks than wealthier ones, as they have substantially younger populations but also less developed healthcare systems. Moreover, systematic differences in the sectoral composition of employment make some countries better able to preserve income while mitigating health risks through social distancing or lockdowns. Low-income countries may benefit from their large agricultural sectors and rural populations, which provide a resilient source of income that can be sustained while limiting social contacts. On the other hand, Gottlieb, Grobovšek, Poschke, and Saltiel (2021b) show that in urban areas, the ability to work from home is far more limited in lower income countries. Combining their estimates with data on urbanization rates, we can measure the total share of labor in *social sector* employment, as in Kaplan, Moll, and Violante (2020), to capture cross country differences in the ability to work from home or with limited in-person interactions. The composite measure shows that emerging markets have the highest share of workers in social employment, due to their large urban workforce concentrated in high-contact sectors such as manufacturing and retail trade. In contrast, low-income countries have the smallest social employment shares, due to the predominance of rural agricultural work.

To investigate the extent to which these factors can explain observed differences in mortality and output, this paper follows the newly emerged literature on the macroeconomics of pandemics by combining a variant of the SICR model standard in epidemiology with a workhorse macro model. In particularly, our framework builds on the heterogeneous-agent incomplete-markets model of Aiyagari (1994), Bewley (1977) and Huggett (1996), which allows us to capture the individual-level trade-offs between consumption and health that have been the focus of most economic analysis during the pandemic. The model distinguishes between social and non-social jobs, differentiating individuals by their ability to work from home or while socially distancing. We incorporate age heterogeneity following Glover, Heathcote, Krueger, and Ríos-Rull (2020) and allow death rates to depend on a person's age, consistent with a vast medical literature. Our model also allows for a time-varying infection rate that captures, in a reduced-form way, the various other non-modeled determinants of disease progression, such as seasonal conditions, improved treatment, or virus mutation. Finally, we include constraints on peak healthcare capacity which represent limits on the ability of certain healthcare systems to treat many patients at once, due to factors like the availability of protective equipment, hospital beds, or supplemental oxygen.

In the model, the propagation of disease depends in large part on individual decisions to stay home during the pandemic or continue working in the market. The model therefore features a public health externality that creates space for welfare improving government interventions. We model lockdown policies in a simple way that is consistent with policy variation observed during the pandemic. Specifically, we feed in time-varying lockdown measures that replicate the changing stringency of government policies over the course of the pandemic, as measured by the Oxford Coronavirus Government Response Tracker (OxCGRT). In the model, lockdown stringency corresponds to the fraction of susceptible individuals who are confined to their home, where they are less likely to become infected but incur income losses depending on their job type. While we do not allow households to disobey lockdowns, individuals can voluntarily elect to work from home at any point in time. Households also receive time-varying public transfers to support or replace lost income. As with lockdowns, we set the level of public financial assistance to match the time-path reported in the OxCGRT financial support index.

To evaluate the quantitative importance of these channels in explaining the facts at hand, we parameterize the model to match key pre-pandemic economic and demographic characteristics of the United States. Parameters governing the epidemiological process are set using estimates from the relevant medical literature. We compute the model's equilibrium response to the COVID-19 pandemic as a surprise "MIT shock," where a small exogenous fraction of the population becomes infected with the virus, and then allow the disease to spread endogenously through the populous. We feed in the time-series of vaccination rates, as reported by OxCGRT, allowing a random fraction of the population to be vaccinated in each period, consistent with rates we observe in the data. We set the non-parametric component of the infection probability so that the model's endogenous disease path (nearly) exactly replicates the time-path of fatalities from COVID-19 in the United States during the pandemic. We calibrate the productivity penalty incurred during lockdowns to match the cumulative 2019-2020 year-on-year employment loss in the United States. We also allow for a one-off shock to aggregate total factor productivity (TFP), which is calibrated to match the cumulative 2019-2020 year-on-year decline in U.S. real GDP per capita.

We use the calibrated model to simulate how the United States would have fared during the pandemic if it counterfactually had the characteristics of emerging market or low-income economies. Comparing the model's predictions to the actual outcomes allows us to assess the importance of each characteristic in explaining cross-country differences in GDP declines and mortality rates. Including all characteristics, the model is able to generate the U-shaped pattern in output losses and mortality rates observed across the world income distribution. The model can fully account for the relatively larger GDP declines and higher fatalities in emerging markets compared with advanced economies. Similarly, the model correctly predicts the more modest output losses and mortality in low income countries, albeit to a quantitatively greater extent than what is observed in the data.

Simulating the contribution of each factor in isolation, we find that variation in the sectoral composition of employment is the most important factor in accounting for cross-country

GDP declines. Emerging markets suffered the greatest output losses in large part because they had high employment shares in close-contact occupations. In contrast, output declines in low income countries were substantially moderated by their large agricultural sector. The sectoral composition of employment also plays an important role in explaining cross-country mortality outcomes, along with variation in age demographics and the size of social insurance programs. In both low income and emerging markets, low levels of public financial assistance during the pandemic substantially amplified fatalities by leading many economically vulnerable individuals to continue working in the market rather than sheltering at home during times of peak infection. Our counterfactuals predict that if the United States had implemented the more limited transfer programs in low income and emerging market economies, cumulative fatalities from the pandemic would have been 50 percent greater. In low income countries, these higher fatalities were avoided largely thanks to their substantially younger populations with greater natural immunity to infection and serious illness. High agricultural employment shares, where transmission is lower while working, also reduces mortality in low income countries. In contrast, emerging market economies experienced much greater mortality because they do not benefit from favorable age demographics and also have high social sector employment shares.

Following the counterfactual simulations, we conclude the analysis by reporting multiple correlations between cross-country changes in GDP per capita during the pandemic and covariates representing the various channels embodied in our model. Consistent with our findings, the data show that the agricultural employment shares are highly correlated with GDP changes during the pandemic, while lockdown stringency exhibits a strong negative correlation. Median age and indices of government economic support show weaker correlations. Altogether, the covariates greatly reduces the observed U-shape pattern in GDP declines across the world income distribution. The result suggests that this parsimonious set of variables, and the economic mechanisms they represent, are empirically relevant in explaining cross-country macroeconomic outcomes during the pandemic.

Taken together, our analysis suggests that the comparatively worse outcomes in emerging markets, and comparably better outcomes of low income countries, were in large part predetermined by underlying economic and demographic conditions, rather than specific policy decisions. This is especially true for cross-country output losses following the pandemic. In particular, differences in quasi-fixed features like age structure, the sectoral composition of employment, and ICU capacity alone can over-account for all output losses in low income and emerging markets economies. In contrast, if all countries adopted the lockdown and transfer policies of advanced economies, GDP declines would be 1.4 and 1.0 percentage points *larger* in emerging market and low-income countries, respectively.

Policy differences also appears to play a modest role in explaining cross-country mortality differences, with the exception of transfer policies in emerging markets. The quasi-fixed features of age, employment composition, and ICU capacity can over-account for the lower fatalities in low-income countries, but only explain roughly one-third of the higher fatalities in emerging market economies. The smaller and more limited public transfer programs in emerging markets can quantitatively account for much of their remaining variation in excess mortality. While not explicitly modelled, the result suggests that financial constraints which limit the ability of governments to enact large scale public transfer programs may be an important determinant of mortality globally and across countries. A valuable goal for future research would be to help refine the quantitative importance of different policy decisions across countries in determining macroeconomic outcomes during and after the pandemic.

Our work builds on the first generation of papers addressing the aggregate effects of COVID-19 in the developing world, which were largely written in the early months of the pandemic (Loayza and Pennings, 2020; Alon, Kim, Lagakos, and VanVuren, 2020; Alfaro, Becerra, and Eslava, 2020; von Carnap, Almås, Bold, Ghisolfi, and Sandefur, 2020; Djankov and Panizza, 2020). The current paper differs in its efforts to explain observed macroeconomic outcomes through the first year and a half of the pandemic, in particular the larger declines in GDP and employment in emerging markets. Sanchez (2021) also notes the larger decline in GDP middle-income countries, but does not attempt to explain this finding. We also emphasize the inability of individuals in emerging market economies to work from home, following Gottlieb, Grobovšek, Poschke, and Saltiel (2021a,b), though we argue that low-income developing countries, on account of their large agriculture sectors, are better able to work without social interactions. As in Ma, Shapira, De Walque, Do, Friedman, and Levchenko (2021), we emphasize the stark difference in the age structure of the population across countries and the role this difference plays in quantiative impact of the pandemic and pandemic policies.

On the modeling front, our study most closely follows the structural macro work on the pandemic using models of heterogeneity in income, age and occupation/sector of employment (e.g. Acemoglu, Chernozhukov, Werning, and Whinston, 2020; Bairoliya and Imrohoroglu, 2020; Kaplan, Moll, and Violante, 2020; Glover, Heathcote, Krueger, and Ríos-Rull, 2020; Brotherhood, Kircher, Santos, and Tertilt, 2021; Chopra, Devereux, and Lahiri, 2021). Our model of disease dynamics features endogenous behavioral responses to changes in infection rates, even in the absence of government intervention, as in Greenwood, Kircher, Santos, and Tertilt (2019); Alvarez, Argente, and Lippi (2020); Krueger, Uhlig, and Xie (2020) and other studies. To our knowledge ours is the first to evaluate the quantitative predictions of a model of this sort for how the experience of emerging markets differed from richer (or poorer) countries.

Our study abstracts from many important features of reality that may also be relevant for the effects of the pandemic outside of the world's advanced economies, such as negative impacts through shocks to global supply chains (Cakmakli, Demiralp, and Ozcan, 2020; Bonadio, Huo, Levchenko, and Pandalai-Nayar, 2021), the ability to issue sovereign debt (Arellano, Bai, and Mihalache, 2020), or the ability to test and trace infections (Berger, Herkenhoff, and Mongey, 2020). We also abstract from differences in the prevalence of co-morbidities, such as diabetes and cardiovascular disease, and differential ability or willingness or ability to mask or get vaccinated. These issues would be valuable to consider in future studies trying to explain cross-country differences in the macroeconomic effects of the pandemic.

3.2 Macroeconomic Effects of the Coronavirus Pandemic by Income Level

This section presents the main facts regarding excess mortality and output losses across the world income distribution resulting from the coronavirus pandemic. Following the IMF classification, we focus in particular on three major income groups: low-income economies, emerging markets, and advanced economies. In 2019, the median GDP per capita of these three country groups was $1,124, $6,700, and $43,144, respectively, in constant 2010 USD[1]. While there is interesting variation even with these group, we focus the main part of our analysis on just the three aggregate groups. Section 5 of the paper looks at empirical patterns in the full set of countries for which data are available. Here, drawing on various data sources, we show that both output losses and excess mortality exhibit hump-shaped outcomes with middle income countries experiencing the worst. We then present in a systematic way the important differences in policy and underlying economic and demographic conditions. For each, we briefly discuss their relevance for the pandemic's impact in order to help motivate the model and quantitative analysis which follows.

3.2.1 The Impact of COVID-19

The first fact we highlight is the differential impact of the pandemic on output losses and employment declines across the world income distribution. Figure 3.1 displays the data by plotting changes in output and employment for low-income, emerging, and advanced economies. While there is considerable variance even within groups, a clear U-shaped patterns emerges in which output losses were greatest in emerging economies. GDP per capita fell by 6.7 percent and employment by 5.4 percent in emerging economies, considerably worse than both wealthier countries where output and employment losses were 4.6 percent and 2.4 percent, respectively,

[1]At the time of this writing, these are the most recent available data for emerging market and low-income countries.

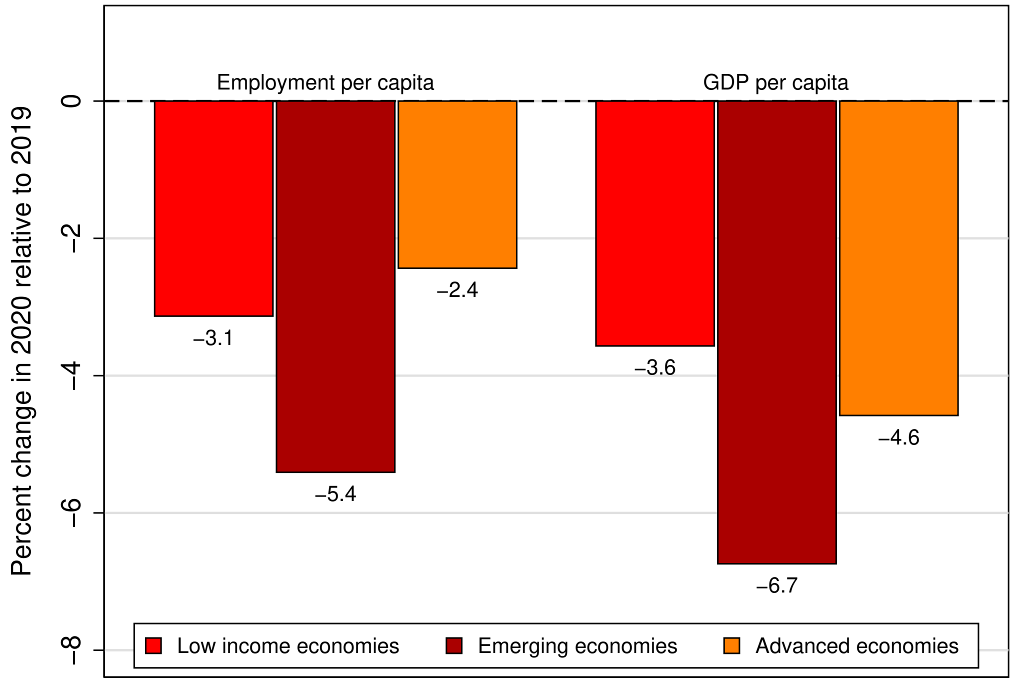


Figure 3.1. GDP and Employment Growth from 2019 to 2020 by National Income

Note: Employment data comes from the ILO Statistical Database and data on GDP per capita is taken from the World Bank World Development Indicators.

and lower income countries where those losses stood at 3.6 and 3.1 percent.[2] Such outcomes are surprising given the tremendous resources and technology that wealthy countries brought to bare in combating COVID-19, resources that low-income countries had no ability to marshal or match in any comparable way.

The second important fact pertains to the fatalities caused by COVID-19. These deaths are commonly measured using excess mortality, the difference between total deaths in a given month of the pandemic and those that would be normally expected, measured as expected deaths during the same month over the previous (typically five) years. Figure 3.2 displays the data by comparing mortality outcomes in advanced and emerging economies. As with output losses, we find that the emerging economies experienced the worst outcomes. According to estimates

[2]Appendix Figures C.1 and C.2 illustrate that the relationship also holds in the un-binned data and Appendix Figure C.3 displays similar trends in cross-country consumption data.

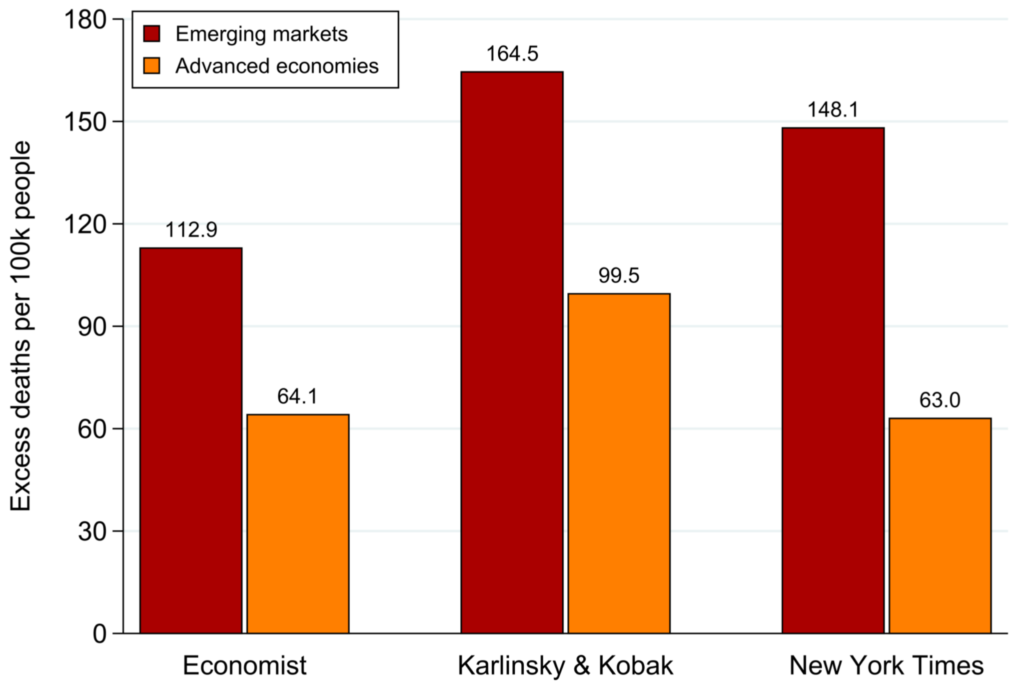


Figure 3.2. Excess Deaths from 2019 to 2020

Note: Data sourced from the New York Times and Economist excess mortality trackers, and Karlinsky and Kobak's (2021) World Mortality Database.

The Economist, excess deaths in emerging economies stands at 112.9 per hundred thousand people, which is around 75 percent higher than the average estimate for advanced economies, which experienced 64.1 excess deaths per hundred thousand. Estimates from the World Mortality Database of Karlinsky and Kobak (2021) show 164.5 excess deaths per hundred thousand people, or 65 percent larger than the 99.5 deaths per hundred thousand of advanced ones. The gap is even wider in the New York Times mortality tracker which records 148.1 deaths per hundred thousand in emerging economies, compared to 63 in advanced ones.

Internationally comparably data on excess mortality in low-income countries are more difficult to find. The most comparable statistics of which we are aware contain very few observations from low-income countries (see Appendix Figures C.4 and C.5). These data, from *The Economist* and Karlinsky and Kobak (2021), have two and five observations from the low-income group respectively. Deaths for this small set of countries average around 100 excess deaths per hundred thousand people, putting them well below the level of the emerging markets. Official data on deaths from COVID-19 in low-income show remarkably low levels of fatalities (see e.g. Appendix Figure C.6), though there is widespread belief that official statistics undercount deaths there. Our read of the literature is that there is still no clear consensus on what the true death rates have been in low-income countries, though it seems unlikely that they are worse than the high rates estimated in emerging markets such India (Deshmukh et al., 2021; Ramachandran and Malani, 2021), Mexico (Dahal et al., 2021) and Brazil (Yamall Orellana et al., 2021).

Taken together, the data reveal that the impact of the COVID-19 pandemic across the world income distribution has been highly non-linear. Emerging economies have been hit the hardest most in terms of output losses and likely in terms of excess mortality as well. Equally surprising is that the data suggest that low-income countries have fared better than advanced economies in terms of output losses, and possibly also in terms of mortality rates, despite the far greater economic and technological resources mustered by the latter to combat the crisis.

3.2.2 Differences in Policy Response

A natural candidate explanation for the cross-country variation is that they reflect differences in policy responses to the COVID-19 pandemic. While nearly all countries implemented some sort of lockdown and transfer programs, they varied widely both in the stringency of restrictions and in the generosity of transfers. The policy distinction matters for how well countries manage the endogenous path of infections through the public health externality and for the ability of households to protect themselves by staying home for prolonged periods without income.

By lockdown policies, we refer to those whose primary aim is to restrict individual behavior and social interactions to stem the spread of disease. These include school closures; workplace closures; public event cancellations; restrictions on public gatherings; closure of public transport; stay-at-home requirements; public information campaigns; and domestic and

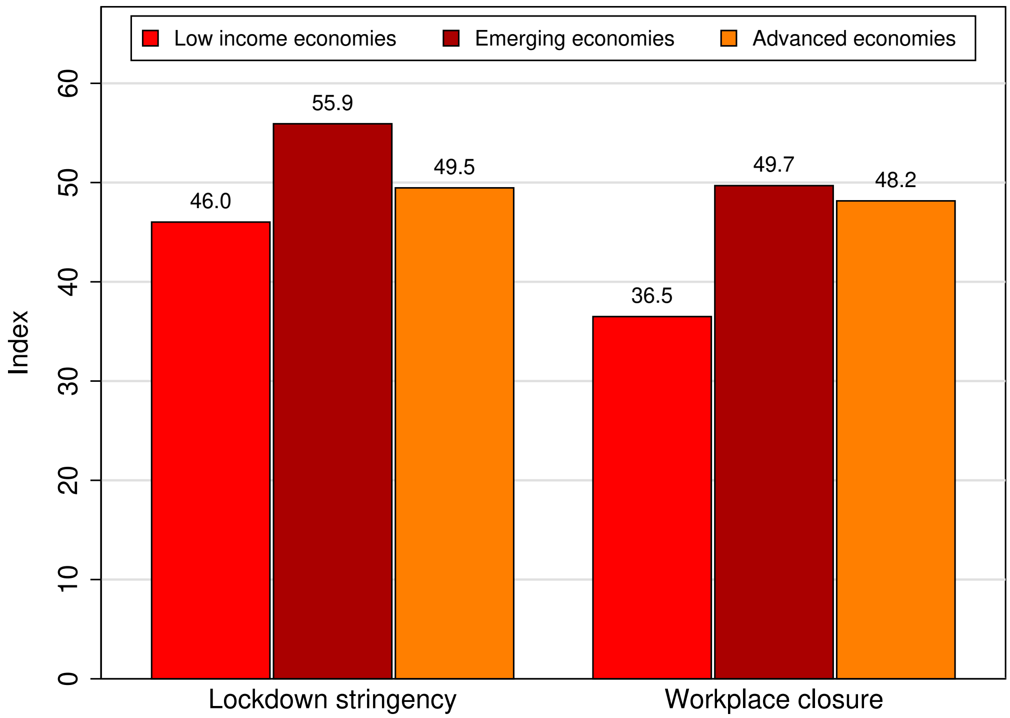


Figure 3.3. Oxford Lockdown Stringency Index

Note: The Government Stringency Index is taken from the Oxford Government Response Tracker (OxCGRT). GDP per capita is expressed at PPP and taken from Penn World Table 9.1 (Feenstra et al., 2015).

international travel restrictions. The Oxford Coronavirus Government Response Tracker's (OxCGRT) *stringency index* provides a parsimonious quantifiable measure of how strict these policies were across countries. Figure 3.3 plots the index of each country group, and shows that the most stringent lockdown policies were implemented by emerging economies (the un-binned data are displayed in Appendix Figure C.7). When we simulate lockdown policies, we implement them using the time-series of workplace closures reported by OxCGRT to be consistent with how such policies are represented in the model. As cross-country and time-series data (see Appendix Figure C.8) show, the variation in workplace lockdowns is similar to the overall stringency of policies across countries. One concern is that these data only represent *de jure* differences in policies, and that *de facto* lockdowns actually varied markedly less. Google workplace mobility data suggests this is not the case, confirming that the largest gap in workplace mobility is between the richest and poorest countries (see Appendix Figure C.9).

Another important dimension of the policy response in nearly all countries was the expansion of social insurance payments, such as unemployment benefits. These payments are viewed as critical to offsetting lost income and make isolating at home economically feasible for those with low savings or little income. However, as the crisis unfolded it quickly became clear that governments in many developing countries lacked the fiscal capacity to sustain substantial transfers to major segments of their population for very long. Consequently, we observe substantially more cross-country variation in the size and scope of social insurance programs than in lockdown policies.

Figure 3.4 provides two measures capturing the scope and generosity of transfer programs implemented in response to COVID-19 across the world income distribution. The left side histogram plots national pandemic spending as a share of GDP, which includes comprehensive measures of budgetary fiscal support to individuals and firms estimated by the IMF. While pandemic spending appears similar in low-income and emerging economies, they are only about one-third the spending undertaken by advanced economies which reached nearly 10 percent of GDP. The right side histogram displays the Oxford's Government *Economic Support Index* which records financial assistance programs such as income replacement and debt relief for individual citizens. The index should be interpreted as an ordinal measure of economic assistance for individual citizens in that it does not include support to firms or business and does not take into account the total fiscal value of economic support programs. Nevertheless, the data reveal a similar pattern with spending on economic support rising monotonically with national income.[3]

These cross-country differences in lockdown policies and public insurance programs are even more apparent when one examines the underlying components of the OxCGRT's indices which are displayed in Table 3.1. The first noticeable feature is that low-income countries have the least stringent policies in every lockdown category, and in all other categories except "Facial

[3]The greater cross-country variation in economic support policies, as compared to lockdown policies, is most apparent in these underlying data. See Appendix Figures C.10 and C.11.

Table 3.1. Oxford Covid-19 Government Response Indices in 2020

.

Notes: Countries are grouped into low income, emerging markets, and advanced economies using the IMF's economic classification of countries. Data in the table is the average level of the Oxford Covid-19 government response tracker by country income group.

	Country Income Group		
Index	Low-Income	Emerging Markets	Advanced Economies
Panel A: Included in both Stringency and Health & Containment Indices			
School closures	53.8	64.8	50.1
Workplace closures	34.6	47.0	45.1
Cancellation of public events	57.0	69.4	63.7
Restrictions on public gatherings	50.9	59.5	61.3
Closure of public transport	22.5	32.0	17.8
Stay at home requirements	25.0	35.7	24.9
Restrictions on internal movements	32.9	47.7	31.8
International travel controls	57.6	63.6	63.4
Public information campaigns	79.7	83.8	87.0
Panel B: Included only in Health & Containment Index			
Contact tracing policy	54.4	61.5	67.6
Facial coverings	43.8	46.4	37.3
Testing policy	37.9	52.2	58.8
Vaccination policy	22.8	31.3	35.3
Protection of the elderly	19.4	40.8	57.3
Panel C: Included only in Economic Support Index			
Income support	17.3	29.3	57.8
Contract/Debt relief	31.0	49.6	58.9
Observations	52	67	33

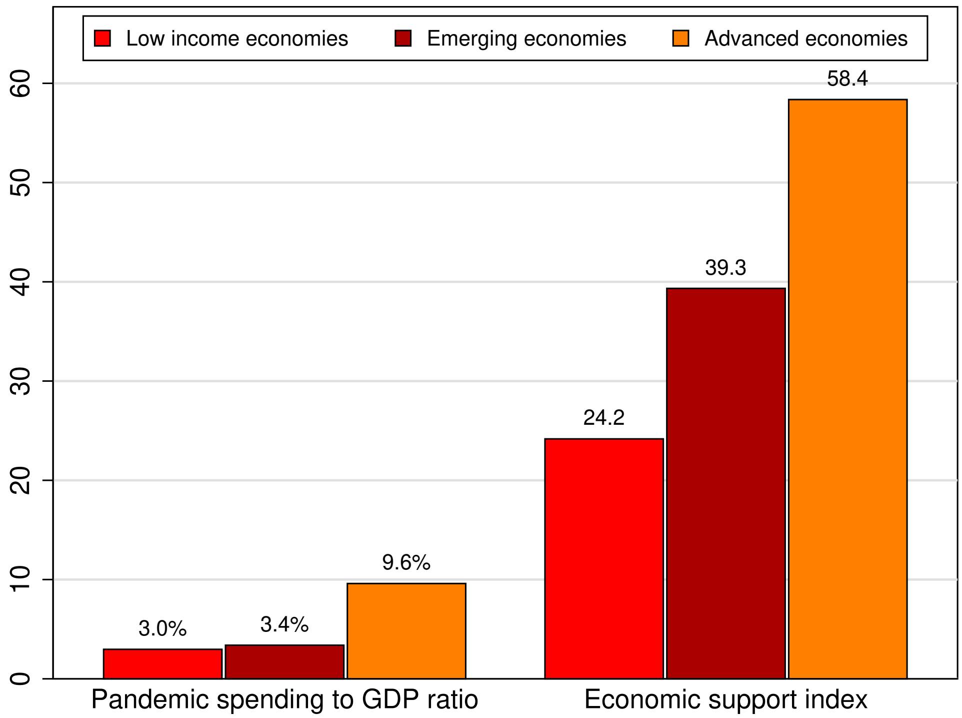


Figure 3.4. Pandemic Spending and Economic Support

Notes: The left side histogram plots the ratio of pandemic spending to GDP, taken from the IMF. The right side histogram displays the Oxford Economic Support Index available through the Oxford Coronavirus Government Response Tracker's (OxCGRT).

Coverings." The near opposite is true for emerging economies which have the most stringent policies across all sub-categories of lockdown measures (Panel A) except "Public Information Campaigns." The largest deviations in emerging economy lockdowns pertain to the closure of public transport, stay at home orders, and restrictions on internal movements. This is notable since these measures likely imposed the largest restrictions on commercial activity, especially in emerging economies where the ability to work from home is not widespread (see Section 3.2.5) and substituting to e-commerce and delivery services is limited by infrastructure. Finally, it is interesting to note that the stringency of emerging economy policies does not extend beyond lockdowns; as Panels B and C show, direct public health interventions and economic support policies were generally less encompassing in emerging economies. Taken altogether, the scope of differences in the stringency and aim of policies across the world income distribution offer ample scope for them to drive the differences in outcomes we observe in the data.

3.2.3 Differences in Population Structure

It has been well known since the beginning of the pandemic that COVID-19 poses dramatically greater health risks to older individuals, in particular those over the age of 65 (Ferguson et al., 2020; Glynn, 2020). Early centers of infection in the west, such as Italy, experienced health impacts concentrated on those in this older age range, with particularly severe fatality rates for those in their 80s and 90s. At the same time, the number of deaths linked to COVID-19 for those under 20 has been negligible, though certainly not zero.

A basic demographic difference between advanced and developing economies is that populations are far younger in the developing world. Since fatality rates from COVID-19 are very low for young individuals but rise sharply with age, these demographic differences suggest much smaller populations of vulnerable individuals in the developing world. One can see these demographic differences starkly when looking at cross-country data on the median age. Figure 3.5 plots the median age against GDP per capita in a set of 158 countries using data from UN Population Division and Penn World Tables. Data from the UN Population Division show that countries in the bottom quartile of the world income distribution have a median age of 19.1 years. Nigeria, Africa's most populous country, has a median age of 17.9, while countries like Angola and the Democratic Republic of the Congo have median ages of just 16.4 and 16.8 years old. By contrast richer countries like Italy, the United Kingdom and France have median ages of 45.9, 40.2 and 41.2, respectively.

Another statistic indicative of the much smaller vulnerable population in the developing world is the cross-country data on the population above 65. In the world's poorest countries the fraction of the population that is above age 65 is negligible, with an average of around 3 percent for countries in the bottom quartile of the world income distribution. The older population is much larger as a fraction of the total in richer economies, and reaches around one quarter of the population in Japan. Among countries in the topic quartile of the world, the average is about 15

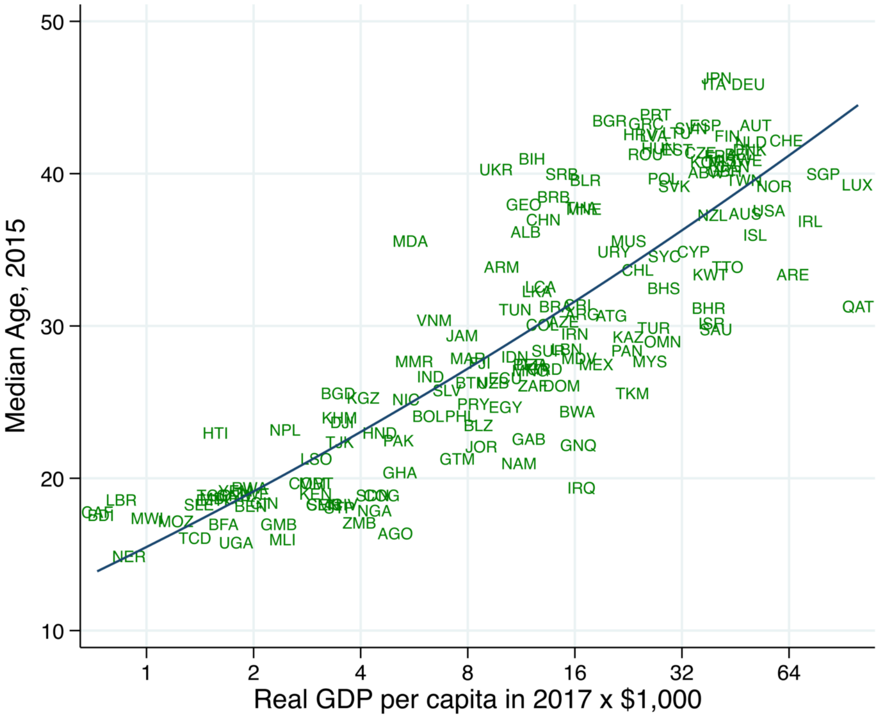


Figure 3.5. Median Age of the Population

Notes: Median age data corresponds to 2015 and is from the UN Population Division. GDP per capita is expressed at PPP and taken from Penn World Table 9.1 (Feenstra et al., 2015).

percent of the population being above age 65 (see Appendix Figure C.12).

It is hard to look at statistics like these and not see how different the impacts of COVID-19 will be in less developed countries. Concretely, while almost everything about COVID-19 suggests a more severe impact in less-developed countries, the far younger demographic is clearly in their favor.

3.2.4 Differences in Healthcare Capacity

Developing countries typically have substantially less ability to control disease than do richer countries. Sanitation and hygiene are more of an issue given the lack of widespread piped water and functioning sewage systems. Health infrastructure, especially hospital and health clinic capacity, is also less developed. For mild cases of COVID-19 infections, this may make little differences, as bed rest is likely to suffice in these mild cases. However, for critical cases, the lack of intensive-care capacity is a clear disadvantage for developing countries in their attempts to save lives during the pandemic.

Figure C.13 plots the number of hospital beds per 10,000 people, as reported by the World Health Organization (WHO), against GDP per capita. The number of hospital beds is an imperfect measure of hospital capacity for many reasons, most importantly because it is not a bed per se that helps critical patients recover from COVID-19 but trained doctors, equipment like ventilators, and appropriate pharmaceuticals. Still, for lack of more comprehensive cross-country data, we take hospital beds as a proxy for medical care capacity.

By this metric there are stark differences in healthcare capacity across countries. Richer countries, which have quite some range amongst themselves, average around 49 hospital beds per 10,000 people. Countries like Japan and Korea have even more beds per capita, having 134 and 115 beds per 10,000 people, respectively. This is still far higher than the capacity in developing countries, which is a paltry 12 beds per 10,000 people on average in the bottom quartile of the income distribution. In Appendix Table C.1, we report the availability of intensive care unit (ICU) beds and per capita healthcare costs across a limited set of countries. Consistent with the patterns observed from the number of hospital beds, it appears that low income countries possess significantly fewer ICU beds than high income countries.

3.2.5 Differences in Sectoral Composition of Employment

It is widely known that the sectoral composition of employment varies systematically with economic development. These differences are important because commercial disruptions brought on by COVID-19 and the resulting lockdowns differed substantially by occupation. Non-essential jobs that could not be performed remotely or while socially distancing experienced the largest and most sustained drops in employment throughout the recession; in contrast, occupations that were amenable to working from home experienced minimal disruption and some even flourished during the pandemic. In our model, we highlight two systematic differences in the composition of employment between advanced and developing economies which are relevant to the pandemic's macroeconomic outcomes across countries: the share of rural employment and the extent to which the urban workforce can work from home.

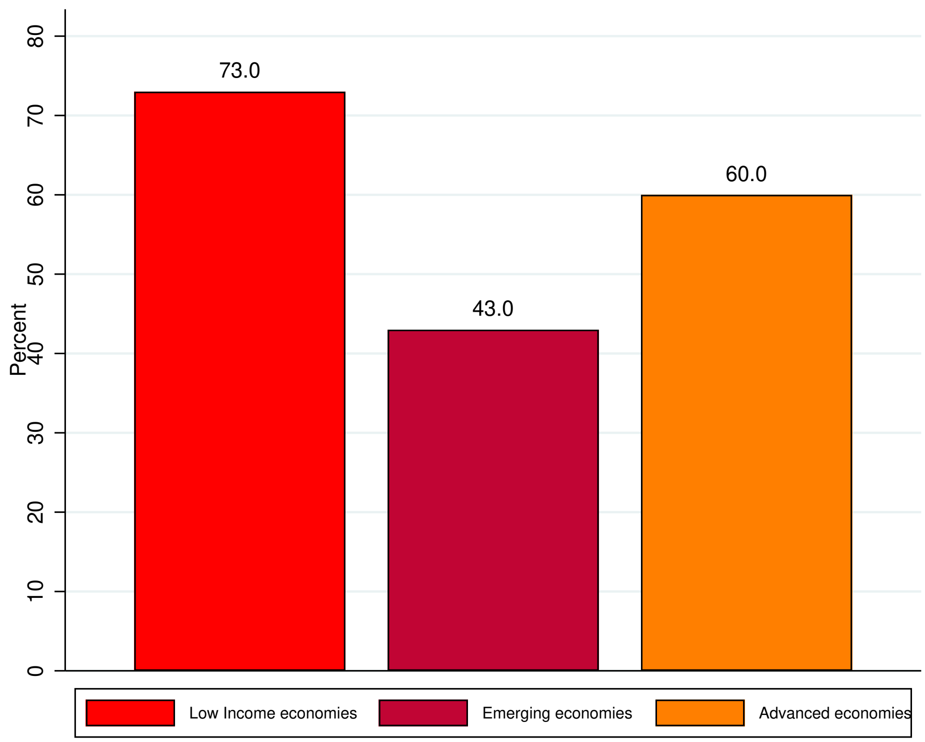


Figure 3.6. Non-Social Sector Employment Share

Notes: The non-social sector includes rural employment and urban jobs that can be done from home, as estimated by Gottlieb et al. (2021b). See text for details. GDP per capita is expressed at PPP and is taken from the Penn World Table 9.1 (Feenstra et al., 2015).

It is well known that the share of agricultural employment varies widely with economic development (see Figure C.14). In the poorest countries, over 70 percent of the population is engaged in agricultural work on average, often subsistence farming on family plots; in advanced economies, that share is in the low single digits. The high agricultural share, while often considered a drag on economic modernization, offers a resilient source of income during pandemics. A good deal of agriculture in the developing world takes place on household-run farms, allowing it to continue during "stay-at-home" orders. Even in the absence of lockdowns, farming can often continue while socially distanced or with contact restricted to household members. Agricultural workers therefore do not face the same stark trade-offs in choosing between protecting their health or incomes since farming can often continue without substantially increasing the risk of infection. Consequently, while agricultural workers may be vulnerable because of low wages, their employment is more resilient to large losses from lockdowns or voluntary self-isolation.

Outside of the agriculture sector, labor markets in lower income countries are characterized by widespread informality and employment concentrated in high-contact sectors. Large informal sectors will generally make economies more vulnerable to COVID-19 since, like agriculture, these jobs generally pay low wages while, unlike agriculture, most informal jobs cannot be performed from home or while socially distancing. To summarize these effects at the country level, we follow Kaplan et al. (2020) and aggregate employment into *social* and *non-social* sectors. Social sector workers have limited ability to work from home and suffer large income losses during lockdowns, while non-social sector workers can substitute more easily to remote work. We calculate the non-social sector share to include rural employment and all urban jobs that can be worked from home. For the latter, we use the cross-country estimates of Gottlieb et al. (2021a) which are constructed using worker level data on the task-content of jobs in urban labor markets. Figure 3.6 displays the resulting estimates of non-social employment and illustrates that it varies substantially across countries. Emerging market economies have the lowest ability to work from home, with only 43 percent employed in non-social, low-contact jobs. In advanced economies, the non-social share is 60 percent, due to the greater number of high skill, professional jobs. However, the non-social share is largest in low-income countries, at 73 percent of aggregate employment, driven by the large agricultural labor force.

As a consequence of theses differences in the sectoral composition of employment, emerging market economies are more exposed to economic losses during the pandemic. Having less jobs that can be done from home or while socially distanced leads to greater economic losses during lockdowns and workplace closures. Moreover, in the absence of robust transfers, many social sector workers can become desperate and so voluntarily elect to continue working, rather than shelter at home, during times of peak infection. Such decisions will generally provide only marginal income gains, while amplifying the infection risk for the whole population through the public health externality. Large social sector employment can therefore be a liability for emerging market countries fighting COVID-19, as these workers are particularly vulnerable with limited options to avoid increasing their risk of becoming infected, or infecting others.

3.3 Model

Our analysis draws on a quantitative heterogenous-agent macroeconomic model, with epidemiology as in the SICR framework, to analyze variation in the macroeconomic consequences of the COVID-19 pandemic across the world income distribution. The model is equipped with several features that vary between advanced and developing economies that are relevant for the pandemic's impact, as motivated by the data presented in the previous section. These include uninsurable idiosyncratic health and income risks, age heterogeneity, government transfers and lockdown policies, healthcare capacity, and the ability to work from home across sectors. This section now presents these features in detail.

3.3.1 Households and Preferences

The economy is populated by a unit mass of heterogenous individuals who make consumption and savings decisions subject to idiosyncratic income and health risks. Individuals differ in their age $j \in \{\text{young adult , old adult}\}$ and permanent labor productivity $z \sim G$. Time is discrete and each period represents two weeks. Preferences are given by:

$$U = \mathbb{E}\left[\sum_{t=0}^{\infty} \beta_j^t \left\{ \log(c_t) + \bar{u} \right\}\right], \tag{3.1}$$

where the discount factors β_j^t capture age heterogeneity in the population, and $\beta_{\text{young}} > \beta_{\text{old}}$. This specification follows the tractable formulation of Glover et al. (2020) that abstracts from explicitly modeling age, appealing to the logic that pandemics are sufficiently short-lived relative to entire lifetimes. It thus suffices to model only the expected number of years left to live, which is captured by the heterogeneity in discount factors. The term $\bar{u}$ represents the flow utility value of being alive, following the specification of Jones and Klenow (2016), and represents the reason that model households try to avoid fatality risk. Once an individual dies, they receive a fixed utility level that potentially depends on their individual characteristics, as we describe below.

There are two sectors, which we denote as social ($s = S$) and non-social ($s = N$). Households in sector s supply labor to a representative firm where they can earn wage w_s per effective hour worked. We assume that households are born with the sector they supply labor to and cannot switch. The social sector represents the workers with occupations that provide limited ability to work remotely. Examples of these occupations in the social sector includes waitresses and hair dressers. The non-social sector represents occupations that can be done with low levels of social contact. Examples of such occupations include farmers in the agricultural sector who can work in their fields while distancing from others and college professors who can easily teach remotely.

At the beginning of life, workers draw their permanent productivity, $z \sim G$. Incomes in both sectors are also subject to idiosyncratic productivity shocks as in Bewley (1977), Huggett (1993) and Aiyagari (1994). Specifically, we assume that individual labor productivity in each sector is composed of the sector-specific permanent component z and an idiosyncratic component v following the stochastic process:

$$\log v_{t+1} = \rho_v \log v_t + \varepsilon_{t+1}, \qquad \varepsilon_{t+1} \sim F(0, \sigma_v). \tag{3.2}$$

We include idiosyncratic income risk because developing countries are far from having full insurance, and so accounting for how people insure themselves in response to policies which may keep them away from work for prolonged periods of time is a first order consideration.

After observing their income realization, households make consumption and savings decisions given the interest rate, r, and subject to a no-borrowing condition, $a \geq 0$. Formally, the budget constraint of a household in sector s before the pandemic is given by:

$$c + a' \leq (1-\tau) w_s z v n + (1+r) a + T \tag{3.3}$$

where τ is the income tax rate and T is government transfers.

3.3.2 Aggregate Production Technology

The economy produces a single final good by combining capital with labor services supplied by the three sectors. The aggregate production technology is given by:

$$Y = A L^{\alpha} K^{1-\alpha},$$

where A is the total factor productivity and $0 < \alpha \leq 1$ is labor's share of value-added. We abstract from the domestic capital market. The aggregate capital stock is composed entirely of foreign sources, $K = K^F$, which can be rented at an exogenously given international rental rate r^F and which depreciates at rate δ. Aggregate labor depends on the total supply of labor services from the social and non-social sector,

$$L = L_S + L_N.$$

The assumption that social and non-social labor are aggregated to produce a single consumption good is a simplifying one. While some labor may be social due to the production technology requiring social interactions with other workers, certainly a substantial portion of social sector labor requires face-to-face interactions with customers and thus require social interactions to consume. We abstract from these notions for simplicity and, to the extent they operate, we allow them to be absorbed by the exogenous component of the public health externality $\beta_t^{\mathbb{I}}$ and the pandemic's shock to TFP A (see section 3.4.1). Furthermore, appendix Figure C.15 plots the change in consumption-related mobility from the Google Community Mobility Report (Bavadekar et al., 2020) for advanced, emerging, and low income economies.

The change in consumption-related mobility over the course of the pandemic looks very similar across country income groups, and so our assumption of a single consumption good likely does not distort our cross-country analysis.

3.3.3 Credit and Capital Markets

Credit market incompleteness prevents households from borrowing against future earnings. As a result, individuals must maintain non-negative assets in formulating their consumption plans subject to (3.3), giving rise to hand-to-mouth consumers as well as a precautionary savings motive in response to idiosyncratic health and income risks. The precautionary motive is important for getting aggregate welfare measurements correct since it creates another feedback between the epidemiological and economic dynamics, as individuals withhold some consumption to increase precautionary savings in response to the pandemic's onset.

3.3.4 Public Health and Hospital Capacity

Households face idiosyncratic health risk which can reduce their labor productivity and increase the probability of dying. Susceptibility to infection is determined in part by economic decisions taken by households. Once infected, progression of the disease depends on an individual's age and the availability of public health infrastructure offering treatments.

Health risks are modeled using an SICR epidemiological model with five health states: susceptible ($\mathbb{S}$), infected ($\mathbb{I}$), critical ($\mathbb{C}$), recovered ($\mathbb{R}$), and deceased ($\mathbb{D}$). We denote by N_t^x the mass of individuals in each health state $x \in \{\mathbb{S}, \mathbb{I}, \mathbb{C}, \mathbb{R}, \mathbb{D}\}$ at time t and use $N_t = N_t^{\mathbb{S}} + N_t^{\mathbb{I}} + N_t^{\mathbb{C}} + N_t^{\mathbb{R}}$ to measure the non-deceased population. Figure 3.7 illustrates how these states evolve:

The probability a susceptible person becomes infected is given by:

$$\pi_t^{\mathbb{I}} = \beta_t^{\mathbb{I}} \times \frac{N_t^{\mathbb{I}}}{N_t}$$

where $\beta_t^{\mathbb{I}}$ is the time-varying infection rate, reflecting the disease's natural progression (e.g. new

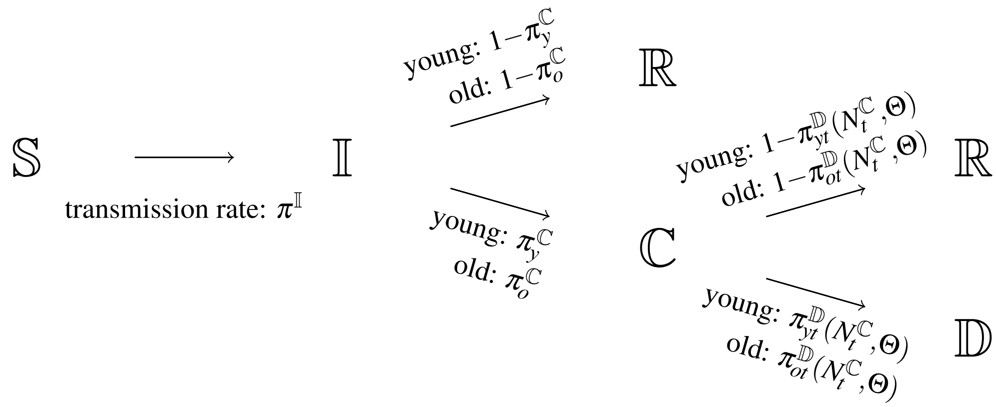


Figure 3.7. Dynamics of Health States and Transition Probabilities

variants), seasonal variations in infection rates, better medical treatments and other un-modeled factors that change infection rates over time.

Individuals who contract the virus experience a proportional drop in productivity of $1-\eta$ for one model period (two weeks), at which point they either recover or enter a critical health state. The probability of becoming critically ill depends on an individual's age and is given by π_j^C. Those in critical health are unable to work and require hospitalization. The likelihood of recovery in the hospital depends again on their age in addition to the availability of public health infrastructure, such as ICU beds and ventilators. In particular, the fatality rate of a critically ill patient of age j is given by:

$$\pi_{jt}^{\mathbb{D}}(N_t^{\mathbb{C}},\Theta) = \begin{cases} \pi_j^{\mathbb{D}} & \text{if assigned ICU bed} \\ \kappa \times \pi_j^{\mathbb{D}} & \text{if not assigned} \end{cases}$$

where π_j^D is a baseline fatality rate for age j individuals in critical health and κ governs the impact on fatality rates of strained hospital resources. Whether or not a critically ill patient receives an ICU bed depends on overall hospital capacity and the number of other patients. Specifically, letting Θ denote hospital ICU capacity, the probability a new patient receives an ICU bed is given by $\min\{\Theta/N_t^C, 1\}$. In other words, all critically-ill patients receive an ICU bed if hospital capacity constraints are not binding, and beds are rationed amongst the critically-ill with probability Θ/N_t^C when constraints bind.

3.3.5 Lockdowns and Voluntary Substitution Away from the Workplace

While the disease's progression is exogenous, the probability a susceptible person becomes infected depends on endogenous economic decisions and the prevalence of infections in the population. To incorporate the feedback from economic behavior to infections, we allow individuals to lower their degree of exposure to the virus by voluntarily reallocating working time from the market to home. Specifically, in each period we allow workers to choose between going to the workplace or working remotely. Remote work involves less social contacts, providing protection from becoming infected, but leads to a productivity penalty.

In particular, the productivity penalty of working remotely depends on an individual's sector of employment and is parameterized by ϕ_s, for $s \in \{S, N\}$. An individual who works n hours remotely in sector s provides only $\phi_s n$ units of effective labor, where $0 \leq \phi_s < 1$. We assume that $\phi_S < \phi_N < 1$ to capture the notion that jobs in the non-social sector are better suited to be conducted remotely. We parameterize the health benefits of remote work by lowering an individual's probability of infection by proportion ξ when working from home. Consequently, the probability a susceptible person becomes infected is given by:

$$\pi_t^{\mathbb{I}} = \begin{cases} \beta_t^{\mathbb{I}} \times N_t^{\mathbb{I}}/N_t & \text{if go to workplace} \\ \beta_t^{\mathbb{I}} \times N_t^{\mathbb{I}}/N_t \times \xi & \text{if work remotely} \end{cases}$$

Given the trade-offs between productivity penalties and lower infection risks, individuals optimally choose whether or not to go to their workplace each period according to,

$$V = \max\{V^w + \varepsilon_w, V^r + \varepsilon_r\}$$

where V^w and V^r represent the value functions of being in the workplace and of working remotely, respectively. For each of the two options, we also introduce a taste shocks ε_w and ε_r, which are drawn *i.i.d* from a Gumbel distribution with variance σ_g. The variance σ_g is calibrated to match the fraction of labor working remotely in the pre-pandemic steady state.

Lockdowns

Infection rates can be further mitigated by containment policies, such as lockdowns. These policies reduce the virulence of infection by forcing some individuals to switch from market work to working from home, which reduces both their individual infection risk as well as the economy-wide infection risk through the public health externality. As in Kaplan et al. (2020), we model lockdown intensity as a certain fraction of the workforce being forced to work remotely through, for instance, stay-at-home orders. The stringency of lockdowns varies both across time and countries. Following Bick et al. (2020), we assume that 70 percent of the workers are forced to work from home under a full lockdown. We also assume lockdowns are not targeted and so are applied with the same intensity to all household groups.[4]

3.3.6 Vaccinations

Susceptible individuals can also obtain immunity through vaccination. In each period, every susceptible individual faces a non-negative probability of receiving a vaccination. Once vaccinated, an individual receives immunity and joins the recovered population. The particular probability of vaccination in any period is taken from the actual time path of vaccinations administered in the United States. Details on the vaccination data series used are provided in the calibration section.

[4]For example, if the lockdown intensity if 70 percent in a given period, then 70 percent of each group (young social, young non-social, old social, old non-social) are required to work remotely.

3.3.7 Government and Taxation

The government has power to tax, transfer, and impose economic lockdowns subject to running a balanced flow budget which satisfies,

$$B_t + \tau \int y(a,x,v)dQ = T$$

where $y(a,x,v)$ is pretax income for individual $(a,x,v) \sim Q$, τ is the prevailing tax rate, and T is aggregate transfers to households. In addition to tax revenue, we allow developing countries access to emergency bonds, B_t, which can be used to finance additional welfare transfers during government imposed lockdowns. The source of these funds is international donors and multinational institutions such as the IMF, World Bank, and World Health Organization. Funds borrowed for emergency transfers accrue interest at rate $1+r^F$ until the pandemic ends, at which they are repaid through annual annuities. Formally, emergency transfers are given by:

$$B_t = \begin{cases} \bar{B} & \text{during the lockdown} \\ -\frac{r^F}{1+r^F} \times \sum_{t_l - t_s}^{t_l - t_e} \left(1+r^F\right)^t \bar{B} & \text{after pandemic ends} \\ 0 & \text{otherwise} \end{cases}$$

where $\bar{B}$ is the size of per-period emergency transfers during lockdown, which we take parametrically, and t_s, t_e, and t_l index the lockdown's start, the lockdown's end, and the pandemic's end, respectively.

3.4 Quantitative Analysis

In this section, we discuss the calibration strategy, validate the model's fit, and present our counter-factual results. To evaluate the quantitative importance of each channel in explaining the cross-country variation in outcomes, we calibrate the model to match the U.S. economy and then vary key economic and demographic characteristics of the U.S. to match those of low-income and emerging economies. For each variation, we display the dynamic path of output and fatalities predicted by the model. To identify the most salient channels, we report the cumulative effects of each counterfactual on the U.S. economy compared to the calibrated benchmark.

3.4.1 Data Sources and Calibration

For expositional clarity, we divide the calibrated targets into three broad categories corresponding to those governing economic mechanisms, those controlling epidemiological dynamics, and those delineating differences between the advanced, emerging, and low-income countries.

Table 3.2. Calibration of Economic Parameters

Var	Description	Value	Source / Target
r^F	Exogenous interest rate	0.0006	Pre-COVID T-Bills rate 1.5%
ρ_v	Persistence of idiosyncratic income shock	0.91	Floden and Lindé (2001)
σ_v	St.Dev of idiosyncratic income shock	0.04	Floden and Lindé (2001)
α	Labor share	0.6	Gollin (2002)
β_y	Discount factor for the young	0.9984	Glover et al. (2020)
β_o	Discount factor for the old	0.9960	Glover et al. (2020))
σ_g	Variance of remote / non-remote work taste shock	0.0101	Pre-COVID Remote Workers 8.2%
ϕ_n	Productivity remote work, non-social sector	1	Barrero et al. (2021)
ϕ_s	Productivity remote work, social sector	0.6	US COVID-19 Employment Declines - 6.4%
$A(P)$	Pandemic Total Factor Productivity	1.02	COVID-19 Output Declines -4.6%

Table 3.2 reports the parameters that govern the core economic dynamics of the model. Population demographics are modeled using age dependent discount factors accounting for differences in the remaining years of life for young and old workers. The age specific discount factors are taken from Glover et al. (2020), and the stochastic income processes are taken from Floden and Lindé (2001), who estimate similar income processes in the United States and Sweden. The taste-shock for remote work σ_g is chosen so that 8.2 percent of the pre-pandemic laborforce works remotely, consistent with the estimates in Bick et al. (2020). Labor's share of income comes from Gollin (2002), and the rental rate of capital is set to the two-week return on pre-COVID Treasury Bills. We set the productivity penalty for remote work in the nonsocial

Table 3.3. Calibration of Epidemiological Parameters

Var	Description	Value	Source or Target
η	Effect of infection on productivity	0.3	Alene et al. (2021)
ξ	Reduction of infection probability by working from home	0.6	Mossong et al. (2008)
κ	Impact of hospital overuse on fatality	2	Glover et al. (2020)
$\pi_y^{\mathbb{C}}$	Rate of young entering $\mathbb{C}$ from $\mathbb{I}$	6.7%	Ferguson et al. (2020)
$\pi_o^{\mathbb{C}}$	Rate of old entering $\mathbb{C}$ from $\mathbb{I}$	38.0%	Ferguson et al. (2020)
$\pi_y^{\mathbb{D}}$	Rate of young entering $\mathbb{D}$ from $\mathbb{C}$	2.7%	Glynn (2020)
$\pi_o^{\mathbb{D}}$	Rate of old entering $\mathbb{D}$ from $\mathbb{C}$	9.0%	Glynn (2020)

sector, ϕ_n, to unity, consistent with evidence of small productivity losses for these workers in most cases, and potentially even productivity gains in some cases (Barrero et al., 2021). The penalty for remote work in the social sector, ϕ_s, is chosen to match the employment decline in the US. Ideally, this parameter would be chosen to match the observed average employment decline among all advanced economies of 2.4 percent; however, this decline is less than the average advanced economy output decline of 4.6 percent. Due to the assumption that production technology is Cobb-Douglas, the model is unable to generate a decline in output greater than the decline in employment. Hence we opt instead to use the US number, which is consistent with the production technology.[5] Finally, The TFP shock accompanying the pandemic $A(P)$ is chosen to match the aggregate 2019-2020 year-on-year output decline in advanced economies.[6]

Table 3.3 reports parameters controlling the epidemiological transmission of disease and their interactions with public health infrastructure and lockdown policies. We take parameters governing the fatality infection rates from Glynn (2020) and the rates of infected cases becoming critical from Ferguson et al. (2020). The reduction in infection probability is set to 0.6 to match evidence from Mossong et al. (2008) that about 40 percent of person-person contacts happen at the workplace. The effect of hospital congestion on disease fatality rates, κ, is taken from Glover

[5]The larger decline in advanced economy output than employment may be reflective of the fact that many advanced economies implemented policies that explicitly rewarded employers that did not furlough workers, leading to many workers who were technically employed but idle. Notably, the United States did not implement many policies of this nature.

[6]Appendix Table C.2 summarizes the internally calibrated parameters and the model's fit to the data. Note that TFP in normal times, $A(N)$ is set to one, so that $A(P)$ should be interpreted as a relative TFP shock in effect during the Pandemic.

et al. (2020). The productivity penalty of becoming infected, η, is set to match a 30 percent share of asymptomatic infection cases, as estimate in the meta-analysis of Alene et al. (2021). Such a choice is motivated by the observation that those known to be infected cannot work, and so have productivity of zero, while those who are infected but asymptomatic may continue to work unhindered. Finally, we choose the time-varying behavioral-adjusted infection probability, $\beta_t^{\mathbb{I}}$, so that the model's endogenous path of fatalities precisely matches the experience of the United States. The simulated endogenous path of the virus also account the time path of vaccinations and lockdowns in the U.S.. Vaccination data is taken from the COVID-19 Data Repository by CSSE at John Hopkins University, and we assume vaccination rates continue to grow at 1% per period after the last available data point, until period 60. The time path of lockdown policies comes from the Oxford Coronavirus Government Response Tracker (see Appendix Figure C.8). We assume lockdown policies are gradually lifted starting in the last period of available data until they are completely discontinued by period 60. Figure 3.8 plots the fitted results and validates the model's ability to replicate these dynamics exactly.

The mortality dynamics (and output losses) are shaped by both government lockdown policies and voluntary household substitution away from market work and consumption. Figure 3.9 illustrates that both margins play an important role in the calibrated model by plotting the equilibrium population share under lockdowns or voluntarily sheltering from home. The dashed purple line represents the strictness of prevailing lockdowns, reporting the share of the susceptible population forced to stay home. Any mass above this "Lockdown" curve represents voluntary substitution to working from home. The figure shows that there is a considerable amount of voluntary sheltering at home, above and beyond what is required by lockdowns, especially during times of peak infection (such as in Winter of 2020-2021). Voluntary substitution to working from home also varies substantially across age and sector of employment. Consistent with differences in health risks and economic costs, we see more voluntary working from home among the older population and in the social sector, where the health risks of the pandemic are most acute. The result also highlights the importance of accounting for heterogeneity in age and the sectoral

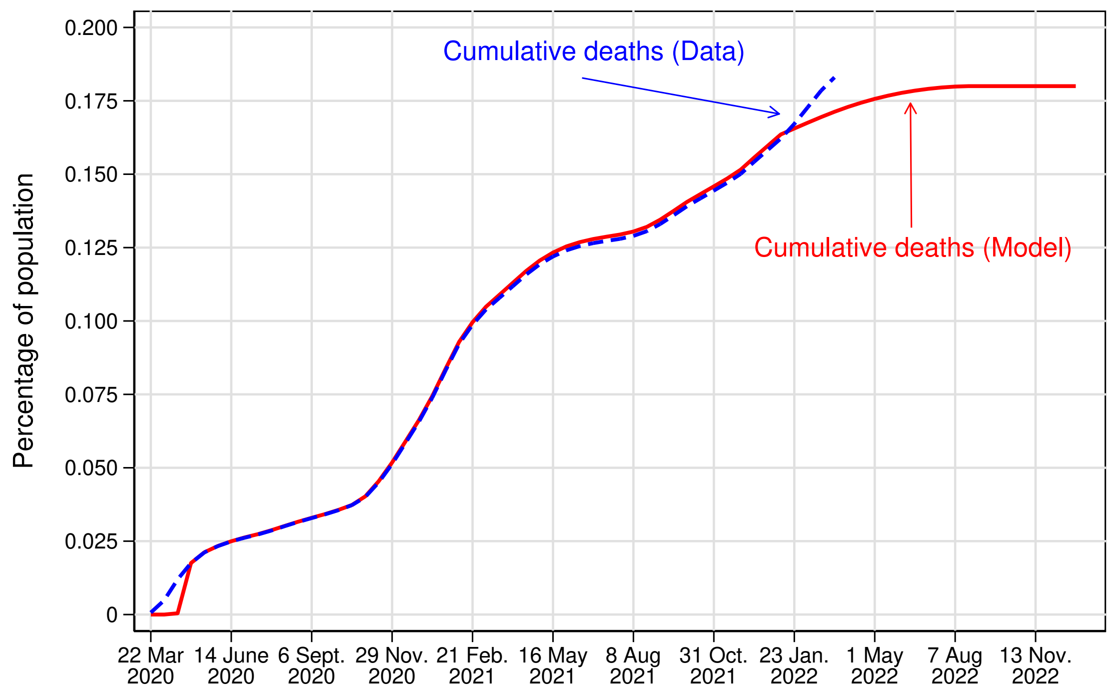


Figure 3.8. Predicted and Actual COVID-19 Mortality in the Advanced Economies

Note: Time path of U.S. COVID-19 mortality taken from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at John Hopkins University.

composition of work alongside government policies in evaluating output and mortality dynamics throughout the pandemic.

As a simple sanity check of the model's infection and mortality dynamics, we can compare the model-predicted proportion of COVID deaths occuring among those older than 65 to un-targeted data from the real world. According to the CDC, this proportion is 74 percent[7]. In the quantative model, it is fairly close at 82 percent, providing confidence that the model's predictions for the distribution of deaths across age groups are closely aligned with reality.

Table 3.4 summarizes parameters which vary across advanced and developing countries. Although this is a parsimonious set of parameters, we argue that it captures the key fundamental differences relevant for COVID-19 that vary across countries. In Appendix C.3, we explore the robustness of our quantitative results to variation in other parameters that one might intuitively

[7]Taken from https://data.cdc.gov/d/vsak-wrfu/visualization. Accessed May 26th, 2022.

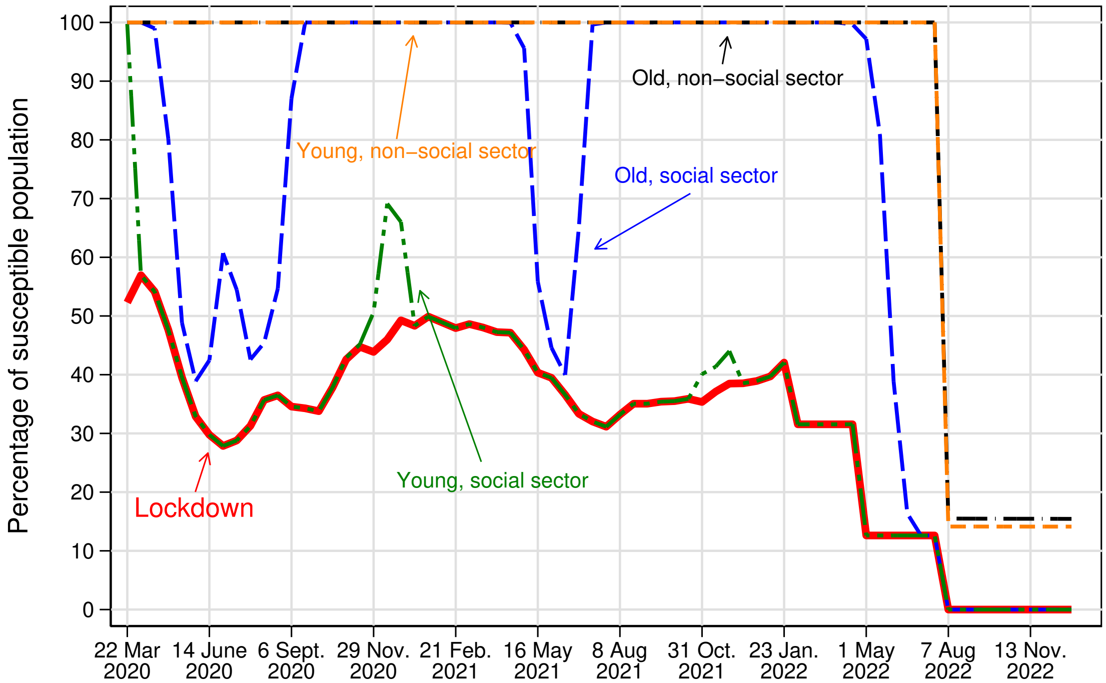


Figure 3.9. Lockdowns and Voluntary Working from Home

expect to vary between countries of different income levels.

The tax rates for the advanced and developing countries are taken from Besley and Persson (2013). Age demographics ω_y come from the World Bank and measure the share of the population under 65. The youth share in advanced economies corresponds to the U.S. economy, as it is our benchmark calibration, and we set the shares for emerging and low-income countries to their group averages. The share of workers in the social sector, ω_s, is constructed using estimates from Gottlieb et al. (2021b) on the share of urban labor that can work from home and adjusting the ratio to account for the rural population. Specifically, using data on urban and rural labor supply from the UN Population Division, we include all rural jobs in non-social employment and then assign urban workers to the non-social sector based on the proportion who can work from home. From this we calculate ω_s. For example, Gottlieb et al. (2021b) estimates that 7.4 percent of urban workers in Kenya can work from home, and 25.7 percent of the Kenyan

Table 3.4. Calibration of Parameters Varying Across Advanced and Developing Economies

Var	Description	Advanced Economies	Emerging Economies	Low-Income Economies	Source or Target
$\bar{u}$	Flow value of being alive	$11.4\bar{c}^{US}$	$11.4\bar{c}^{MID}$	$11.4\bar{c}^{DEV}$	Glover et al. (2020)
τ	Marginal tax rate	0.25	0.20	0.15	Besley and Persson (2013)
ω_y	Share of young in population	73%	84%	92%	UN Population Division
ω_s	Share of social sector workforce	40%	57%	27%	Gottlieb et al. (2021b)/IPUMS
Θ	Hospital capacity per capita	0.00042	0.00025	0.00011	Glover et al. (2020) / WHO

population resides in urban areas, resulting in a social sector share of $0.257 \times (1-0.074)$ or 23.8 percent. For advanced economies (which are absent from the dataset of Gottlieb et al., 2021b), we use the work-from-home estimates reported by Barrero, Bloom, and Davis (2021).

The flow value of life, $\bar{u}$, is calibrated using the value of statistical life (VSL) approach. Following Glover et al. (2020), we set the per-period statistical value of life to \$515,000 for advanced economies, equal to 11.4 times average US consumption. The value for $\bar{u}$ is then computed so that the behavioral response to a marginal increase in the risk of death is consistent with the VSL. Specifically, we get $\bar{u}$ by solving,

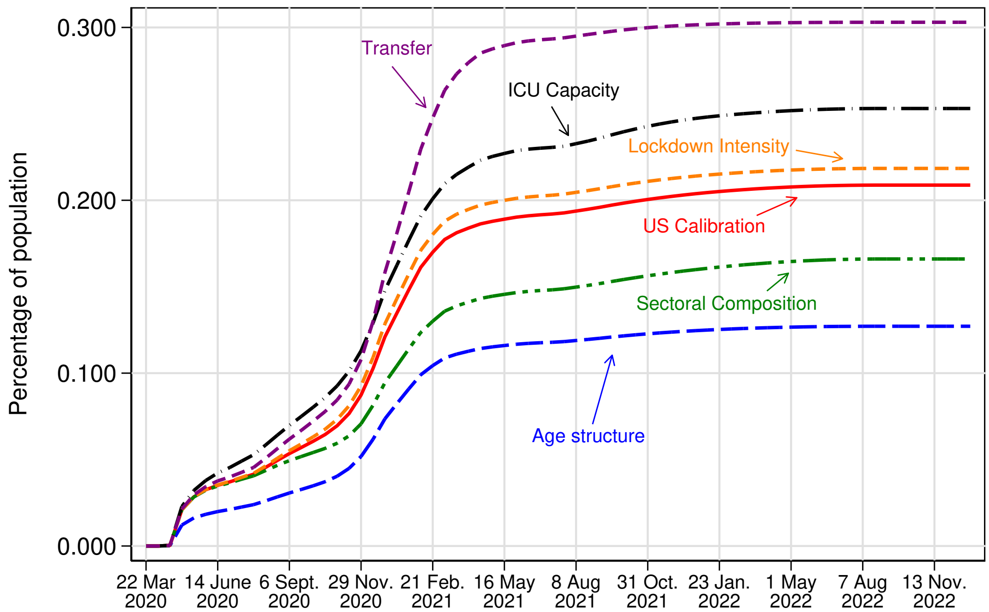
$$\text{VSL} = \frac{dc}{d\rho}|_{E(u)=k,\rho=0} = ln(\bar{c}) - \bar{u}$$

where ρ is the risk of death and $\bar{c}$ is average consumption. Absent better evidence, we assume the VSL has unitary income elasticity and adjust $\bar{u}$ for developing countries accordingly.

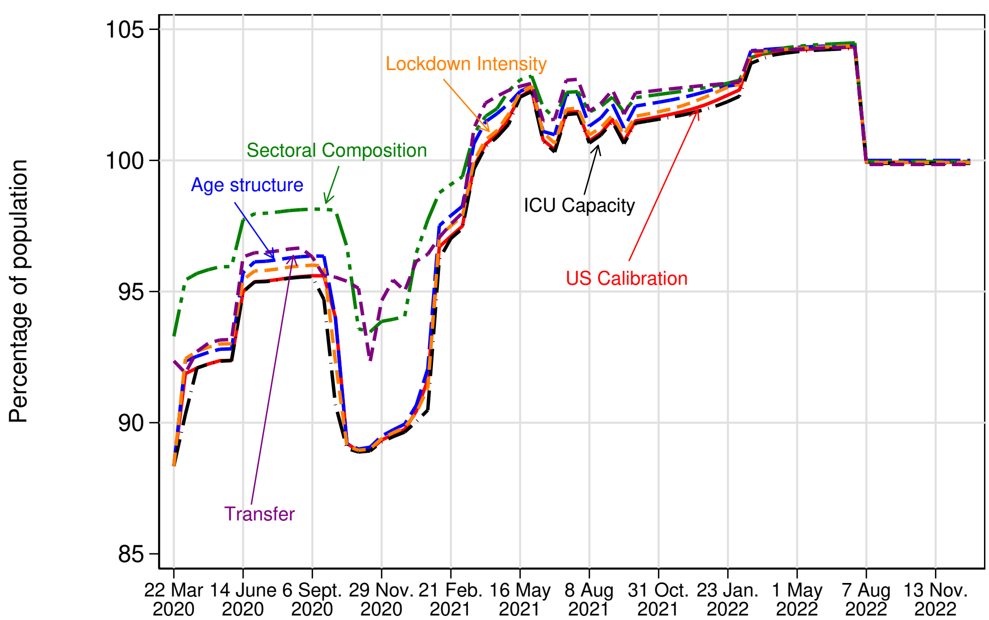
The final cross-country parameter to be set govern the ICU hospital capacity in developing and developed countries. One challenge is that while many countries report hospital bed capacity, few developing countries distinguish explicitly between general hospital capacity and ICU capacity in the data. To address this, we assume the ratio of hospital beds to ICU beds is constant across countries, and calibrate Θ by adjusting WHO data on the availability of hospital beds in the top and bottom quartiles of country income levels (as in Figure C.13) by the ratio of hospital beds to ICU beds taken from Glover et al. (2020).

3.4.2 Economic and Demographic Sources of Cross-Country Differences

Figures 3.10 and 3.11 plot the dynamic path of GDP per capita and fatalities as a percentage of the population during the COVID-19 pandemic in the United States in each of our counterfactual simulations. The panels on top display cumulative fatalities and those on the bottom plot GDP. Each figure provides six simulated paths: the benchmark U.S. calibration and the five counterfactual exercises which vary demographics, the sectoral composition of employment, public healthcare capacity, government transfer programs, and the stringency of lockdowns. Figure 3.10 reports counterfactuals that endow the U.S. economy with the characteristics of

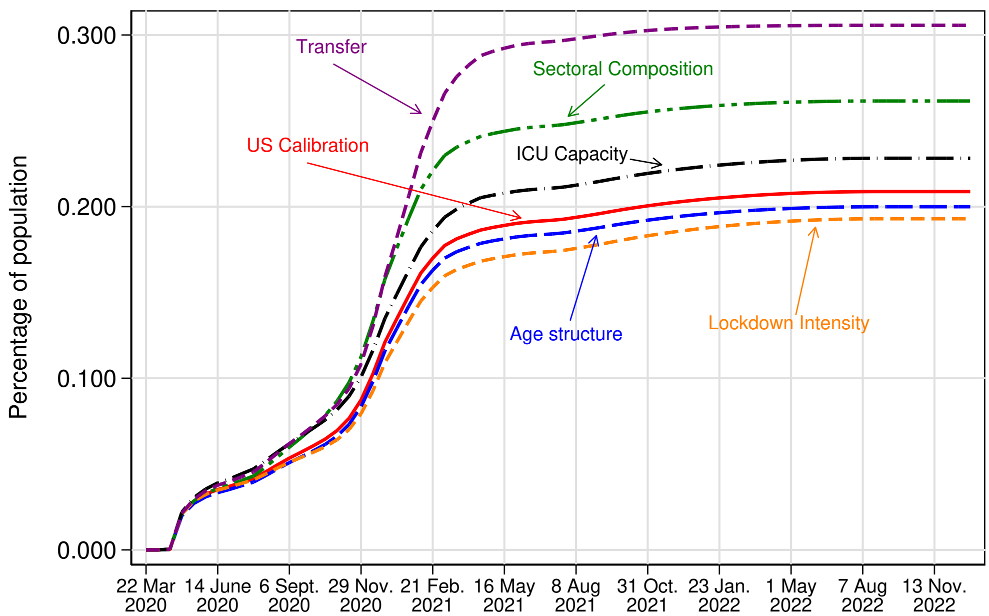


(a) Cumulative Death, US with Low Income Economies' Features

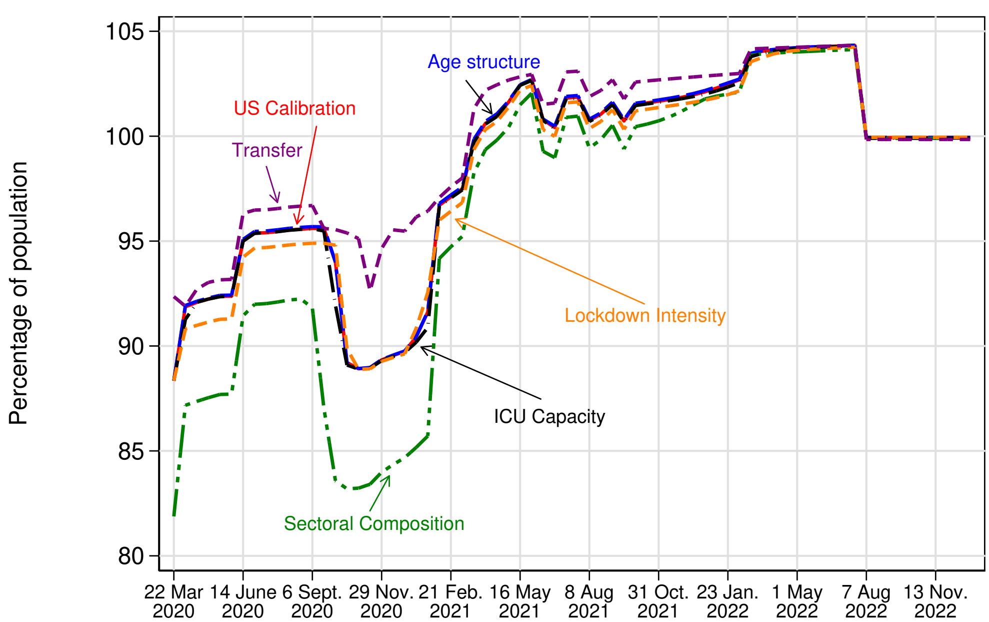


(b) GDP per capita, US with Low Income Economies' Features

Figure 3.10. Time Path of Cumulative Deaths and GDP: Low Income Economies



(a) Cumulative Death, US with Emerging Economies' Features



(b) GDP per capita, US with Emerging Economies' Features

Figure 3.11. Time Path of Cumulative Deaths and GDP: Emerging Economies

low-income countries; Figure 3.11 reports the results of endowing the U.S. with emerging market economy characteristics.

Looking across the panels, one can see that all five mechanisms play an important role to some degree, but differences in age demographics, the sectoral composition of employment, and the size of public transfer programs are the most quantitatively prominent. In both low income and emerging market economies, low levels of public financial assistance during the pandemic lead to much higher levels of fatalities. Without transfers to support or replace income lost in the pandemic, many households are not able to shelter at home during times of peak infection and instead must work outside the home, further propagating the spread of disease which increases fatalities. Quantitatively, low levels of public transfers are the largest factor pushing fatalities higher outside advanced economies. The simulations show that if the United States scaled back its public transfer programs to the levels of low income and emerging market economies, that alone would lead to cumulative fatalities from the pandemic to grow by 50 percent. Despite the effect of low transfers, fatalities in low income countries remained modest because of the offsetting effect of its substantially younger population. The high agricultural employment share in low-income countries also substantially reduces fatalities. In contrast, emerging markets experienced far higher mortality rates because they do not benefit from the favorable demographics of low income countries and have a high social sector employment share, making it difficult to control the spread of disease while working. ICU constraints exacerbated fatalities in both low and emerging economies, though played a secondary role overall. Differences in lockdown intensity play the smallest role, suggesting the more important cross-country policy difference during the pandemic was in the size of public insurance programs.

The output counterfactuals exhibit less variation than what we see in fatalities, suggesting the mechanisms we study contribute more equally to observed economic declines. Among the channels, only the sectoral composition of employment and public transfers stand out as having an especially important quantitative role. In low-income countries, economic losses were moderated by a large agricultural sector that was minimally disrupted by lockdowns and social distancing requirements. In emerging markets, high levels of urban employment in jobs that cannot be done from home explains a substantial part of their larger economic losses. Somewhat perversely, the low levels of public transfers which amplify fatalities in low income and emerging market economies also serve to *reduce* output losses by causing financially vulnerable households to continue working outside the home. The impact is most pronounced at times of peak infection, as is visible in the transition paths during the winter months of 2020.

To assess what may be driving the especially bad outcomes observed in emerging markets, Table 3.5 reports the cumulative effect of our counterfactuals on 2019-2020 year-on-year changes in GDP and fatalities. For comparison, the first data column displays the data for advanced and emerging economies discussed in the introductory sections (see Appendix Figures 3.1 and 3.2). The second data column reports the model-predicted simulated outcomes for both the advanced and emerging market economies. The third column reports outcomes when we endow the advanced economy with only the age demographics, sectoral employment, and ICU capacity of emerging economies. We distinguish these features since we view them as largely immutable throughout the pandemic's duration. To facilitate comparisons, the final row of each column reports the ratio of outcomes in emerging markets relative to advanced economies.

In panel (a) we see that the model does relatively well at replicating variation in GDP. In the data, GDP in advanced economies contracted by -4.6 percent while emerging economies shrank by -6.7 percent. The benchmark model nearly replicates these data, predicting GDP declines of -4.6 percent and -6.3 percent in advanced and emerging economies, respectively. While the model predicts contractions in GDP that are slightly smaller than in the data, it accurately replicates the relative severity of the pandemic across countries. In the model, emerging markets experience contractions in GDP that are 34 percent larger, in line with the data, at 47 percent larger.

Panel (b) reports excess mortality per hundred thousand people in advanced economies and emerging markets, both in the data and full counterfactual. As with GDP, the model accurately reproduces total mortality in both advanced and emerging economies. The model

Table 3.5. Cumulative Effect of the COVID-19 Pandemic in Emerging Markets

Panel (a): GDP Changes from 2019 to 2020			
	Data	Model	
		All Features	Age/Sector/ICU
Advanced Economies	-4.6	-4.7	-4.7
Emerging Economies	-6.7	-6.3	-8.3
Ratio	1.47	1.34	1.77
Panel (b): Excess Mortality			
	Data	Model	
		All Features	Age/Sector/ICU
Advanced Economies	100	105	105
Emerging Economies	165	159	95
Ratio	1.65	1.51	0.90

also does a good job at replicating the relative severity of the pandemic in emerging markets. Endowing the advanced economy with all the features of an emerging market economy leads to a 51 percent rise in excess mortality. In the data, emerging market economies registered excess mortality that was on average 65 percent greater than advanced economies. The model can therefore almost completely account for the relatively higher number of fatalities in emerging markets during the pandemic.

Finally, in light of the particularly severe outcomes in emerging economies, it is natural to ask if there is anything governments in emerging markets could have done differently to improve their outcomes. While we do not model the optimality of different policies, our framework allows us to study the extent to which differences in outcomes depend on features that are outside the control of policymakers. In particular, we view a country's age demographics, sectoral composition of employment, and healthcare capacity to be largely fixed over the duration of the pandemic. That governments cannot choose the age of their population is obvious. Similarly, it's generally widely held that the industrial composition of the economy is rigid in the short-run. While public healthcare capacity can in principle be expanded (and was, rather rapidly in a few places like China), we believe that emerging market economies by and large only had limited ability to do so during the pandemic, especially given the concurrent global competition for medical equipment, oxygen, and protective gear.

The final column of Table 3.5 reports the cumulative counterfactual impact on output and fatalities if only these immutable characteristics varied between emerging markets and advanced economies while differences which depend on policy choices, namely transfers and lockdowns, remain fixed at advanced economy levels.

For output, the immutable characteristics alone predict a -8.3 percent decline in GDP for emerging markets, over-accounting for the -6.7 percent decline observed in the data. This simulation suggests that the especially large GDP declines in emerging markets may have been largely outside the control of policymakers, depending instead on prevailing demographic and structural conditions that cannot be easily changed. The larger GDP losses in emerging economies under advanced economy transfer and lockdown policies signifies that the number of workers induced to voluntarily stay home through more generous transfer programs in advanced economies more than outweighed the countervailing effect of their less stringent lockdowns.

For mortality, these fixed features lead to a 10 percent decline in mortality, a modest share of the 65 percent increase in mortality observed in the data. In the model, much of the mortality gap between advanced and emerging economies is accounted for by the smaller transfers implemented by emerging market economies (we discuss this in more detail in subsection 3.4.3). Limited public financial assistance results in many lower-income households continuing to work outside the home during the pandemic, propagating infections that lead to higher mortality. The result suggests that constraints on the ability of emerging markets to support large scale public transfer programs, such as limited fiscal capacity and borrowing costs, were an important determinant of the greater excess mortality they experienced during the pandemic.

Table 3.6 reports the cumulative counterfactual effects of the COVID-19 pandemic in low income countries. The model correctly predicts the more modest GDP declines and mortality rates in low income economies relative to advanced ones, as in the data, albeit with larger

Table 3.6. Cumulative Effect of the COVID-19 Pandemic in Low Income Economies

Panel (a): GDP Changes from 2019 to 2020			
	Data	Model	
		All Features	Age/Sector/ICU
Advanced Economies	-4.6	-4.7	-4.7
Low Income Economies	-3.6	-2.3	-2.8
Ratio	0.78	0.48	0.60
Panel (b): Excess Mortality			
	Data	Model	
		All Features	Age/Sector/ICU
Advanced Economies	100	105	105
Low Income Economies	-	43	38
Ratio	-	0.41	0.36

differences. In the counterfactual, low income countries experience GDP declines that are 48 percent the size of contractions in advanced economies, while the data show declines that were 78 percent the size of advanced economy losses. The result suggests that there may be other important factors driving GDP declines in low income countries that are not present in our current model, such as foreign demand shocks, supply chain disruptions, constraints on sovereign debt, and limited fiscal capacity. The model also predicts mortality rates in low income countries that are 41 percent the size of mortality rates in advanced economies, primarily due to their younger age demographic and higher agriculture employment share. While systematic data on COVID-19 fatalities in low income countries is not yet available, the result is consistent with the limited evidence on excess deaths available for some countries in Africa (see Appendix Figures C.4 and C.5).

As with emerging markets, we can assess the extent to which outcomes were driven by policy choices or fixed short-run characteristics of low income countries by comparing the cumulative counterfactuals in the last two columns of Table 3.6. Endowing advanced economies with only the immutable characteristics of low income countries generates a GDP declines of

-2.8 percent compared to a decline of -2.3 percent in the full low income economy. For mortality, these fixed factors alone lead fatalities to fall by 64 percent, accounting for slightly more than the entire model-predicted decline in mortality between advanced and low income economies of 59 percent. Taken together, the results suggest that younger populations and high agricultural employment shares predisposed low-income countries to have fewer fatalities from COVID-19, but public lockdowns and transfer policies played an important role in moderating the economic fallout accompanying the pandemic.

3.4.3 Key Takeaways

Table 3.7. Cumulative Effect of the COVID-19 Pandemic: Counterfactual Scenarios

Panel (a): Emerging Economies		
	GDP Changes	Excess Mortality
Emerging Economies	-6.3	159
with Advanced Economies Transfer	-8.5	88
with Advanced Economies Lockdown	-5.6	186
Panel (b): Low Income Economies		
	GDP Changes	Excess Mortality
Low Income Economies	-2.3	43
with Advanced Economies Transfer	-2.7	39
with Advanced Economies Lockdown	-2.5	41

Despite its simplicity, the model is able to accurately replicate the observed differences in macroeconomic outcomes following COVID-19 across the world income distribution. In this subsection, we summarize the key takeaways from the various outcomes and counterfactuals produced by the model.

The first takeaway is that the large output declines in emerging market economies were largely a function of pre-existing characteristics rather than cross-country differences in policy. The second column of Panel A in Table 3.5 reports a model-predicted GDP decline of 6.9 percent for the typical emerging market economy following typical lockdown and transfer policies. For comparison, the third column displays the decline in GDP for a typical emerging market economy (i.e. an economy with typical demographics, sectoral composition, and healthcare capacity) implementing the lockdown and transfer policies of the advanced economies. Mimicking the lockdown and transfer policies of the advanced economy leads the emerging market economy to experience a 1.4 percentage point higher decline in GDP, suggesting that there was little emerging market policymakers could have done differently to avoid high output losses.

Table 3.7 further investigates the extent to which policy could have mitigated outcomes in emerging market economies by displaying the model-predicted GDP losses and mortality for the emerging market and low income economies if they implemented the transfer or lockdown policies of advanced economies. Looking at Panel (a), it is clear that large GDP losses in emerging markets were unavoidable. Implementing the advanced economy's less restrictive lockdown policies reduces GDP losses by only 0.7 percent and implementing larger transfers actually increases GDP losses by encouraging more workers to stay home.

In contrast to GDP losses, the model predicts that mortality in emerging market economies could have been substantially reduced through policy. While less stringent lockdowns, unsurprisingly, result in a slight increase in mortality, increasing the size of transfers to the level seen in the advanced economy substantially reduces mortality from 159 people per 100,000 to 88. This substantial reduction results from the fact that larger transfers make up for lost incomes and allow the most vulnerable social sector workers, who make up a substantial share of the population in emerging market economies, to avoid going to the workplace.

While the impact of policy differences in low income countries move outcomes in similar directions, the overall quantitative contribution of these changes is small. The stringency of lockdowns in low-income and advanced economies were relatively similar, and so account for only a small part of the cross-country variation. Differences in transfer programs were much more substantial, but still only have modest effects on output losses and mortality. These outcomes are consistent with the notion that the relatively better macroeconomic outcomes experienced by low income countries during the pandemic were largely the result of pre-existing favorable demographics and high shares of non-social sector employment, rather than particular policy choices.

Finally, while we don't model optimal policy decisions, it seems reasonable to suggest that the lower transfers in emerging market economies were not an intentional choice by local policymakers but rather were an outcome of the limited fiscal capacity and high borrowing costs faced by many governments. Interpretted through this lens, the model's policy counterfactuals suggest that the ability of governments in emerging markets to borrow and fund emergency transfers was an important determinant of their overall mortality outcomes. Likewise, this result suggests that expansions in international emergency financial assistance during the pandemic, including debt relief and lending programs such as the IMF's Rapid Credit Facility (RCF) and Rapid Financing Instrument (RFI), contributed meaningfully to reducing mortality globally and particularly so in emerging market and lower income countries.

3.5 Empirical Correlates of GDP Declines During the Pandemic

In this section we explore the empirical correlates of changes in GDP per capita from 2019 to 2020, focusing on the same variables emphasized in the model. We make no claim at uncovering causal patterns in this section. Instead, we assess the extent to which correlations between aggregate income changes during the pandemic and a country's demographic, economic, and policy characteristics are consistent with the model's predictions and quantitative exercises.

We begin with the basic relationship between declines in GDP per capita and pre-pandemic level of GDP per-capita. The first column of Table 3.8 shows that this relationship is U-shaped, as we argued earlier. Both the level and quadratic coefficients on GDP per capita in 2019 are statistically significant at the five-percent level, with the former negative and latter positive. The second column includes controls for the agricultural employment share. The variable exhibits a significant positive correlation with changes in GDP, holding constant differences in national

Table 3.8. Correlates of GDP per Capita Change from 2019 to 2020

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

	Dependent variable: GDP per capita change from 2019 to 2020					
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)
GDP per capita in 2019	-0.10**	0.037	-0.17*	-0.076*	-0.11	-0.052
	(0.046)	(0.068)	(0.094)	(0.044)	(0.068)	(0.11)
GDP per capita in 2019^2	0.0014**	0.00021	0.0020*	0.0011*	0.0014*	0.00084
	(0.00066)	(0.00071)	(0.0010)	(0.00063)	(0.00080)	(0.0011)
Agriculture emp. share		0.076***				0.062**
		(0.027)				(0.030)
Median age			0.083			0.074
			(0.079)			(0.082)
Lockdown stringency				-0.13***		-0.13**
				(0.043)		(0.053)
Economic support					0.0042	0.024
					(0.036)	(0.038)
Constant	-4.21***	-8.03***	-5.67***	2.38	-4.29***	-2.97
	(0.60)	(1.66)	(1.48)	(2.07)	(1.09)	(3.34)
Observations	144	144	144	140	140	140
R^2	0.031	0.071	0.037	0.129	0.030	0.163

income, means that countries with larger percentages of their workforce in agriculture also experienced smaller declines in national income, all else equal. Interesting, the coefficients on GDP per capita and its square are now statistically indistinguishable from zero, with the former switching signs. The third column includes median age as a control which exhibits no significant correlation, somewhat puzzlingly. The fourth column controls for the stringency of lockdowns, which is positive and statistically significant. The fifth column adds controls for the generosity of economic support programs during the pandemic, which turns out to be statistically insignificant.

Column six of Table 3.8 adds all the covariates at once. This specification shows that agriculture's employment share remains a strong positive correlate of GDP changes, while lockdown stringency remains a strong negative correlate. Median age and the economic support index continue to be insignificant. Collectively, the inclusion of these controls eliminates the statistical significance of the original U-shape pattern in GDP per capita, and substantially reduce the magnitude of the original correlations. We take this as suggestive evidence that these variables are important empirical determinants of macroeconomic performance across the world income distribution, at least thus far, during the pandemic.

3.6 Conclusion

The macroeconomic impact of the COVID-19 pandemic was most severe in emerging market economies, which represent the middle of the world income distribution. This paper provides a quantitative economic theory to explain why these economies fared so poorly compared to both poorer and wealthier nations. Our model is motivated by key economic and demographic differences across the world income distribution, including variation in lockdown policies, public insurance, demographics, healthcare capacity, and the sectoral composition of employment.

Our quantitative model does well in predicting the greater output declines and higher mortality rates in emerging market economies, as in the data. The model also quantitatively predicts the more modest output losses and fatalities in low income countries, albeit to a greater extent than what is observed in the data. Among the factors we consider, the size of public transfer programs, age demographics, and the sectoral composition of employment are the most quantitatively important. Low levels of public financial assistance and a high share of jobs which require social interaction explains most of the greater GDP losses and higher fatalities in emerging markets. Low income countries also suffered from low levels of public transfers, but the negative consequences were largely blunted by their substantially younger populations, whom are naturally more resistant to illness, and large agricultural employment share, which provide a resilient source of income during lockdowns and while socially distancing. Quantitatively, the results suggest that cross-country differences are mostly driven by variation in public transfer programs and factors outside the short-term control of government officials. The outsized role of public transfer programs in explaining cross-country differences highlights the importance of constraints which may limit the ability of governments to enact and sustain large scale social insurance programs during emergencies. A valuable avenue for future research is to better identify the sources of these policy differences and what impact they had on macroeconomic outcomes throughout the pandemic.

Overall, our findings suggest that much of the variation in aggregate outcomes across country income groups during the pandemic can be attributed to a small set of economic characteristics and broad policy choices. However, the model is stylized in many ways and does not attempt to analyze the many more granular policy choices that surely mattered for the first year of the pandemic. Absent from this study are policy decisions regarding school closings (e.g. Fuchs-Schündeln, Krueger, Ludwig, and Popova, 2022), mask use (e.g. Abaluck et al., 2021; Karaivanov et al., 2021), testing and tracing policies (e.g. Berger et al., 2020), and vaccine provision (e.g. Arellano, Bai, and Mihalache, 2021). Future research could also fruitfully assess the quantitative importance of other policy choices for cross-country macroeconomic performance during the pandemic.

One salient cross-country feature missing from our analysis is informality; low-income and emerging market economies have much higher fractions of the population engaged in informal work than advanced economies do. Informality may be important for two reasons. First, informal workers in the social sector may be able to avoid lockdown policies due to being outside of the reach of government enforcement, reducing the effectiveness of lockdowns. Second, informal workers may be more difficult to reach with government transfer programs due to not being registered with social security. However, data from Latin American suggests that these concerns may be quantitatively small. Acevedo, Castellani, Lotti, and Székely (2021) document that both the formal and informal sector experience similar losses in jobs, suggesting that informal work did not serve as as a means through which workers shirked lockdowns, and Busso, Camacho, Messina, and Montenegro (2021) find that governments created many new transfer programs that increased the coverage of social insurance, particularly for those in the lowest income quintile, and helped cover informal workers.

To the extent that informal workers are less likely or unable to receive government transfers, our model will underpredict fatalies in emerging market economies relative to low income and advanced economies. While low income countries have the highest proportion of informal workers, the vast majority of these workers are engaged in non-social work due to living in rural agricultural areas and thus do not need transfers as they are able to maintain their income. Emerging market economies possess a much higher fraction of informal workers employed in the social sector. The inability to receive transfers means that these workers will be much less likely to choose to reduce their income and protect themselves from the virus, leading to higher contagion and higher spread. Alon et al. (2020) perform a thorough analysis of the effect of informality on COVID-19 outcomes for countries of varying income levels.

Another key limitation of our analysis is that it relies on a large exogenous time-varying component of the infection rate in order to match the observed path of excess deaths in the United States. In reality, however, much of the time variation in infection probabilities is likely due public policy choices that are not modeled here. These include policies that increase the prevalence of mask wearing, the development of better treatments for the infected, the rate of vaccination, or general knowledge about how COVID-19 can and cannot be transmitted. Future research should more explicitly consider the role these factors play in determining cross-country differences in the pandemic's consequences.

3.7 Acknowledgements

This chapter, in full, is a reprint of the paper ”Macroeconomic Effects of Covid-19 Across the World Income Distribution”, IMF Economic Review (2023), 71(1), 99-147. Kim, Minki; Alon, Titan; Lagakos, David; VanVuren, Mitchell. The dissertation author was the primary investigator and author of this paper.

Appendix A

Appendix for Chapter 1

A.1 Appendix Figures for Chapter 1

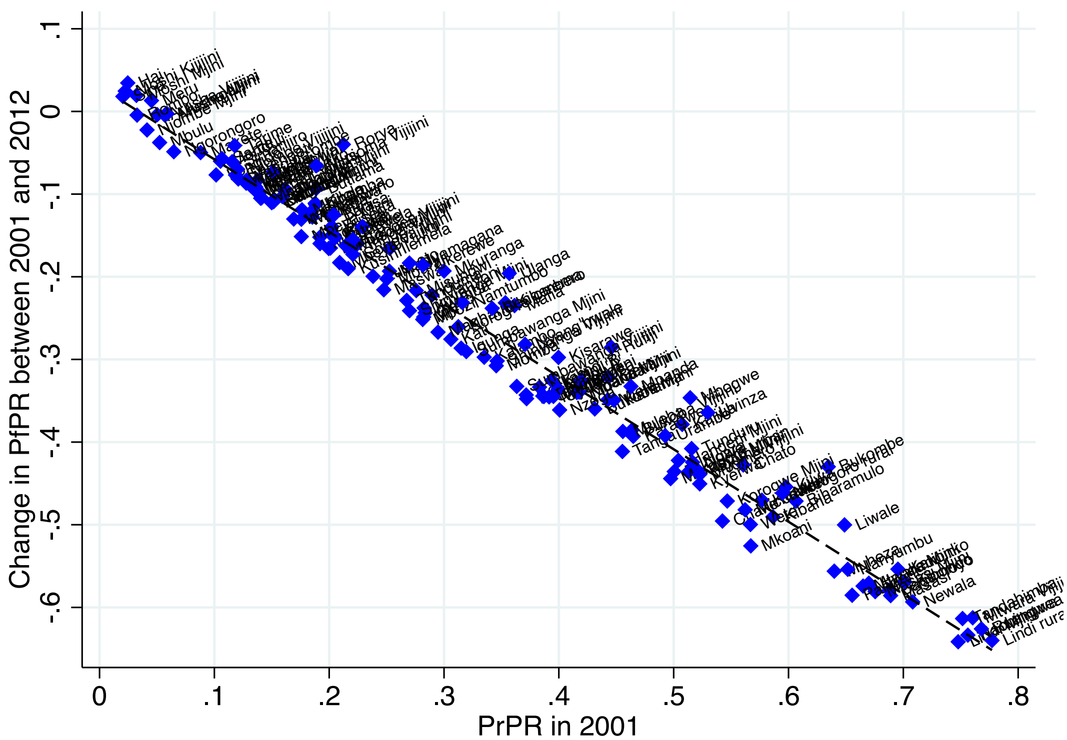


Figure A.1. Change in regional malaria risk conditional on malaria risk in 2001

Notes: Each point represents a district at GIS-2 level. Yearly malaria risk (PfPR) is obtained from the Malaria Atlas Project (MAP). Fitted line is calculated by regressing the changes in malaria risk over 2001 and 2012 on initial malaria risk in 2001.

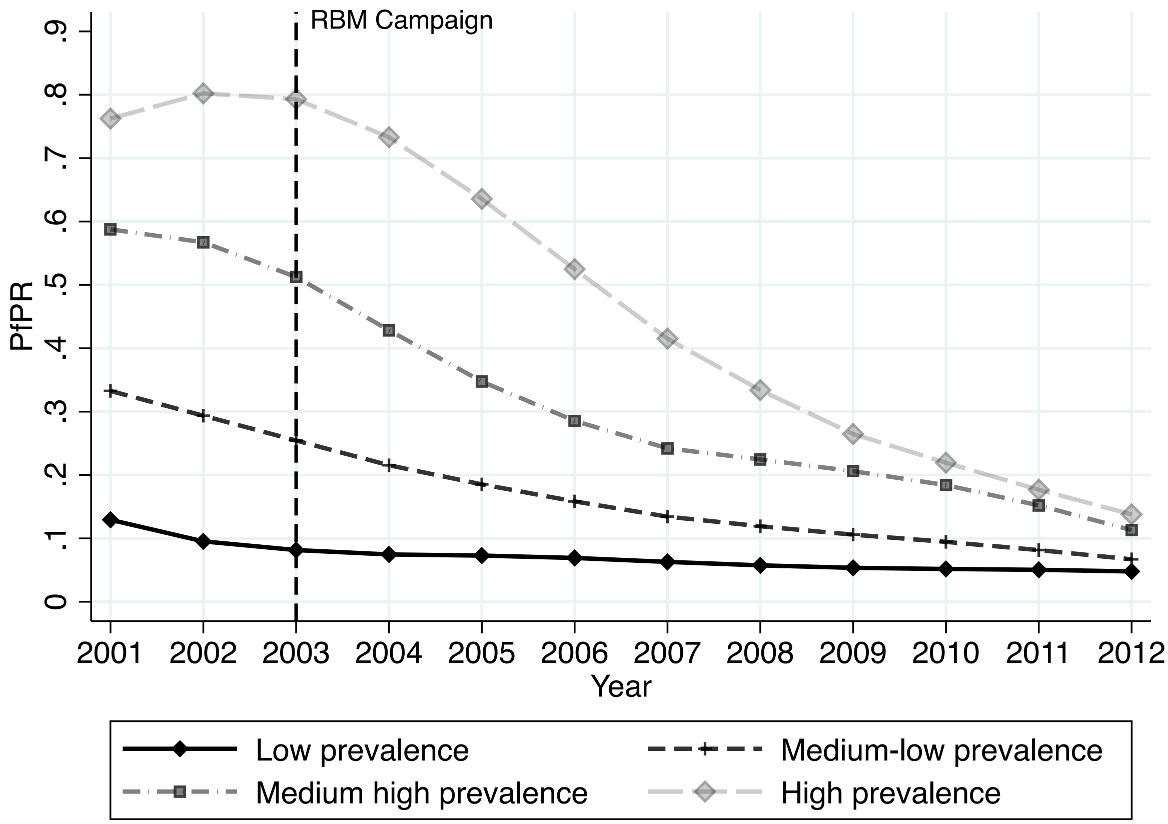


Figure A.2. Time trend of malaria risk of the four pre-campaign malaria prevalence categories

Notes: Each point represents a within-category population-weighted mean of PfPR. Regional malaria prevalence data obtained from the Malatia Atlas Project (MAP).

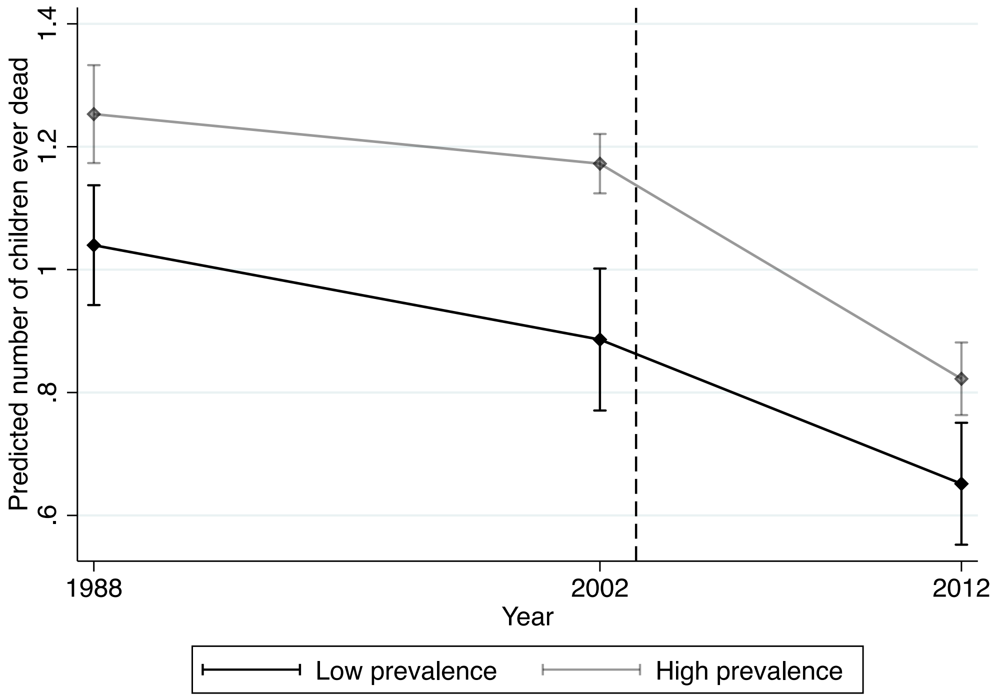


Figure A.3. Parallel trend in child mortality

Notes: This figure illustrates the parallel trend between the high- and low- prevalence regions by plotting the mean of the number of children ever dead conditional on the covariates used in regression (1.5). Three waves of the Tanzania National Census (1988, 2002, 2012) are used. Samples are restricted to women between the ages 30 and 49 in 2012 and those who were born and residing (surveyed) in the same region in 2012. Control variables included are the respondents' age and years of schooling and urban-rural residential status. Standard errors are clustered at the region level. 95% confidence intervals are plotted. The vertical dashed line indicates the timing of the RBM campaign.

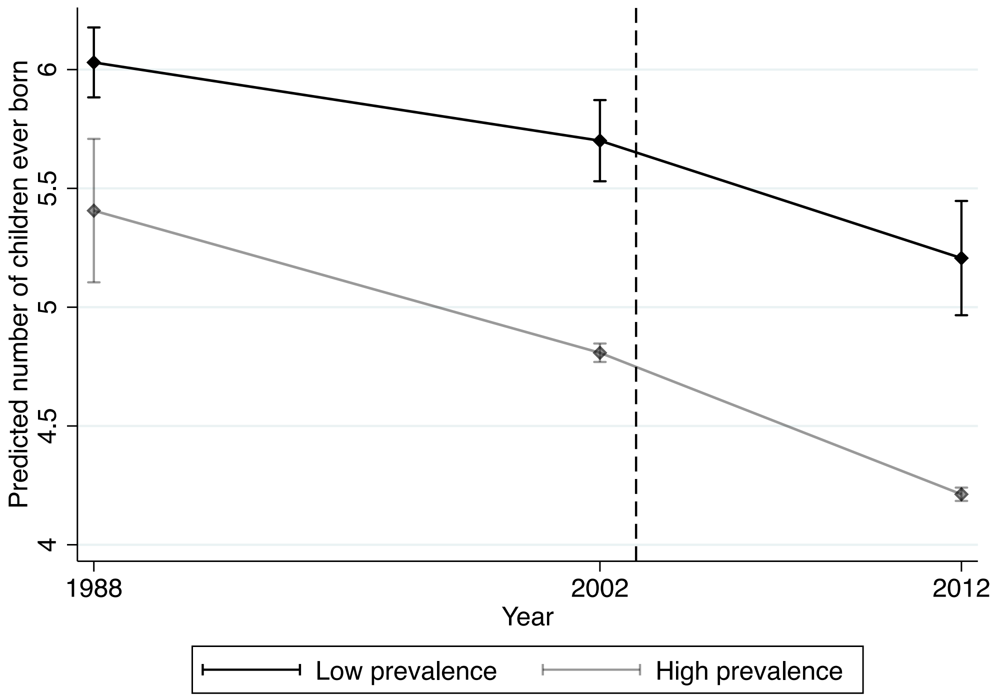


Figure A.4. Parallel trend in fertility

Notes: This figure illustrates the parallel trend between the high- and low- prevalence regions by plotting the mean of the number of children ever born conditional on the covariates used in regression (1.6). Three waves of the Tanzania National Census (1988, 2002, 2012) are used. Samples are restricted to women between the ages 30 and 49 in 2012 and those who were born and residing (surveyed) in the same region in 2012. Control variables included are the respondents' age and years of schooling and urban-rural residential status. Standard errors are clustered at the region level. 95% confidence intervals are plotted. The vertical dashed line indicates the timing of the RBM campaign.

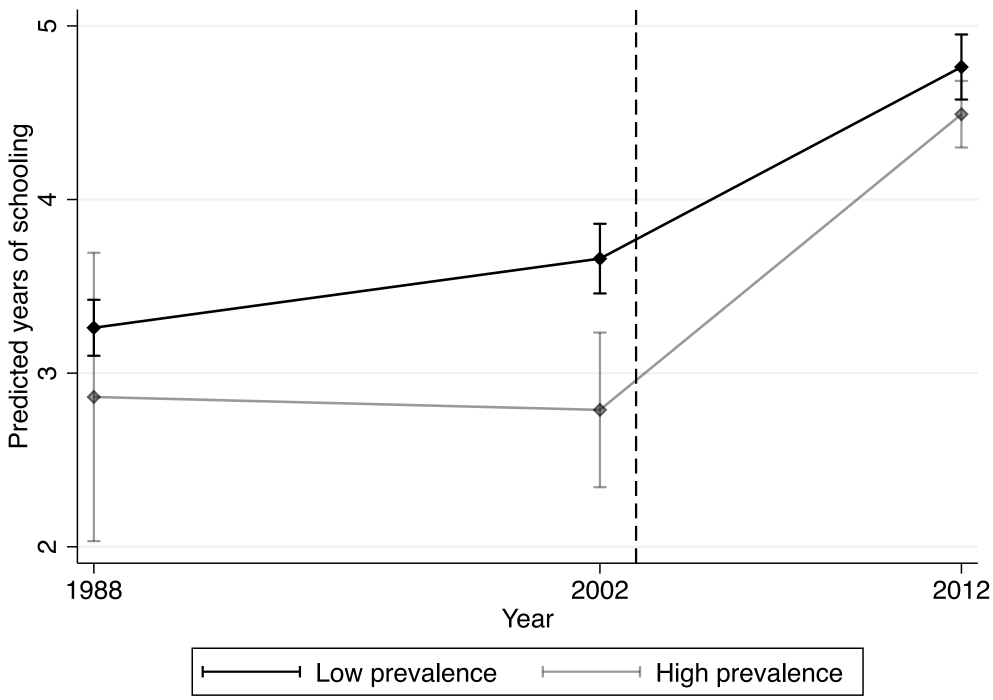


Figure A.5. Parallel trend in children's years of schooling, children aged 10-15 in 2012

Notes: These figures illustrate the parallel trend between the high- and low- prevalence regions by plotting the mean of the number of years of schooling conditional on the covariates used in regression (1.7). Three waves of the Tanzania National Census (1988, 2002, 2012) are used. Samples are restricted to women between the ages 30 and 49 in 2012 and those who were born and residing (surveyed) in the same region in 2012. Control variables included are the respondents' age and urban-rural residential status. Standard errors are clustered at the region level. 95% confidence intervals are plotted. The vertical dashed line indicates the timing of the RBM campaign.

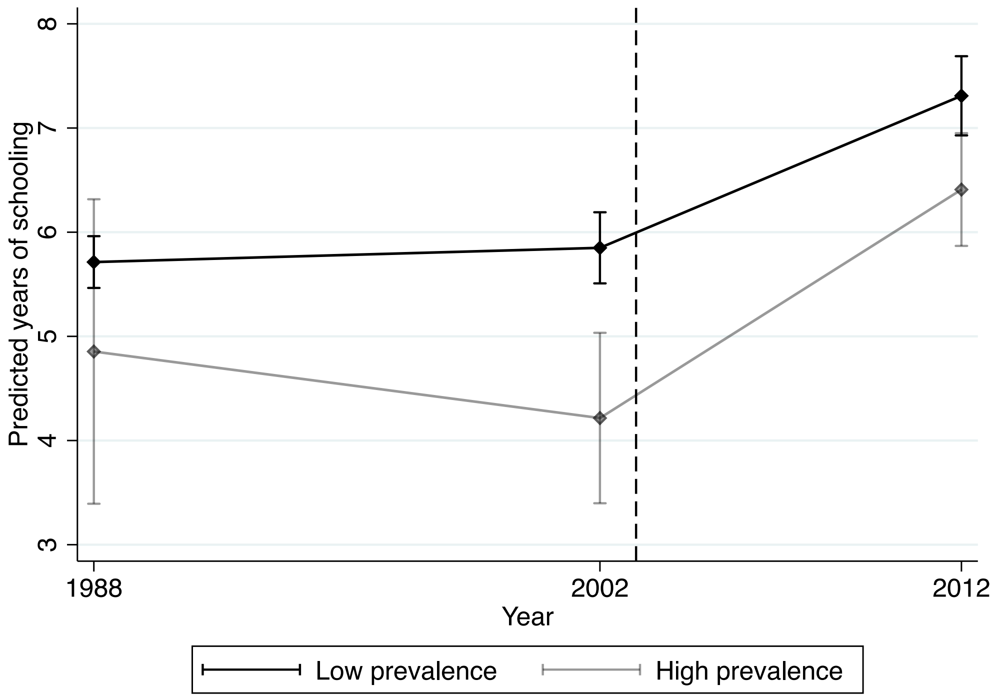


Figure A.6. Parallel trend in children's years of schooling, children aged 15-20 in 2012

Notes: These figures illustrate the parallel trend between the high- and low- prevalence regions by plotting the mean of the number of years of schooling conditional on the covariates used in regression (1.7). Three waves of the Tanzania National Census (1988, 2002, 2012) are used. Samples are restricted to women between the ages 30 and 49 in 2012 and those who were born and residing (surveyed) in the same region in 2012. Control variables included are the respondents' age and urban-rural residential status. Standard errors are clustered at the region level. 95% confidence intervals are plotted. The vertical dashed line indicates the timing of the RBM campaign.

A.2 Appendix Tables for Chapter 1

Table A.1. Descriptive Statistics

Notes: Calculated from 2002 Census data. Sample is restricted to women between age 35 and 49. Low-prevalence corresponds to the regions with PfPR lower than 10% in 2001, while high-prevalence corresponds to the regions with PfPR higher than 75% in 2001. Having water supply is defined as having access to piped water either within or outside the dwelling, including the public piped water. Mean coefficients; standard deviation in parentheses.

	Entire Sample	**Low Prevalence Regions**	**High Prevalence Regions**
Age	37.48	37.55	37.81
	(5.624)	(5.621)	(5.666)
Number of children ever born	5.751	5.622	4.855
	(3.018)	(2.887)	(2.697)
Number of children dead	0.985	0.856	1.204
	(1.386)	(1.316)	(1.550)
Years of schooling	4.407	4.672	3.973
	(3.484)	(3.420)	(3.362)
% Household w/ electricity	0.0695	0.0791	0.0288
	(0.254)	(0.270)	(0.167)
% Household w/ water supply	0.347	0.421	0.185
	(0.476)	(0.494)	(0.388)
PfPR in 2001	0.336	0.135	0.763
	(0.194)	(0.0510)	(0.00664)
Number of families in household	1.365	1.323	1.492
	(0.936)	(0.866)	(1.041)
Labor force participation	0.832	0.855	0.880
	(0.374)	(0.352)	(0.325)
Urban-rural status	0.355	0.340	0.400
	(0.479)	(0.474)	(0.490)
Observations	244,343	76,836	6,209

Table A.2. Effects of the RBM on Years of Schooling (Full Table)

Notes: This table reports the estimation restuls from OLS regression (1.7). Brackets contain standard errors clustered at the region level. PfPR_{2-10} (75%+) × Post indicates the interaction between the indicator of high-prevalence regions (PfPR in 2001 exceeding 0.75) and the post-treatment indicator. Other interaction terms are defined similarly. Samples are restricted to the individuals who were born and residing (surveyed) in the same region in 2012. Variable Urban indicates whether the respondent reside in the urban part within the region. All columns include region fixed effects. *, **, and *** indicate significance at the 10, 5, 1% levels.

	Age group in 2012			
	Age 10-15	**Age 15-20**	**Age 20-25**	**Age 25-30**
Dependent variable mean in 2012	4.28	6.85	6.68	5.78
Post	1.144***	1.431***	1.040***	0.358***
	(0.041)	(0.068)	(0.085)	(0.079)
PfPR_{2-10} $(20\% - 50\%) \times$ Post	0.006	0.032	0.012	-0.130
	(0.059)	(0.088)	(0.109)	(0.099)
PfPR_{2-10} $(50\% - 75\%) \times$ Post	-0.019	0.140	-0.027	-0.290**
	(0.093)	(0.122)	(0.137)	(0.140)
PfPR_{2-10} $(75\%+) \times$ Post	0.633***	0.974***	0.473***	-0.172
	(0.098)	(0.122)	(0.096)	(0.181)
Age	0.711***	0.130***	-0.081***	-0.075***
	(0.008)	(0.008)	(0.004)	(0.005)
Urban	0.759***	1.383***	1.566***	1.433***
	(0.040)	(0.068)	(0.071)	(0.067)
Observations	1,096,274	856,753	674,743	607,976

Table A.3. Effects of the RBM on Fertility (Full Table)

Notes: This table reports the estimation results for the Poisson regression (1.5) and (1.6). PfPR_{2-10} (75%+) × Post indicates the interaction between the indicator of high-prevalence regions (PfPR in 2001 exceeding 0.75) and the post-treatment indicator. Other interaction terms are defined similarly. Samples are restricted to women between age 30 and 49 in 2012, and those who were born and residing (surveyed) in the same region in 2012. Control variables included are age and years of schooling of the respondents and urban-rural residential status. Variable Urban indicates whether the respondent reside in the urban part within the region. All columns include region fixed effects. *, **, and *** indicate significance at the 10, 5, 1% levels.

Dependent variable	Gross Fertility Children ever born	Mortality Children ever dead	Net Fertility Surviving children
Dependent variable mean in 2012	5.30	0.71	4.57
Post	–0.104***	–0.316***	–0.073***
	(0.008)	(0.015)	(0.009)
PfPR_{2-10} (20% − 50%) × Post	0.012	0.019	0.016
	(0.010)	(0.021)	(0.011)
PfPR_{2-10} (50% − 75%) × Post	0.009	0.012	0.017
	(0.013)	(0.022)	(0.013)
PfPR_{2-10} (>75%) × Post	–0.0596***	–0.107***	–0.0255
	(0.008)	(0.035)	(0.017)
Age	0.026***	0.044***	0.022***
	(0.000)	(0.001)	(0.000)
Years of schooling	–0.016***	–0.057***	–0.008***
	(0.001)	(0.002)	(0.001)
Urban	–0.141***	–0.234***	–0.126***
	(0.005)	(0.015)	(0.006)
Observations	586,836	586,836	586,836

Table A.4. Malaria in Tanzania Among Children under age 10

Note: From Tanzania Household Panel Survey, wave 1 (2008) – wave 4 (2014). The survey questions were ”What is the 1st type of illness or injury did [NAME] had that led to his/her hospitalization?” for hospitalization, and ”What was the illness that caused [NAME]’s death?” for the death. Responses from the parents who were unsure of the cause of deaths are excluded.

Panel A: Top 5 illnesses that led to hospitalization (%)					
	Wave 1	Wave 2	Wave 3	Wave 4	Average
Malaria	**-**	**41.21**	**49.1**	**39.62**	**43.27**
Fever	-	21.21	15.77	21.92	19.68
Stomach	-	7.58	3.58	4.23	5.29
Diarrhea	-	5.45	6.45	1.92	4.72
Headache	-	0.91	0	0.38	0.46
Panel B: Top 5 illnesses that caused death (%)					
Malaria	**55.56**	**42.39**	**60.62**	**46.81**	**51.35**
Diarrhea	7.78	15.22	0.00	4.26	7.69
Vomiting	0.00	1.63	0.00	0.00	0.62
Flu	0.00	0.54	0.62	0.00	0.42
Asthma	2.22	1.09	0.62	0.00	1.04

A.3 Additional Empirical Results for Chapter 1

Table A.5. Child Quantity Regression for Different Age Groups

Notes: This table reports the estimation results from Poisson regression (1.5) and (1.6) from different age groups for women. PfPR_{2-10} (75%+) × Post indicates the interaction between the indicator of high-prevalence regions (PfPR in 2001 exceeding 0.75) and the post-treatment indicator. Other interaction terms are defined similarly. Samples are restricted to women who were born and residing (surveyed) in the same region in 2012. Control variables included are age and years of schooling of the respondents and urban-rural residential status. Variable Urban indicates whether the respondent reside in the urban part within the region. All columns include region fixed effects. *, **, and *** indicate significance at the 10, 5, 1% levels.

	Age group of women in 2012			
	Age 30-39	Age 40-49	Age 50-59	Age 60-69
Panel A: Child Mortality				
Dependent variable: Number of children ever died				
Post	-0.399^{***}	-0.234^{***}	-0.233^{***}	-0.180^{***}
	(0.019)	(0.023)	(0.023)	(0.020)
PfPR_{2-10} (75%+) × Post	-0.160^{***}	-0.074	0.020	-0.025
	(0.047)	(0.064)	(0.038)	(0.063)
Panel B: Gross Fertility				
Dependent variable: Number of children ever born				
Post	-0.077^{***}	-0.137^{***}	-0.146^{***}	-0.054^{***}
	(0.008)	(0.010)	(0.012)	(0.013)
PfPR_{2-10} (75%+) × Post	-0.060^{***}	-0.053^{***}	0.004	0.007
	(0.009)	(0.014)	(0.018)	(0.028)
Panel C: Net Fertility				
Dependent variable: Number of surviving children				
Post	-0.034^{***}	-0.126^{***}	-0.137^{***}	-0.033^{**}
	(0.010)	(0.011)	(0.012)	(0.015)
PfPR_{2-10} (75%+) × Post	-0.007	-0.035^{**}	0.010	0.035^{*}
	(0.019)	(0.017)	(0.013)	(0.018)
Observations	355,644	231,192	133,687	90,455

Table A.6. Heterogeneous Effects of the RBM on Years of Schooling by Gender

Notes: This table reports the estimation restuls from OLS regression (1.7), run separately for male and female. Brackets contain standard errors clustered at the region level. PfPR_{2-10} (75%+) × Post indicates the interaction between the indicator of high-prevalence regions (PfPR in 2001 exceeding 0.75) and the post-treatment indicator. Other interaction terms are defined similarily. Samples are restricted to the individuals who were born and residing (surveyed) in the same region in 2012. Variable Urban indicates whether the respondent reside in the urban part within the region. All columns include region fixed effects. *, **, and *** indicate significance at the 10, 5, 1% levels.

	Age group in 2012			
	Age 10-15	**Age 15-20**	**Age 20-25**	**Age 20-30**
Panel A: Male				
Dependent variable: Years of schooling				
Post	1.128***	1.358***	1.183***	0.404***
	(0.044)	(0.070)	(0.086)	(0.078)
PfPR_{2-10} (75%+) × Post	0.561***	0.997***	0.639***	-0.006
	(0.091)	(0.107)	(0.120)	(0.102)
Observations	551,298	414,836	296,759	269,074
Panel B: Female				
Post	1.161***	1.492***	0.914***	0.324***
	(0.043)	(0.073)	(0.089)	(0.085)
PfPR_{2-10} (75%+) × Post	0.710***	0.947***	0.378***	-0.272
	(0.105)	(0.144)	(0.120)	(0.276)
Observations	544,976	441,917	377,984	338,902

Appendix B

Appendix for Chapter 2

Table B.1. Summary Statistics for NLSY Sample

Notes: This table provides summary statistics for the NLSY97 sample that we use in our instrumental variables regression. The top panel includes all individuals in the sample, while the bottom panel includes only those with positive student debt.

	All Individuals Used in IV Sample						
Variable	Mean	Std. Dev.	Median	Min.	P25	P75	Max.
% Male	0.40	0.49	0.00	0.00	0.00	1.00	1.00
% White	0.70	0.46	1.00	0.00	0.00	1.00	1.00
Age at BA	23.16	2.73	22.00	19.00	21.00	24.00	34.00
Year of BA	2006	3	2006	2001	2004	2007	2015
HH Networth in 1997	138,384	134,914	95,375	250	33,000	197,751	599,001
Avg. HH Income	69,890	48,552	59,676	30	36,253	90,254	285,805
Ability Quartile	3.27	0.86	4.00	1.00	3.00	4.00	4.00
$ Student Loans	17,990	25,203	11,500	-	-	25,750	351,000
	Conditional on Positive Student Debt						
Variable	Mean	Std. Dev.	Median	Min.	P25	P75	Max.
% Male	0.38	0.49	0.00	0.00	0.00	1.00	1.00
% White	0.67	0.47	1.00	0.00	0.00	1.00	1.00
Age at BA	23.26	2.71	22.00	19.00	21.00	24.00	34.00
Year of BA	2006	3	2006	2001	2004	2008	2015
HH Networth in 1997	116,109	115,873	79,620	250	27,500	162,500	588,000
Avg. HH Income	62,417	40,421	55,200	30	34,000	80,350	285,805
Ability Quartile	3.24	0.87	3.00	1.00	3.00	4.00	4.00
$ Student Loans	27,259	26,643	20,975	300	12,000	35,000	351,000

Table B.2. IV First Stage Estimates

Effect of 1sd (10ppt) increase in college grant share on:	Total Funding	Grants	Debt	Family Aid	Work Study Aid	Tuition costs
Coefficient (pvalue)	-\$160 0.95	\$7,670 0.00	-\$5,076 0.00	-\$23 0.43	\$197 0.09	-\$863 0.56

Table B.3. IV Robustness Estimates

Effect of 1sd (10ppt) increase in college grant share on:	Years at college	Completion rate	1(Full-time)	Age starting college	Ability (percentile)	Parental income	1(White)
Coefficient (pvalue)	0.11 0.35	0.14% 0.28	0.01 0.21	-0.02 0.71	1.72% 0.23	\$1,145 0.60	0.38 0.11

Table B.4. Parameter Values

Category	**Notation**	**Parameter value**
Preferences	$\beta, u(c) = \frac{c^{1-\rho}-1}{1-\rho}$	$(\beta,\rho) = (0.985, 2)$
Human capital technology	$h' = \theta(sh)^{\alpha} + (1-\delta)h$	$(\delta,\alpha) = (0.05, 0.7)$
Interest rates	(r, r_d)	$(r, r_d) = (0.04, 0.042)$
Initial conditions	$\begin{pmatrix} a_0 \\ \tau \\ x \end{pmatrix} \sim \ln N \left[\begin{pmatrix} \mu_a \\ \mu_\tau \\ \mu_x \end{pmatrix}, \begin{pmatrix} \sigma_a^2 & \rho_{a\tau} & \rho_{ax} \\ \rho_{a\tau} & \sigma_\tau^2 & 0 \\ \rho_{ax} & 0 & \sigma_x^2 \end{pmatrix} \right]$	$(\mu_a,\mu_\tau,\mu_x,\rho_{a\tau},\rho_{ax},\sigma_\tau,\sigma_a)$ $=(3.67, 3.12, 0.0, -0.22, 0.5, 0.88, 1.76)$
	(μ_x, ρ_{ax})	$(\mu_x, \rho_{ax}) = (0.0, 0.37)$
Student debt repayment	$\bar{T}$	$\bar{T} = 10$
Social security system	$\pi(\hat{y}), \bar{y}$	See equation (2.7), $\bar{y} = \$71,700$
High school taste	$\zeta \sim \text{Frechet}(\varepsilon)$	$\varepsilon = 145$
Lifecycle risk of unemployment	$P(z_{t+1} = u \mid z_t = e)$	See Figure 2.5
Occupational heterogeneity	(K, σ_θ)	$(K, \sigma_\theta) = (18, 0.33)$
	$\{w_k, \mu_k, \nu_k\}_{k=1,\ldots,18}$	See Table B.5

Table B.5. Occupation-Specific Parameter Values

Occupation	k	w_k	μ_k	ν_k
Executive Admin	1	5.397	-1.529	0.0385
Management	2	5.387	-1.429	0.0376
Math and Computer Science	3	5.984	-1.523	0.0354
Architects and Engineers	4	5.848	-1.516	0.0365
Counslers	5	4.343	-1.714	0.0449
Teachers	6	4.332	-1.593	0.0449
Education	7	2.922	-1.897	0.0522
Sports	8	3.828	-1.419	0.0396
Media	9	4.277	-1.792	0.0448
Health Practitioners	10	5.192	-1.773	0.0404
Health Support	11	4.145	-1.880	0.0459
Food Services	12	3.127	-1.793	0.0514
Cleaning	13	3.748	-1.923	0.0478
Service Workers	14	3.078	-1.814	0.0516
Sales	15	4.897	-1.483	0.0410
Office & Admin	16	4.119	-1.585	0.0461
Maintenance	17	5.152	-1.970	0.0406
Transportation	18	3.519	-1.530	0.0419

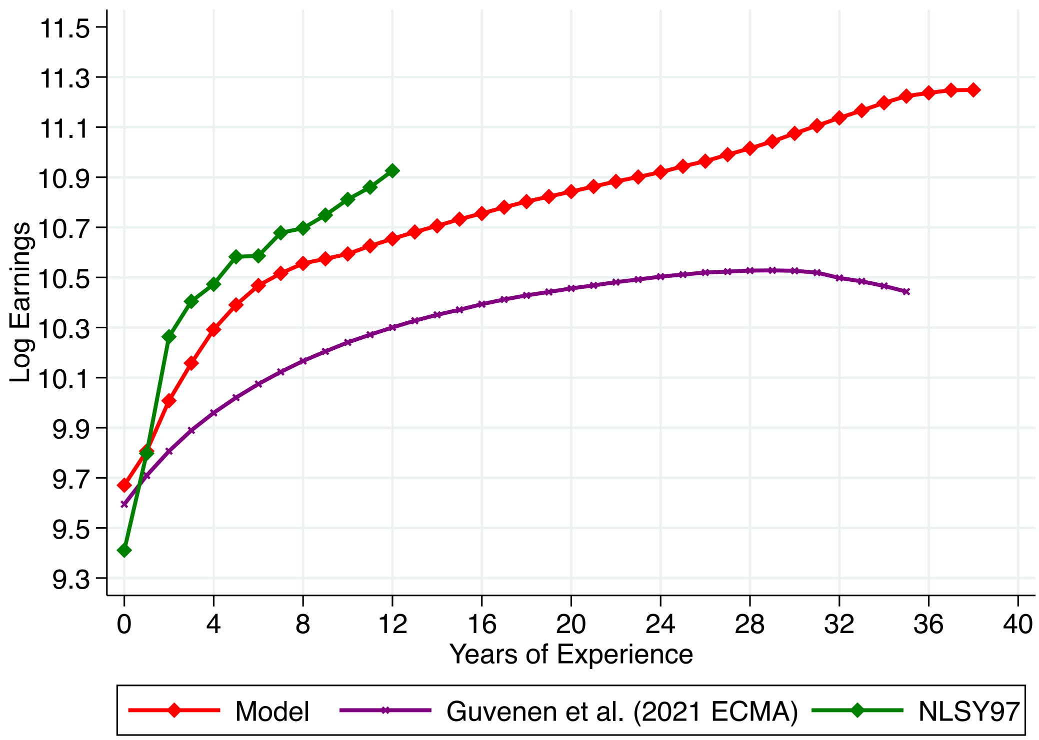


Figure B.1. Validation: Aggregate Lifecycle Profile

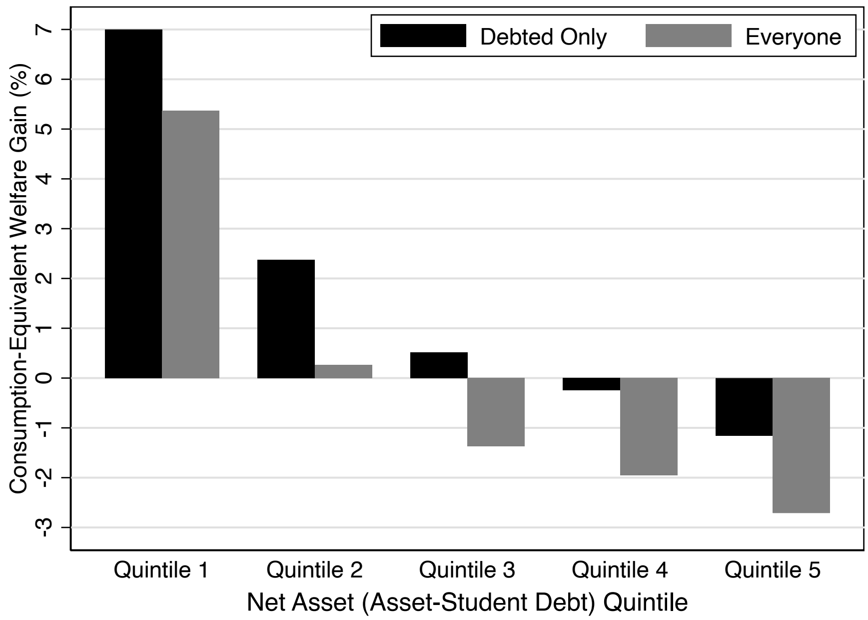


Figure B.2. Welfare by Net Assets under 50k Debt Forgiveness

Notes: This figures shows the distribution of welfare gains under the 50K-cap debt forgiveness policy across the household net asset distribution. Net asset is defined as initial asset minus total amount of student debt at the initial period. Black bars represent the welfare gains among individual who had positive student debts. Gray bars represent the welfare gains among everyone, including individuals without student debts.

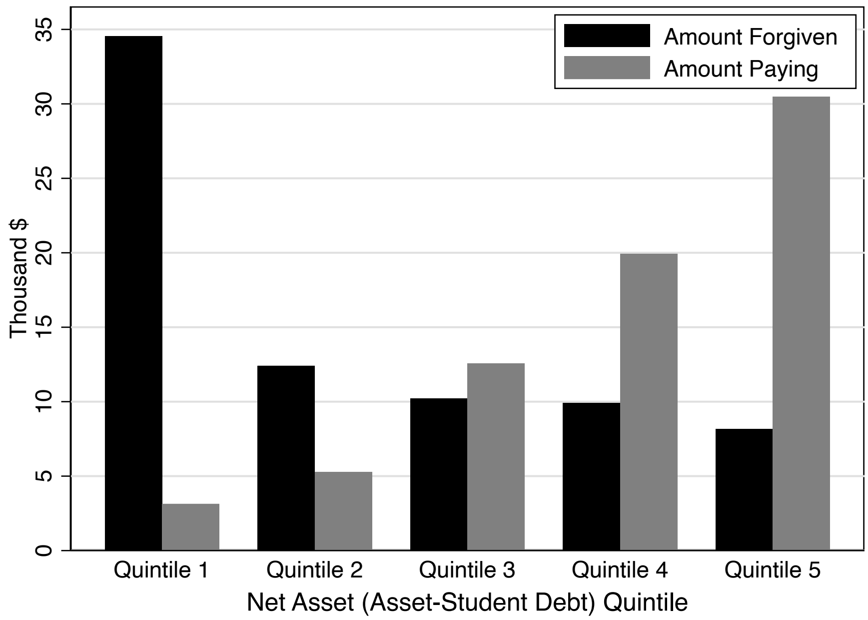


Figure B.3. Distribution of Costs and Benefits, 50k Debt Forgiveness

Notes: This figures shows the average amount of student debts forgiven and average amount of lump-sum taxes paid by the individuals within each net asset quintile. Net asset is defined as initial asset minus total amount of student debt at the initial period.

Appendix C

Appendix for Chapter 3

C.1 Appendix Figures for Chapter 3

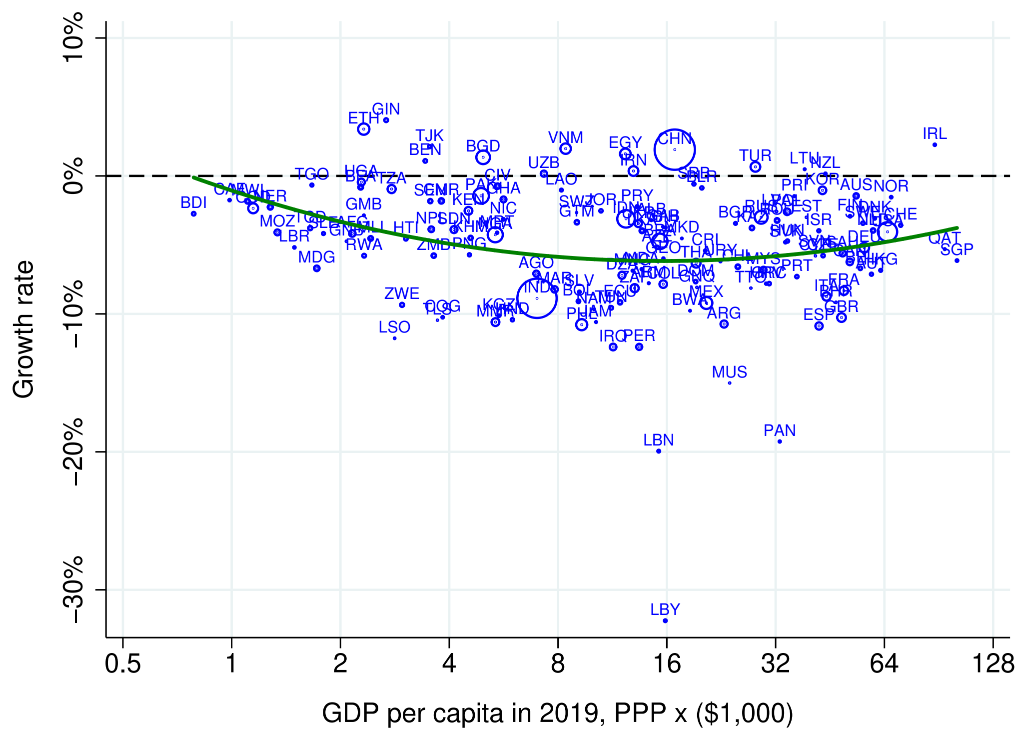


Figure C.1. GDP per capita Growth from 2019 to 2020

Notes: GDP per capita data comes from the World Bank World Development Indicators. GDP per capita is expressed at PPP and is taken from the Penn World Table 9.1 (Feenstra et al., 2015).

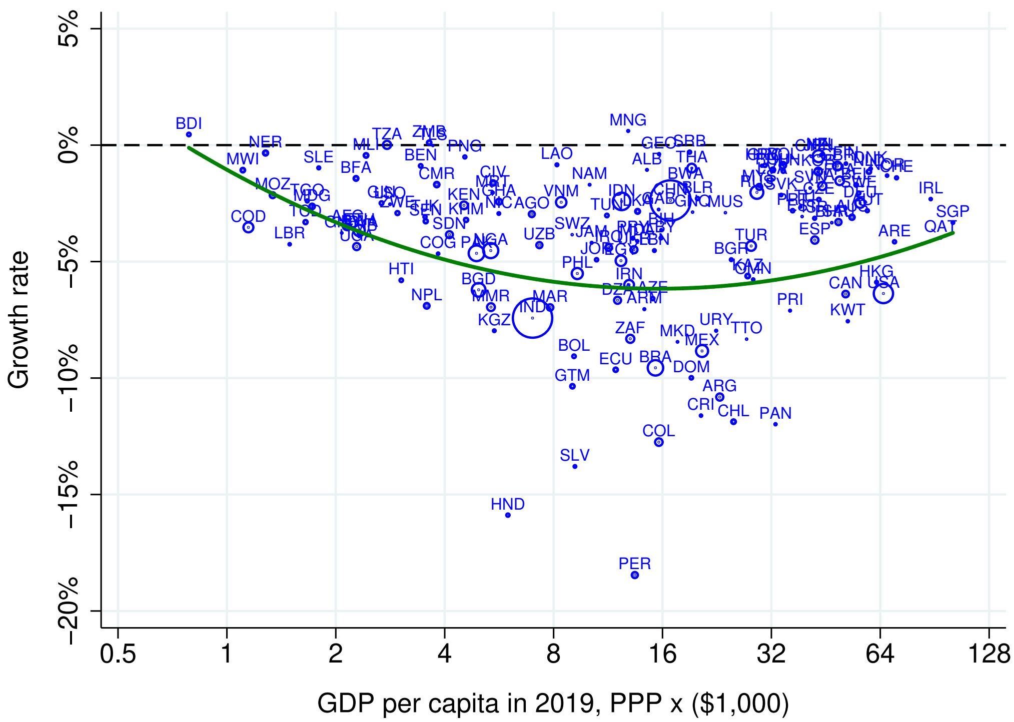


Figure C.2. Employment Growth from 2019 to 2020

Notes: Employment data comes from the ILO Statistical Database. GDP per capita is expressed at PPP and is taken from the Penn World Table 9.1 (Feenstra et al., 2015).

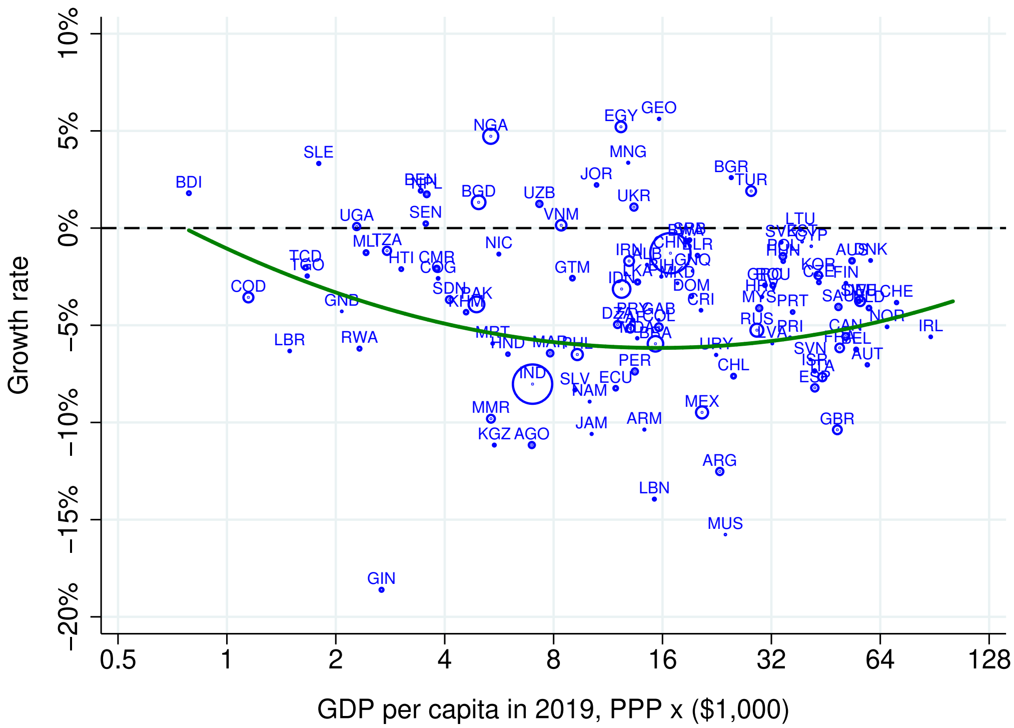


Figure C.3. Consumption-per-capita Growth from 2019 to 2020

Notes: Consumption data comes from the World Bank World Development Indicators. GDP per capita is expressed at PPP and is taken from the Penn World Table 9.1 (Feenstra et al., 2015).

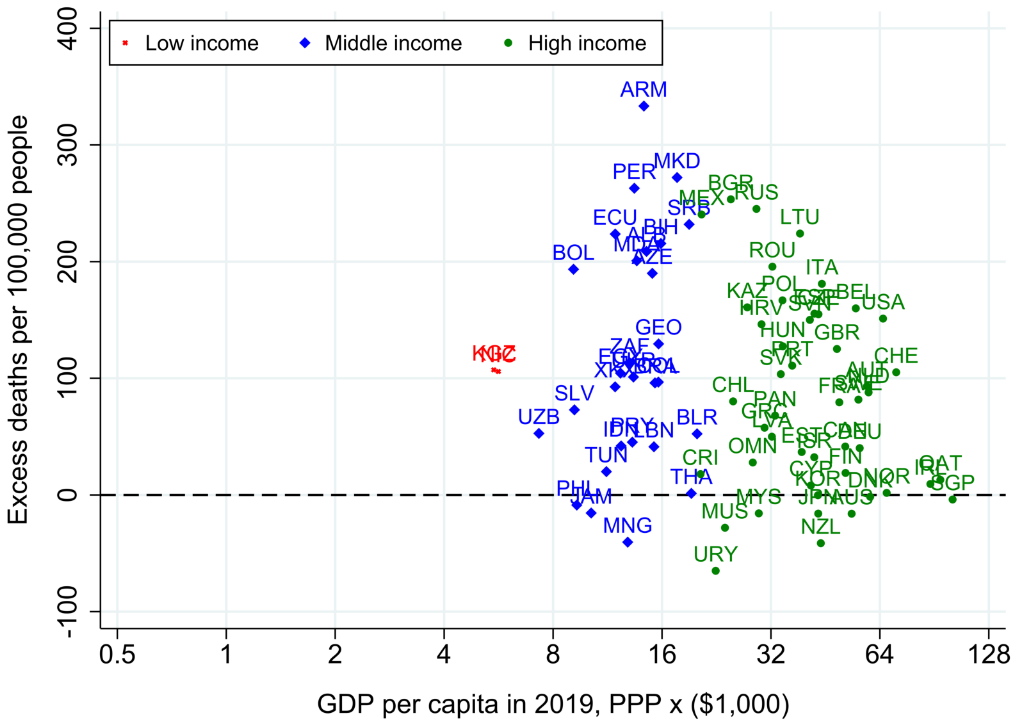


Figure C.4. Excess Deaths Estimated by *The Economist*

Notes: Data sourced from the Economist excess mortality tracker.

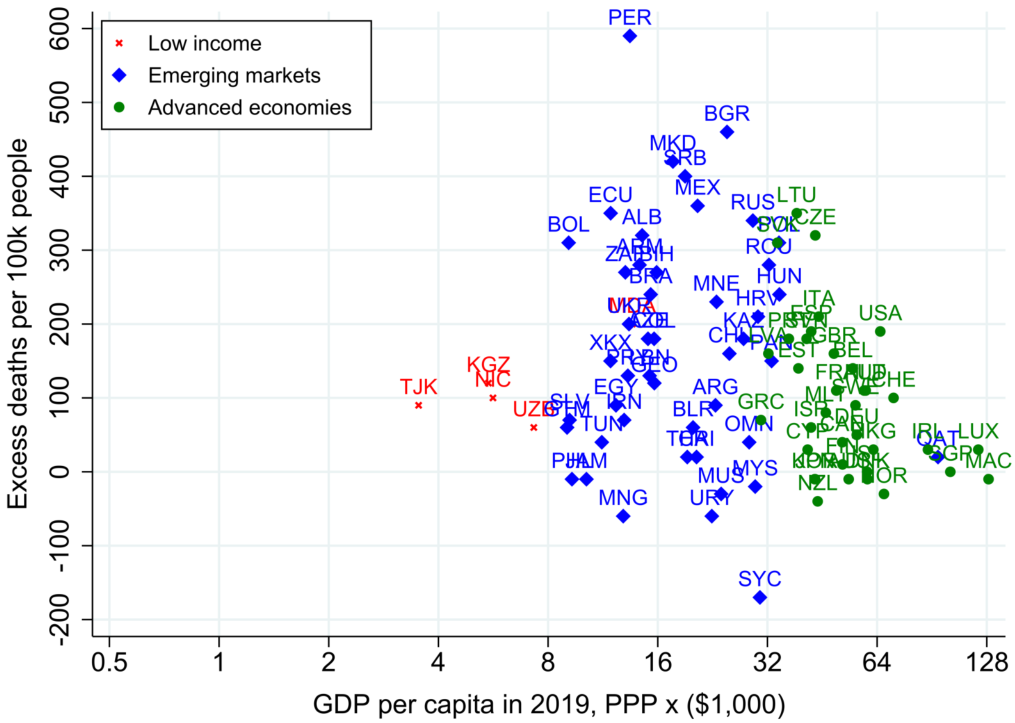


Figure C.5. Excess Deaths Estimated by Karlinsky & Kobak (2021)

Notes: Data sourced from Karlinsky & Kobak (2021)'s World Mortality Database.

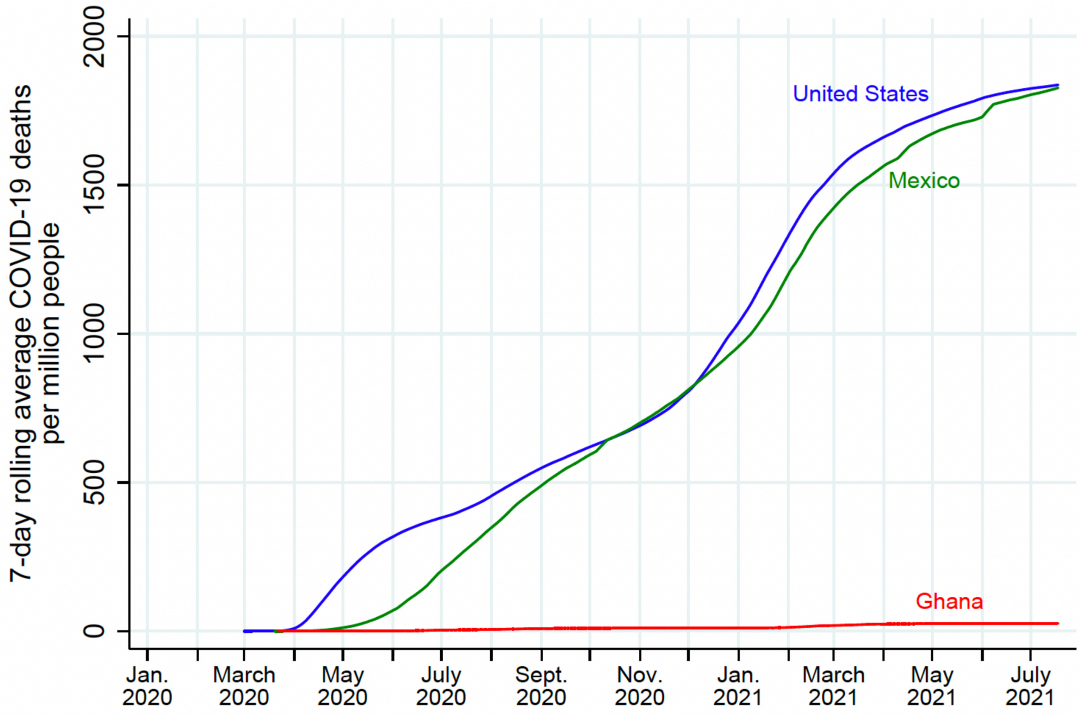


Figure C.6. Official COVID-19 Deaths in the United States, Mexico and Ghana

Notes: This figure plots cumulative official deaths from COVID-19, according to Our World in Data, in the three focus countries: the United States, Mexico and Ghana.

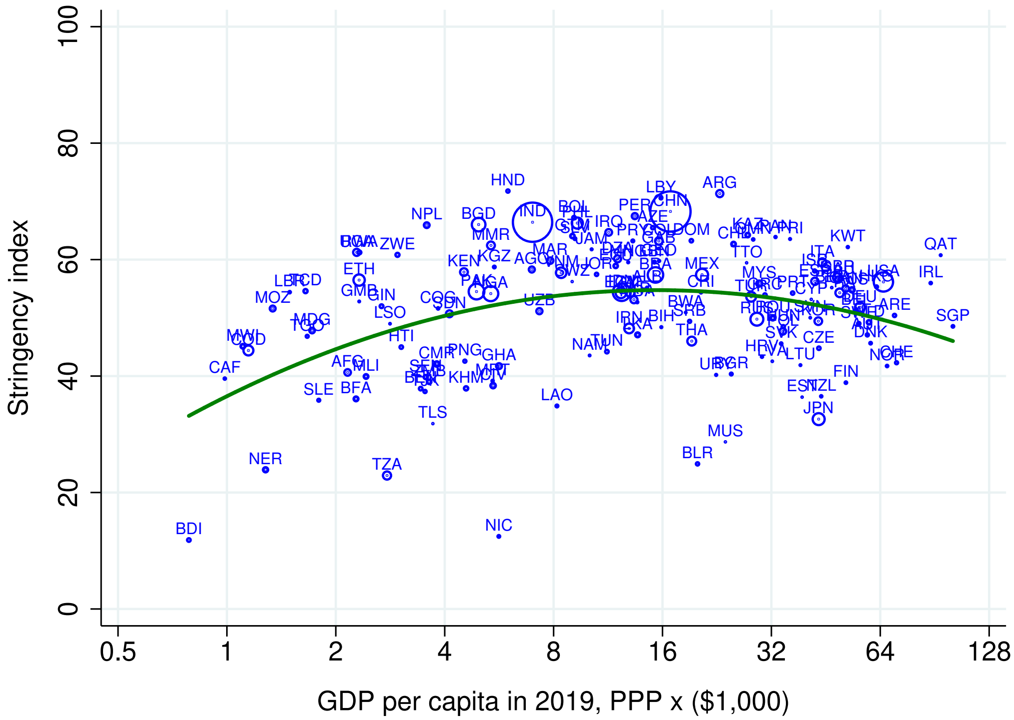


Figure C.7. Oxford Lockdown Stringency Index

Notes: The Government Stringency Index is taken from the Oxford Government Response Tracker (OxCGRT). GDP per capita is expressed at PPP and taken from Penn World Table 9.1 (Feenstra et al., 2015).

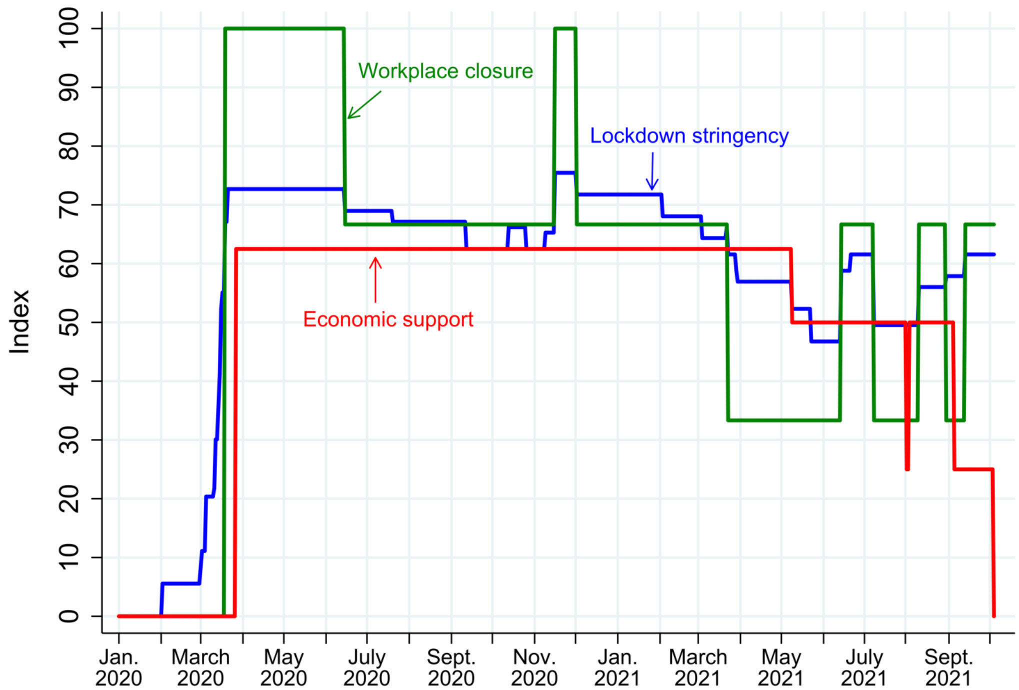


Figure C.8. Time-Series of Lockdown Policies and Economic Support in the United States

Notes: This figure displays the time-series of Oxford Lockdown Stringency Index, Economic Support Index, and Workplace Closures for the United States.

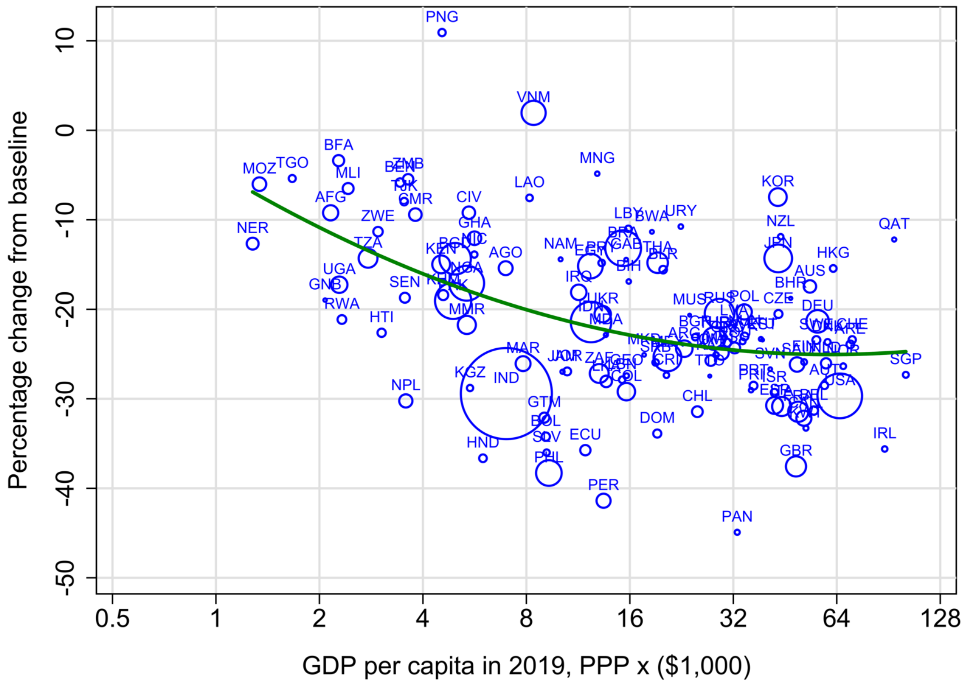


Figure C.9. Changes in Workplace Mobility in 2020

Notes: This figure plots the average percent change in visits and time spent at workplaces from baseline in 2020 against GDP per capita in 2019. The baseline is the median value, for the corresponding day of the week, during the 5-week period Jan 3–Feb 6, 2020. GDP per capita is expressed at PPP and taken from Penn World Table 9.1 (Feenstra et al., 2015). Percent change in visits and time spent at home and workplace in 2020 comes from the Google COVID-19 Community Mobility Reports.

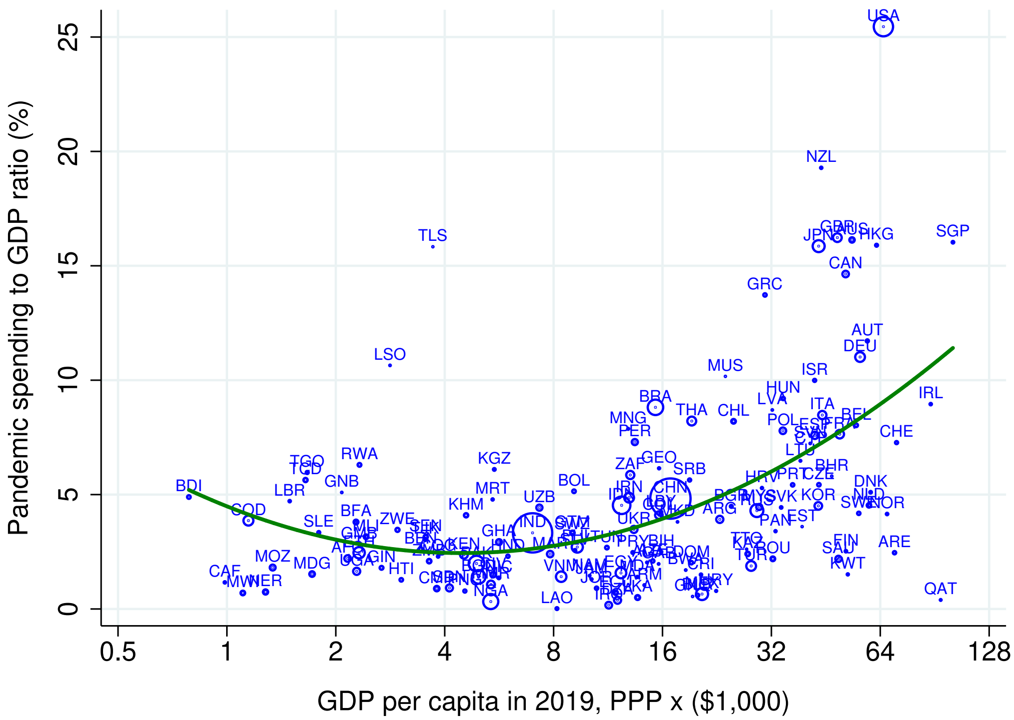


Figure C.10. Pandemic Spending as Share of GDP

Notes: Data on pandemic spending come from the IMF Fiscal Monitor Database. GDP per capita is expressed at PPP and taken from Penn World Table 9.1 (Feenstra et al., 2015).

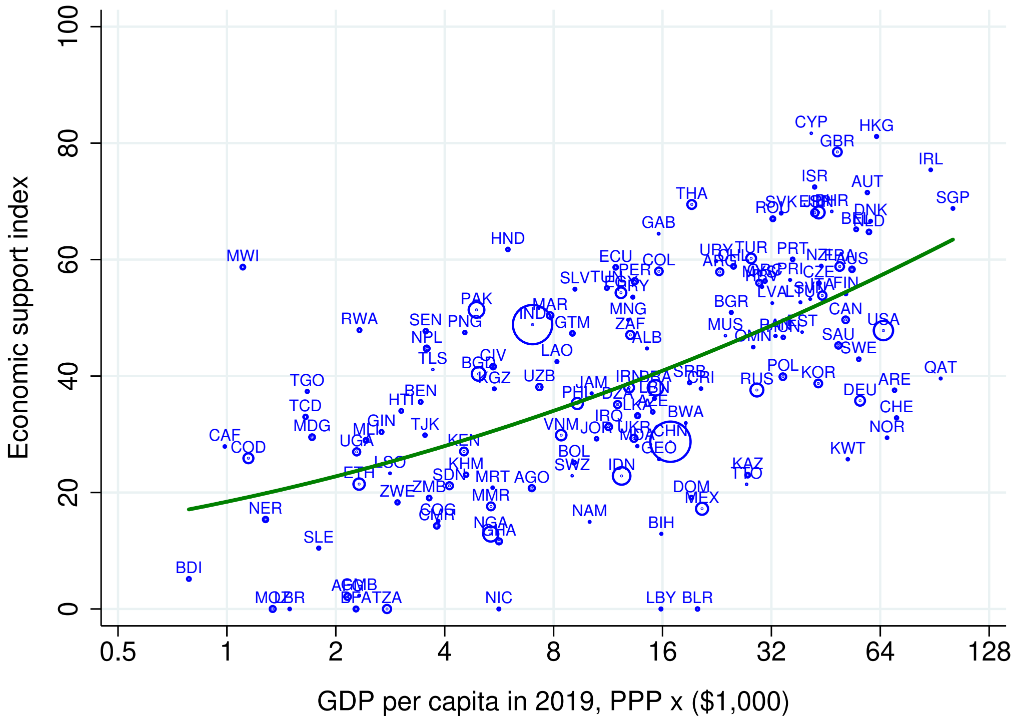


Figure C.11. Economic Support Index

Notes: Oxford Coronavirus Government Response Tracker's *Economic Support Index*.

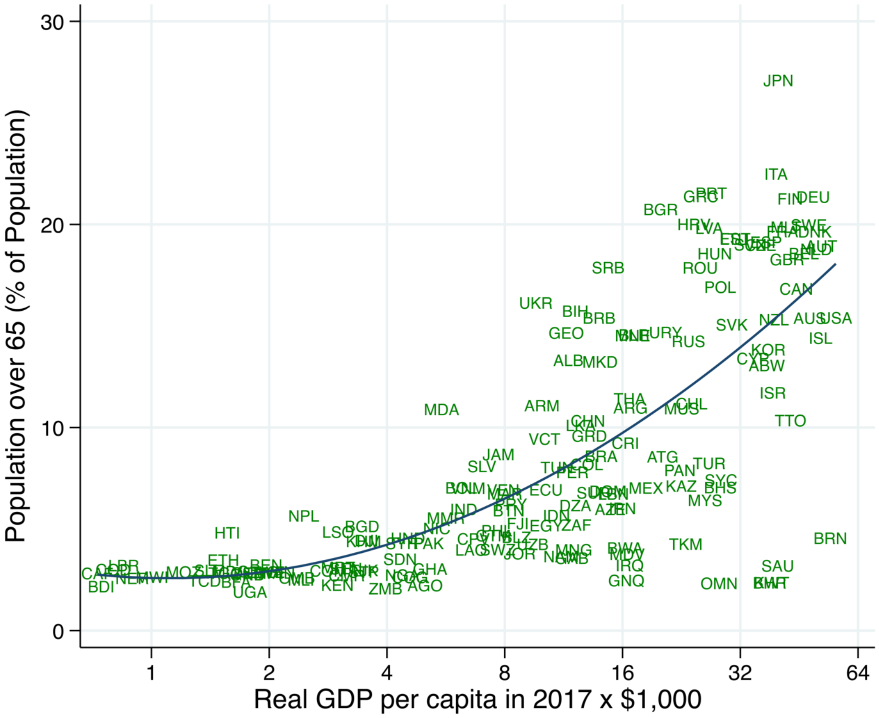


Figure C.12. Fraction of the Population Older than Age 65

Note: This figure plots the proportion of population ages over 65 and above as a percentage of total population across 162 countries. GDP per capita is from Penn World Table 9.1 (Feenstra et al., 2015). Population data is World Bank staff estimates using the World Bank's total population and age/sex distributions of the United Nations Population Division's World Population Prospects: 2019 Revision.

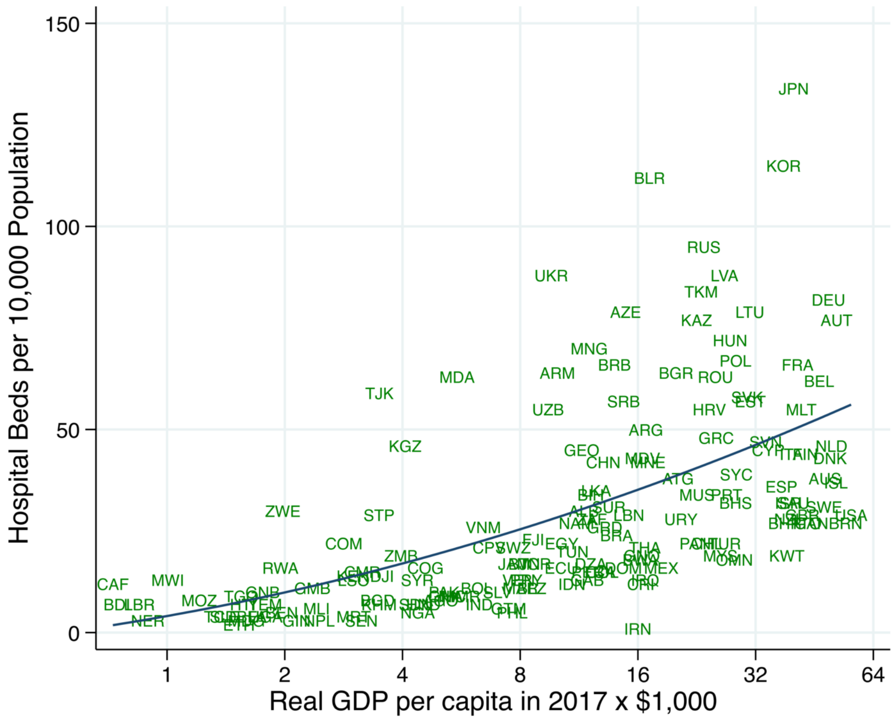


Figure C.13. Hospital Beds per 10,000 People

Note: This figure plots the number of hospital beds available per 10,000 inhabitants in 153 countries. GDP per capita is at PPP and taken from the Penn World Table 9.1 (Feenstra et al., 2015). The hospital bed data are from the World Health Organization's Global Health Observatory.

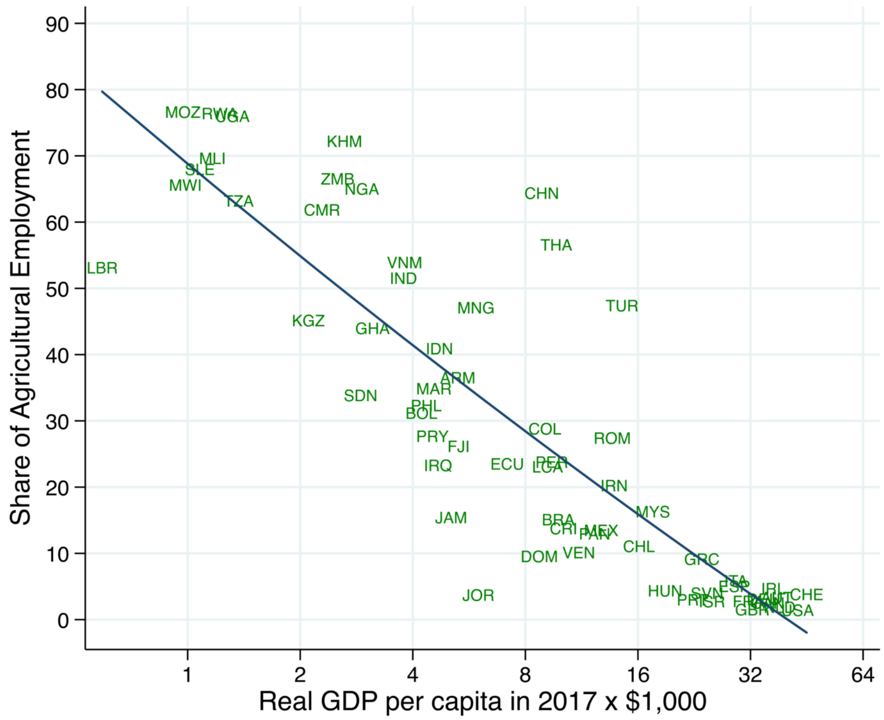


Figure C.14. Size of the Agricultural Sector

Note: Agriculture employment data is taken from the IPUMS database. GDP per capita is expressed at PPP and is taken from the Penn World Table 9.1 (Feenstra et al., 2015).

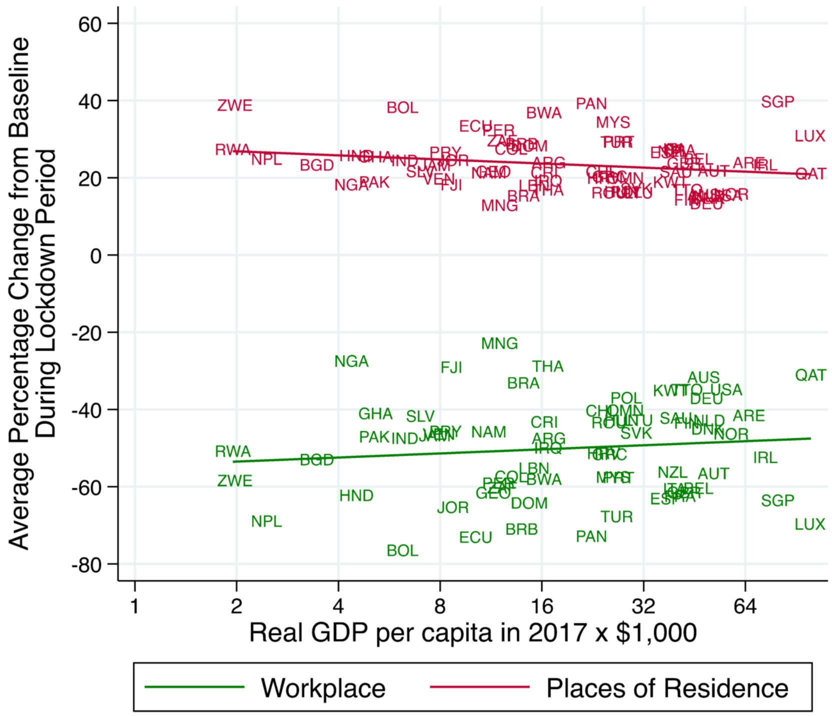


Figure C.15. Changes in Mobility Across Countries During Lockdown Periods

Note: This figure plots the average percentage changes of the mobility metric in the 'Places of Residence' and 'Workplace' categories in the Google Community Mobility Report (Bavadekar et al., 2020), during the lockdown periods for the 65 countries which had implemented or are implementing lockdown. GDP per capita is from Penn World Table 9.1 (Feenstra et al., 2015). The average across all 65 countries is 23.44 percent. The slope of the fitted line is 1.52, with p-value of 0.354 for the 'Workplace' category. For the 'Places of Residence' category, the slope of the fitted line is -1.52, with p-value of 0.083.

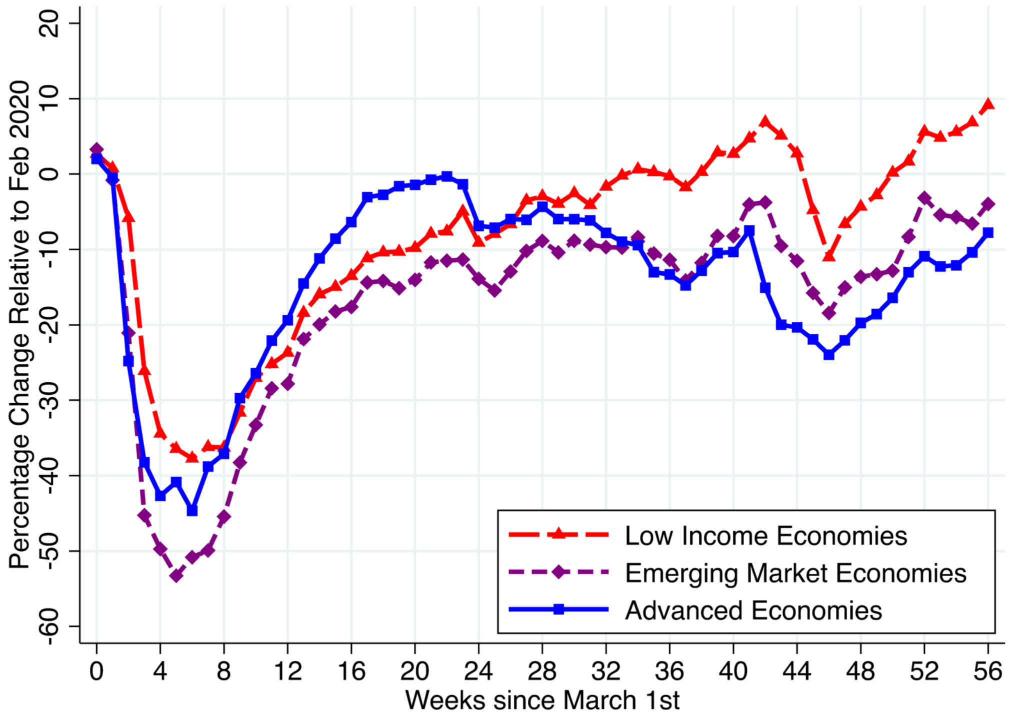


Figure C.16. Changes in Consumption-Related Mobility Across Countries

Note: This figure plots the weekly weighted averaged percentage changes of the mobility metric in the 'Grocery and Pharmacy' and 'Retail and Recreation' categories in the Google Community Mobility Report (Bavadekar et al., 2020), during the lockdown periods for the low-income, emerging market, and advanced economies (total 115 countries). Log GDP per capita for each country was used as a weight in pooling countries withint each group.

C.2 Appendix Tables for Chapter 3

Table C.1. ICU Bed Availability Across Countries

Source: Table 1 in Prin and Wunsch (2012). Healthcare cost includes all public and private expenditures.

Country	ICU beds per 100,000 population	Per capita healthcare cost
United States	20.0-31.7	$7,164
Canada	13.5	$3,867
Denmark	6.7-8.9	$3,814
Australia	8.0-8.9	$3,365
South Africa	8.9	$843
Sweden	5.8-8.7	$3,622
Spain	8.2-9.7	$2,941
Japan	7.9	$2,817
UK	3.5-7.4	$3,222
New Zealand	4.8-5.5	$2,655
China	2.8-4.6	$265
Trinidad and Tobago	2.1	$1,237
Sri Lanka	1.6	$187
Zambia	0	$80

Table C.2. Internally Calibrated Parameters and Model Fit

	Data	Model	Parameters	Description
Advanced Economies GDP Decline, '19-'20	-4.60%	-4.70%	$A(P)$	Pandemic TFP
US Employment Decline, '19-'20	-6.40%	-5.89%	ϕ_s	Productivity of remote work, social sector
Fraction Remote Workers pre COVID	8.20%	8.14%	σ_g	Variance of remote work taste shock

C.3 Robustness of Results for Chapter 3

In this section, we briefly check the robustness of our results to other parameters that we imagine could vary between Advanced and Emerging Market (EM) Economies. We focus on

Table C.3. Calibration of Economic Parameters

Param. Change	Description	GDP Decline (%)	Mortality (Per 100k)	Source
	Baseline	-6.3	159	
$\beta_o \to 0.9955$	Decrease Discount Factor for Old	-6.3	159	UN Population Division
$\xi \to 0.606$	Higher Infection Prob. when WFH	-6.3	172	Alon et al. (2020)
	Path of Vaccinations Reduced by 25%	-6.3	158	Our World in Data
$r^F \to 0.0003$	Reduced Return to Savings	-6.3	158	
$\phi_S \to 0.5$	Higher Penalty of Working Remotely	-8.0	161	

five additional model features that may vary and discuss each one in turn.

The first parameter we imagine might vary across advanced and EM economies is the discount factor of the old. In our model, this discount factor embeds the notion that individuals over the age of 65 have a lower expected lifespan than individuals under the age of 65 and thus carry a lower discount factor. Because life expectancy is lower in EE economies, it is intuitive that the ”gap” between the discount factor of the young and the old should be larger. Quantitatively, we discipline this variation by noting that individuals at age 65 have 20 percent lower life expectancy in EE economies (according to a UN Population Division note), so the gap between the young and old discount rates should be 20 larger. This results in the implied value for β_o of 0.9955.

We also imagine that working from home may not provide as much protection from infection in EM economies due to larger household sizes (i.e. more contacts at home) and participation in crowded markets. We refer to Alon et al. (2020) who use contact matrices constructed by Mossong et al. (2008) and measure that individuals in advanced economies make about 36 percent of their contacts at work while individuals in EM economies make about 35 percent at work. Thus working from home reduces contacts by about 1 percent ($\frac{36.1-35.8}{36.1}$) less in EM economies. Consequently, we inflate the parameter governing (relative) infection risk while working from home by 1 percent to 0.606.

Finally, vaccine rollout was substantially delayed in EM economies relative to advanced. We use vaccination data from Our World in Data (OWID). In particular, OWID reports the time path of vaccinations for countries grouped into ”High income”, ”Upper middle income”, and ”Lower middle income”. We consider the ”High income” series to correspond to advanced economies and the average of the ”Upper middle income” and ”Lower middle income” series to correspond to EM economies. From this, we see that EM economies tend to have roughly 25 percent fewer vaccinations at any given point in time, which we feed into the model.

We also may be concerned that EM economies are more borrowing constrained and have more hand-to-mouth households than advanced economies. We capture this variation by reducing the interest rate faced by these economies, which reduces savings and leads more binding borrowing constraints. We also might be concerned that infrastructure for remote work is less robust in EM economies, resulting in social-sector workers earning a larger penalty from choosing to work remotely. We capture this variation by reducing the fraction of income that social sector workers keep when choosing to work from hope. Lacking good data for both of these types of cross-country variation, we simply decrease r_F and ϕ_S by the arbitrary amounts of 50 percent and 20 percent respectively to explore the robustness of the model conclusions to these parameters.

Encouragingly, very few of these parameter changes have a substantial impact on model-predicted outcomes for EM economies. The two largest impacts are the increase in infection probability when working from home leading to slightly more deaths and the higher penalty of working remotely leading to slightly more GDP loss. Both of these results are intuitive as these parameters directly govern these outcomes essentially by construction.

Bibliography

ABALUCK, J., L. H. KWONG, A. STYCZYNSKI, A. HAQUE, M. A. KABIR, E. BATES-JEFFRIES, E. CRAWFORD, J. BENJAMIN-CHUNG, S. RAIHAN, S. RAHMAN, S. BEN-HACHMI, N. ZAMAN, P. J. WINCH, M. M. HOSSAIN, H. M. REZA, A. A. JABER, S. G. M. F. L. B. A. RAHMAN, T. S. HUQ, S. P. LUBY, AND A. M. MOBARAK (2021): "The Impact of Community Masking on COVID-19: A Cluster Randomized Trial in Bangladesh," Unpublished Manuscript, Yale University.

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., V. CHERNOZHUKOV, I. WERNING, AND M. D. WHINSTON (2020): "A Multi-Risk SIR Model with Optimally Targeted Lockdown," NBER Working Paper No. 27102.

ACEMOGLU, D. AND S. JOHNSON (2007): "Disease and Development: the Effect of Life Expectancy on Economic Growth," *Journal of Political Economy*, 115, 925–985.

ACEVEDO, I., F. CASTELLANI, G. LOTTI, AND M. SZÉKELY (2021): "Informalidad en los tiempos del COVID-19 en América Latina: implicaciones y opciones de amortiguamiento," Tech. rep., IDB Working Paper Series.

AIYAGARI, S. R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics*, 109, 659–684.

AL SEROURI, A. W., S. M. GRANTHAM-MCGREGOR, B. GREENWOOD, AND A. COSTELLO (2000): "Impact of Asymptomatic Malaria Parasitaemia on Cognitive Function and School Achievement of Schoolchildren in the Yemen Republic," *Parasitology*, 121, 337–345.

ALENE, M., L. YISMAW, M. A. ASSEMIE, D. B. KETEMA, B. MENGIST, B. KASSIE, AND T. Y. BIRHAN (2021): "Magnitude of Asymptomatic COVID-19 Cases Throughout the Course of Infection: A Systematic Review and Meta-Analysis," *PloS One*, 16, e0249090.

ALESINA, A., S. HOHMANN, S. MICHALOPOULOS, AND E. PAPAIOANNOU (2021): "Intergenerational Mobility in Africa," *Econometrica*, 89, 1–35.

ALFARO, L., O. BECERRA, AND M. ESLAVA (2020): "EMEs and COVID-19: Shutting Down in a World of Informal and Tiny Firms," NBER Working Paper No. 27360.

ALON, T. M., M. KIM, D. LAGAKOS, AND M. VANVUREN (2020): "How Should Policy Responses to the Covid-19 Pandemic Differ in the Developing World?" NBER Working Paper No. 27273.

ALVAREZ, F. E., D. ARGENTE, AND F. LIPPI (2020): "A Simple Planning Problem for COVID-19 Lockdown," NBER Working Paper No. 26981.

ARELLANO, C., Y. BAI, AND G. P. MIHALACHE (2020): "Deadly Debt Crises: COVID-19 in Emerging Markets," NBER Working Paper No. 27275.

——— (2021): "COVID-19 Vaccination in Emerging Markets," Unpublished Working Paper, University of Rochester.

ASHRAF, Q. H., A. LESTER, AND D. N. WEIL (2008): "When Does Improving Health Raise GDP?" *NBER Macroeconomics Annual*, 23, 157–204.

BAIROLIYA, N. AND A. IMROHOROGLU (2020): "Macroeconomic Consequences of Stay-at-Home Policies During the COVID-19 Pandemic," *COVID Economics Vetted and Real-Time Papers*, 13, 71–90.

BANERJEE, A. V. AND E. DUFLO (2005): "Growth Theory Through the Lens of Development Economics," *Handbook of Economic Growth*, 1, 473–552.

BAROFSKY, J., T. D. ANEKWE, AND C. CHASE (2015): "Malaria Eradication and Economic Outcomes in Sub-Saharan Africa: Evidence from Uganda," *Journal of Health Economics*, 44, 118–136.

BARRERO, J. M., N. BLOOM, AND S. J. DAVIS (2021): "Why Working from Home Will Stick," NBER Working Paper No. 28731.

BARRO, R. J. AND G. S. BECKER (1989): "Fertility Choice in a Model of Economic Growth," *Econometrica*, 481–501.

BAUM, S. AND M. O'MALLEY (2003): "College on Credit: How Borrowers Perceive Their Education Debt," *NASFAA Journal of Student Financial Aid*, 33, 7–19.

BAVADEKAR, S., A. DAI, J. DAVIS, D. DESFONTAINES, I. ECKSTEIN, K. EVERETT, A. FABRIKANT, G. FLORES, E. GABRILOVICH, K. GADEPALLI, S. GLASS, R. HUANG, C. KAMATH, D. KRAFT, A. KUMOK, H. MARFATIA, Y. MAYER, B. MILLER, A. PEARCE,

I. MILINDA PERERA, V. RAMACHANDRAN, K. RAMAN, T. ROESSLER, I. SHAFRAN, T. SHEKEL, C. STANTON, J. STIMES, M. SUN, G. WELLENIUS, AND M. ZOGHI (2020): "Google COVID-19 Community Mobility Reports: Anonymization Process Description (version 1.0)," .

BEN-PORATH, Y. (1967): ""The Production of Human Capital and the Life Cycle of Earnings"," *Journal of Political Economy*, 75, 352–365.

BERGER, D. W., K. F. HERKENHOFF, AND S. MONGEY (2020): "An SEIR Infectious Disease Model with Testing and Conditional Quarantine," NBER Working Paper No. 26901.

BESLEY, T. AND T. PERSSON (2013): "Taxation and Development," in *Handbook of Public Economics*, Elsevier, vol. 5, 51–110.

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

BICK, A., A. BLANDIN, AND K. MERTENS (2020): "Work from Home After the COVID-19 Outbreak," Federal Reserve Bank of Dallas Working Paper 2017.

BILS, M., B. KAYMAK, AND K.-J. WU (2022): "Labor Substitutability among Schooling Groups," NBER Working Paper No. 29895.

BLEAKLEY, H. (2010): "Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure," *American Economic Journal: Applied Economics*, 2, 1–45.

BONADIO, B., Z. HUO, A. A. LEVCHENKO, AND N. PANDALAI-NAYAR (2021): "Global Supply Chains in the Pandemic," *Journal of International Economics*, 133.

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

BROTHERHOOD, L., P. KIRCHER, C. SANTOS, AND M. TERTILT (2021): "An Economic Model of the Covid-19 Pandemic with Young and Old Agents: Behavior, Testing and Policies," Unpublished Manuscript, University of Mannheim.

BUERA, F. J., J. P. KABOSKI, AND Y. SHIN (2021a): "The Macroeconomics of Microfinance," *The Review of Economic Studies*, 88, 126–161.

BUERA, F. J., J. P. KABOSKI, AND R. M. TOWNSEND (2021b): "From Micro to Macro Development," NBER Working Paper No. 28423.

BUSSO, M., J. CAMACHO, J. MESSINA, AND G. MONTENEGRO (2021): "Social protection and informality in Latin America during the COVID-19 pandemic," *PloS one*, 16, e0259050.

CAKMAKLI, C., S. DEMIRALP, AND S. K. OZCAN (2020): "COVID-19 and Emerging Markets: An Epidemiological Model with International Production Networks and Capital Flows," IMF Working Paper No. 20/133.

CALLAWAY, B. AND P. H. SANT'ANNA (2021): "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 225, 200–230.

CASELLI, F. (2005): "Accounting for Cross-Country Income Differences," *Handbook of Economic Growth*, 1, 679–741.

CHAKRABARTI, R., V. FOS, A. LIBERMAN, AND C. YANNELIS (2020): "Tuition, Debt, and Human Capital," Staff Reports 912, Federal Reserve Bank of New York.

CHAPMAN, S. (2015): "Student Loans and the Labor Market: Evidence from Merit Aid Programs," Unpublished Manuscript, Northwestern University.

CHEN, I., S. E. CLARKE, R. GOSLING, B. HAMAINZA, G. KILLEEN, A. MAGILL, W. O'MEARA, R. N. PRICE, AND E. M. RILEY (2016): ""Asymptomatic" Malaria: a Chronic and Debilitating Infection that Should be Treated," *PLoS Medicine*, 13, e1001942.

CHOPRA, A., M. B. DEVEREUX, AND A. LAHIRI (2021): "Pandemics Through the Lens of Occupations," *Canadian Journal of Economics*, 54.

CONSUMER FINANCIAL PROTECTION BUREAU (2022): "Report of the CFPB Education Loan Ombudsman," .

CURRIE, J. AND T. VOGL (2013): "Early-Life Health and Adult Circumstance in Developing Countries," *Annuual Review of Economics*, 5, 1–36.

CUTLER, D., W. FUNG, M. KREMER, M. SINGHAL, AND T. VOGL (2010): "Early-Life Malaria Exposure and Adult Outcomes: Evidence from Malaria Eradication in India," *American Economic Journal: Applied Economics*, 2, 72–94.

DAHAL, S., J. M. BANDA, A. I. BENTO, K. MIZUMOTO, AND G. CHOWELL (2021): "Characterizing All-Cause Excess Mortality Patterns During COVID-19 Pandemic in Mexico," *BMC Infectious Diseases*, 21.

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

DARUICH, D. AND J. KOZLOWSKI (2020): "Explaining Intergenerational Mobility: The Role of Fertility and Family Transfers," *Review of Economic Dynamics*, 36, 220–245.

DATOO, M. S., H. M. NATAMA, A. SOMÉ, D. BELLAMY, O. TRAORÉ, T. ROUAMBA,

M. C. Tahita, N. F. A. Ido, P. Yameogo, D. Valia, A. Milogo, F. Ouedraogo, R. Soma, S. Sawadogo, F. Sorgho, K. Derra, E. Rouamba, F. Ramos-Lopez, M. Gairns, S. Provstgaard-Morys, J. Aboagye, A. Lawrie, R. Roberts, I. Valéa, H. Sorgho, N. Williams, G. Glenn, L. Fries, J. Reimer, K. Ewer, U. Shaligram, A. V S Hill, and H. Tinto (2022): "Efficacy and Immunogenicity of R21/Matrix-M Vaccine Against Clinical Malaria After 2 years' Follow-up in Children in Burkina Faso: a Phase 1/2b Randomised Controlled Trial," *The Lancet Infectious Diseases*.

De Philippis, M. and F. Rossi (2021): "Parents, Schools and Human Capital Differences Across Countries," *Journal of the European Economic Association*, 19, 1364–1406.

Deshmukh, Y., W. Suraweera, C. Tumbe, A. Bhowmick, S. Sharma, P. Novosad, S. H. Fu, L. Newcombe, H. Gelband, P. Brown, and P. Jha (2021): "Excess Mortality in India from June 2020 to June 2021 During the COVID Pandemic: Death Registration, Health Facility Deaths, and Survey Data," Unpublished Working Paper, University of Toronto.

Djankov, S. and U. Panizza, eds. (2020): *COVID-19 in Developing Economies*, London, UK: Centre for Economic Policy Research Press.

Doepke, M. (2005): "Child Mortality and Fertility Decline: Does the Barro-Becker Model Fit the Facts?" *Journal of Population Economics*, 18, 337–366.

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

Erosa, A., T. Koreshkova, and D. Restuccia (2010): "How Important is Human Capital? A Quantitative Theory Assessment of World Income Inequality," *The Review of Economic Studies*, 77, 1421–1449.

Federal Reserve Bank of New York (2022): "Quarterly Report on Household Debt and Credit," .

Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015): "The Next Generation of the Penn World Table," *American Economic Review*, 105, 3150–82.

Ferguson, N., D. Laydon, G. Nedjati Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunuba Perez, G. Cuomo-Dannenburg, A. DIghe, I. Dorigatti, H. Fu, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, L. C Okell, S. van Elsland, H. Thompson, R. Verity, E. Volz, H. Wang, Y. Wang, P. GT Walker, C. Walters, P. Winskill, C. Whittaker, C. A Donnelly, S. Riley, and A. C Ghani (2020): "Report 9: Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand," Unpublished Working Paper, Imperial College London.

FERNANDO, S., D. GUNAWARDENE, M. BANDARA, D. DE SILVA, R. CARTER, K. MENDIS, AND A. WICKREMASINGHE (2003): "The Impact of Repeated Malaria Attacks on the School Performance of Children," .

FERNANDO, S. D., C. RODRIGO, AND S. RAJAPAKSE (2010): "The 'Hidden' Burden of Malaria: Cognitive Impairment Following Infection," *Malaria Journal*, 9, 1–11.

FIELD, E. (2009): "Educational debt burden and career choice: Evidence from a financial aid experiment at NYU Law School," *American Economic Journal: Applied Economics*, 1, 1–21.

FLODEN, M. AND J. LINDÉ (2001): "Idiosyncratic Risk in the United States and Sweden: Is There a Role for Government Insurance?" *Review of Economic Dynamics*, 4, 406–437.

FOLBRE, N. (2008): *Valuing Children: Rethinking the Economics of the Family*, Harvard University Press.

FRIED, S. AND D. LAGAKOS (Forthcoming): "Electricity and Firm Productivity: A General-Equilibrium Approach," *American Economic Journal: Macroeconomics*.

FU, C., H.-C. LIN, AND A. TANAKA (2021): "College Loans and post-Schooling Human Capital Investment," Unpublished Manuscript, University of Wisconsin-Madison.

FUCHS-SCHÜNDELN, N., D. KRUEGER, A. LUDWIG, AND I. POPOVA (2022): "The long-Term Distributional and Welfare Effects of Covid-19 School Closures," *The Economic Journal*, 132, 1647–1683.

GLOVER, A., J. HEATHCOTE, D. KRUEGER, AND J.-V. RÍOS-RULL (2020): "Health Versus Wealth: On the Distributional Effects of Controlling a Pandemic," NBER Working Paper No. 27046.

GLYNN, J. R. (2020): "Protecting Workers Aged 60–69 Years from COVID-19," *The Lancet Infectious Diseases; (published online April 16)*.

GOLLIN, D. (2002): "Getting Income Shares Right," *Journal of Political Economy*, 110, 458–474.

GONÇALVES, B. P., C.-Y. HUANG, R. MORRISON, S. HOLTE, E. KABYEMELA, D. R. PREVOTS, M. FRIED, AND P. E. DUFFY (2014): "Parasite Burden and Severity of Malaria in Tanzanian Children," *New England Journal of Medicine*, 370, 1799–1808.

GOODMAN, S. F., A. ISEN, AND C. YANNELIS (2018): "A Day Late and a Dollar Short : Liquidity and Household Formation among Student Borrowers," Finance and Economics Discussion Series 2018-025, Board of Governors of the Federal Reserve System (U.S.).

GOTTLIEB, C., J. GROBOVŠEK, M. POSCHKE, AND F. SALTIEL (2021a): "Lockdown Accounting," *The B.E. Journal of Macroeconomics*, 22, 197–210.

——— (2021b): "Working from Home in Developing Countries," *European Economic Review*, 133, 103–679.

GREENWOOD, J., P. KIRCHER, C. SANTOS, AND M. TERTILT (2019): "An Equilibrium Model of the African HIV/AIDS epidemic," *Econometrica*, 87, 1081–1113.

GUVENEN, F., F. KARAHAN, S. OZKAN, AND J. SONG (2021): "What Do Data on Millions of US Workers Reveal about Lifecycle Earnings Dynamics?" *Econometrica*, 89, 2303–2339.

HANUSHEK, E. A. AND L. WOESSMANN (2008): "The Role of Cognitive Skills in Economic Development," *Journal of Economic Literature*, 46, 607–68.

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

HERKENHOFF, K., G. PHILLIPS, AND E. COHEN-COLE (2016): "How Credit Constraints Impact Job Finding Rates, Sorting and Aggregate Output," NBER Working Paper No. 22274.

——— (2021): "The Impact of Consumer Credit Access on Self-Employment and Entrepreneurship," *Journal of Financial Economics*, 141, 345–371.

HERKENHOFF, K. F. (2019): "The Impact of Consumer Credit Access on Unemployment," *The Review of Economic Studies*, 86, 2605–2642.

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.

——— (1996): "Wealth Distribution in Life-Cycle Economies," *Journal of Monetary Economics*, 38, 469–94.

HUGGETT, M., G. VENTURA, AND A. YARON (2011): "Sources of Lifetime Inequality," *American Economic Review*, 101, 2923–54.

IONESCU, F. (2009): "The Federal Student Loan Program: Quantitative Implications for College Enrollment and Default Rates," *Review of Economic Dynamics*, 12, 205–231.

JOHANSSON, E. W., R. E. CIBULSKIS, AND R. W. STEKETEE (2010): *Malaria Funding and*

Resource Utilization: the First decade of Roll Back Malaria, World Health Organization.

JONES, C. I. (2022): "Taxing Top Incomes in a World of Ideas," *Journal of Political Economy*, 130, 2227–2274.

JONES, C. I. AND P. J. KLENOW (2016): "Beyond GDP? Welfare across Countries and Time," *American Economic Review*, 106, 2426–57.

JONES, L. E. AND A. SCHOONBROODT (2010): "Complements versus Substitutes and Trends in Fertility Choice in Dynastic Models," *International Economic Review*, 51, 671–699.

JOSEPH, C. (2020): "Education and Labour Market Earnings in Low Income Countries: Empirical Evidence for Tanzania," *Tanzania Journal for Population studies and Development*, 26.

KALEMLI-OZCAN, S. (2003): "A Stochastic Model of Mortality, Fertility, and Human Capital Investment," *Journal of Development Economics*, 70, 103–118.

KALEMLI-OZCAN, S., H. E. RYDER, AND D. N. WEIL (2000): "Mortality Decline, Human Capital Investment, and Economic Growth," *Journal of Development Economics*, 62, 1–23.

KAPLAN, G., B. MOLL, AND G. L. VIOLANTE (2020): "The Great Lockdown: Macroeconomic and Distributional Effects of Covid-19," Unpublished Manuscript, University of Chicago.

KARAIVANOV, A., S. E. LU, H. SHIGEOKA, C. CHEN, AND S. PAMPLONA (2021): "Face Masks, Public Policies and Slowing the Spread of COVID-19: Evidence from Canada," *Journal of Health Economics*, 78.

KARLINSKY, A. AND D. KOBAK (2021): "The World Mortality Dataset: Tracking Excess Mortality across Countries During the COVID-19 Pandemic," Https://doi.org/10.1101/2021.01.27.21250604.

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

KRUEGER, D., H. UHLIG, AND T. XIE (2020): "Macroeconomic Dynamics and Reallocation in an Epidemic," NBER Working Paper No. 27047.

KUECKEN, M., J. THUILLIEZ, AND M.-A. VALFORT (2021): "Disease and Human Capital Accumulation: Evidence from the Roll Back Malaria Partnership in Africa," *The Economic Journal*.

LAGAKOS, D., A. M. MOBARAK, AND M. E. WAUGH (2023): "The Welfare Effects of Encouraging Rural-Urban Migration," *Econometrica*, 91, 803–838.

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

LEYARO, V., P. TWUMASI BAFFOUR, O. MORRISSEY, AND T. OWENS (2014): “Determinants of Urban Labour Earnings in Tanzania, 2000/01-06,” Unpublished Manuscript, The University of Nottingham.

LOAYZA, N. AND S. M. PENNINGS (2020): “Macroeconomic Policy in the Time of COVID-19: A Primer for Developing Countries,” World Bank Research and Policy Brief 147291.

LOCHNER, L. AND A. MONGE-NARANJO (2012): “Credit constraints in education,” *Annu. Rev. Econ.*, 4, 225–256.

LOCHNER, L., T. STINEBRICKNER, AND U. SULEYMANOGLU (2021): “Parental Support, Savings, and Student Loan Repayment,” *American Economic Journal: Economic Policy*, 13, 329–71.

LU, F. AND T. VOGL (2023): “Intergenerational Persistence in Child Mortality,” *American Economic Review: Insights*, 5, 93–109.

LUCAS, A. M. (2010): “Malaria Eradication and Educational Attainment: Evidence from Paraguay and Sri Lanka,” *American Economic Journal: Applied Economics*, 2, 46–71.

LUO, M. AND S. MONGEY (2019): “Assets and Job Choice: Student Debt, Wages and Amenities,” NBER Working Paper No. 25801.

MA, L., G. SHAPIRA, D. DE WALQUE, Q.-T. DO, J. FRIEDMAN, AND A. A. LEVCHENKO (2021): “The Intergenerational Mortality Tradeoff of COVID-19 Lockdown Policies,” Tech. rep., National Bureau of Economic Research.

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

MICHELACCI, C. AND H. RUFFO (2015): “Optimal Life Cycle Unemployment Insurance,” *American Economic Review*, 105, 816–59.

MINCER, J. (1974): *Schooling, Experience, and Earnings*, National Bureau of Economic Research.

MINNESOTA POPULATION CENTER (2020): *Integrated Public Use Microdata Series, International: Version 7.3 [Machine-readable database]*, Minneapolis: University of Minnesota.

MLACHA, C. J. AND M. O. NDANSHAU (2018): “Education and Labour Earnings Inequality in Tanzania: Evidence from Quantile Regression Analysis,” Unpublished Manuscript, University of Dar es Salaam.

Mossong, J., N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska, and J. Edmund (2008): "Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases," *PLoS Med*, 5, e74.

National Malaria Control Programme (2010): *Tanzania Malaria Programme Review 2010: Programme Review Proposal*, Dar es Salaam United Republic of Tanzania NMCP.

Ogbo, F. A., K. E. Osita, O. A. Akorede, K. I. Ifegwu, T. Lawrence, J. Emmanuel, C. Deborah, and E. A. Kingsley (2019): "Determinants of trends in neonatal, post-neonatal, infant, child and under-five mortalities in Tanzania from 2004 to 2016," *BMC Public Health*.

Prin, M. and H. Wunsch (2012): "International Comparisons of Intensive Care: Informing Outcomes and Improving Standards," *Current Opinion in Critical Care*, 18, 700.

Ramachandran, S. and A. Malani (2021): "All-Cause Mortality During SARS-CoV-2 Pandemic in India: Nationally-Representative Estimates Independent of Official Death Registry," Unpublished Working Paper, University of Toronto.

Rothstein, J. and C. E. Rouse (2011): "Constrained after college: Student loans and early-career occupational choices," *Journal of Public Economics*, 95, 149–163.

Saez, E., J. Slemrod, and S. H. Giertz (2012): "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature*, 50, 3–50.

Sanchez, J. M. (2021): "COVID-19's Economic Impact around the World," Federal Reserve Bank of St. Louis Regional Economist.

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

Shastry, G. K. and D. N. Weil (2003): "How Much of Cross-Country Income Variation is Explained by Health?" *Journal of the European Economic Association*, 1, 387–396.

Shih, H.-H. and M.-J. Lin (2018): "Long-Term Impacts of Early-Life Exposure to Malaria: Evidence from Taiwan's Eradication Campaign in the 1950s," *Health economics*, 27, 1484–1512.

Sicuri, E., F. Yaya Bocoum, J. Nonvignon, S. Alonso, B. Fakih, G. Bonsu, S. Kariuki, O. Leeuwenkamp, K. Munguambe, M. Mrisho, V. Were, and C. Sauboin (2019): "The Costs of Implementing Vaccination with the RTS, S Malaria Vaccine in Five

Sub-Saharan African Countries," *MDM Policy & Practice*, 4, 2381468319896280.

SOARES, R. R. (2005): "Mortality Reductions, Educational Attainment, and Fertility Choice," *American Economic Review*, 95, 580–601.

VENKATARAMANI, A. S. (2012): "Early Life Exposure to Malaria and Cognition in Adulthood: Evidence From Mexico," *Journal of Health Economics*, 31, 767–780.

VON CARNAP, T., I. ALMÅS, T. BOLD, S. GHISOLFI, AND J. SANDEFUR (2020): "The Macroeconomics of Pandemics in Developing Countries: an Application to Uganda," *COVID Economics Vetted and Real-Time Papers Issue 27, CEPR*, 104–122.

WEIL, D. N. (2007): "Accounting for the Effect of Health on Economic Growth," *The Quarterly Journal of Economics*, 122, 1265–1306.

WILDE, J., B. APOUEY, J. COLEMAN, AND G. PICONE (2019): "The Effect of Antimalarial Campaigns on Child Mortality and Fertility in Sub-Saharan Africa," Unpublished Manuscript, Max Planck Institute for Demographic Research.

WORLD HEALTH ORGANIZATION (2021): "World Malaria Report 2020: 20 Years of Global Progress and Challenges," *www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2020*.

YAMALL ORELLANA, J. D., G. M. DA CUNHA, L. MARRERO, R. ISMERIO MOREIRA, L. DA COSTA LEITE, AND B. LESSA HORTA (2021): "Excess Deaths During the COVID-19 Pandemic: Underreporting and Regional Inequalities in Brazil," *Cadernos de Saúde Pública*, 37.

YOUNGER, S. D., F. MYAMBA, AND K. MDADILA (2016): "Fiscal Incidence in Tanzania," *African Development Review*, 28, 264–276.

YUM, M. (2023): "Parental time investment and intergenerational mobility," *International Economic Review*, 64, 187–223.

ZHOU, A. (2022): "Building Future Generations: The Macroeconomic Consequences of Family Policies," Unpublished Manuscript, University of Hong Kong.