*American Economic Review 2022, 112(8): 2774–2814*
*https://doi.org/10.1257/aer.20201157*

# The Human Side of Structural Transformation†

*By* Tommaso Porzio, Federico Rossi, and Gabriella Santangelo*

*We document that nearly half of the global decline in agricultural employment was driven by new cohorts entering the labor market. A new dataset of policy reforms supports an interpretation of these cohort effects as human capital. Using a model of frictional labor reallocation, we conclude that human capital growth led to a sharp decline in the agricultural labor supply, accounting, at fixed prices, for 40 percent of the decrease in agricultural employment. This aggregate effect is halved in general equilibrium and it reflects the role of human capital as both a mediating factor and an independent driver of labor reallocation.* (*JEL* J22, J24, J43, L16, O13, O14, Q10)

Over the last two centuries, economic development has typically been accompanied by a process of structural transformation: as countries grow richer, workers reallocate from agriculture to manufacturing and services. The literature has emphasized two mechanisms to explain this pattern: a decrease in the relative demand of agricultural goods driven by income effects and an increase in the relative productivity of the agricultural sector (Herrendorf, Rogerson, and Valentinyi 2014). Both these forces amount to a shift in the relative demand for agricultural labor: keeping fixed their characteristics, workers are progressively more needed in the nonagricultural sector.

This paper argues that the labor force itself has been subject to a radical transformation, which has contributed significantly to the global decline in agricultural employment by reducing the supply of agricultural labor. Using data from a broad sample of countries at different levels of development, we show that more recent cohorts have a comparative advantage towards the nonagricultural sector. We document that a key reason behind this pattern is the secular increase in human capital, which is relatively more valuable out of agriculture. We conclude that human capital growth played a key role in the process of structural transformation; a mechanism

*Porzio: Columbia University, CEPR, and NBER (email: tommaso.porzio@columbia.edu); Rossi: University of Warwick (email: federico.rossi@warwick.ac.uk); Santangelo: University of Cambridge; (email: gabriella.santangelo@econ.cam.ac.uk). Emi Nakamura was the coeditor for this article. We thank three anonymous referees for excellent comments that improved the paper. We also thank Rachel Ngai for an insightful discussion of an early version of the paper and for helpful comments Andrew Atkeson, Martin Beraja, Francesco Caselli, Oded Galor, Joe Kaboski, David Lagakos, Tim Lee, Jonathan Heathcote, Ben Moll, Simon Mongey, Karthik Muralidharan, Michael Peters, Todd Schoellman, Jonathan Vogel. Finally, we have benefited from the reactions of several seminar and conference audiences, including participants at the CEPR MG Annual Programme Meeting, Barcelona GSE Summer Forum, Brown, MIT, UBC, and UCLA. Angus Lewis and Xiao Ma provided excellent research assistance.
†Go to https://doi.org/10.1257/aer.20201157 to visit the article page for additional materials and author disclosure statements.

first proposed by Caselli and Coleman II (2001) in the US context, which we bring under empirical scrutiny on a global scale.

The analysis builds on an empirical cross-cohort study of labor reallocation, which we then interpret with a two-sector general equilibrium model of structural transformation. Our starting point is that the aggregate labor reallocation can be expressed as the product of two distinct forces: within-cohort reallocation, i.e., individuals in a given cohort progressively leaving agriculture over time, and between-cohort reallocation, i.e., more recent cohorts having a lower share of agricultural employment at a given point in time. The key observation is that the relative role of these two forces can be used to shed light on the underlying drivers of structural transformation. Intuitively, while changes in the demand for agricultural labor should affect different cohorts similarly, supply shifts due to changes in cohort-level characteristics—such as the quantity and quality of education—should be reflected in differences across cohorts. Building on this insight, we show that the cross-cohort variation in sectoral employment at a given point in time can be used to identify the extent to which the supply of agricultural labor has shifted over time. Through the lens of our model, we then quantify the implications of such shift and the resulting general equilibrium effects for the aggregate rate of structural transformation.

We start by formalizing a statistical decomposition of the observed aggregate rate of labor reallocation into changes across and within cohorts. We use repeated cross-sections of micro-level data for 69 countries around the world, covering two thirds of the world population and a large part of the income distribution. We run cohort-level regressions of log agricultural employment on year and cohort dummies and calculate the extent to which changes in the estimated cohort effects for the active cohorts can account for the aggregate rate of reallocation. Naturally, part of the differences in agricultural employment across cohorts might reflect factors associated with age as opposed to fixed cohort-level characteristics; for example, mobility barriers that are likely to be more relevant for older workers. To account for this, we consider a version of our decomposition exercise which separately controls for (properly restricted) age and cohort effects.

We find a substantial role for labor reallocation across cohorts. When age effects are controlled for, changes in cohort effects can account on average for about 40 percent of the overall reallocation out of agriculture. While there is some heterogeneity across countries, the contribution of cohort effects is substantial in the overwhelming majority of them. Overall, our results point towards the importance of cross-cohort changes in workers' characteristics for labor reallocation out of agriculture.

We provide several pieces of evidence that support an interpretation of cohort effects as shifts in human capital. As a starting point, we show that, within each country, cohort-level schooling is negatively associated with the estimated cohort effects, controlling for a country-specific quadratic trend. In other words, within a given labor market, cohorts with higher educational attainment are disproportionately less likely to work in agriculture. We then explore several sources of cross-cohort variation in schooling to support a causal link between educational attainment and sectoral choices. First, we compile a novel dataset on educational reforms and political events (such as independence and democratic transitions) in the countries in our sample. We find that cohorts exposed to reforms or events that increased their schooling during youth have lower agricultural employment during adulthood. Second, we follow the identification strategy in Duflo (2001), and exploit a school construction program in Indonesia as a shock to educational attainment: the cohorts more affected by the program are less likely to be employed in agriculture.

We interpret the empirical results using a general equilibrium model of frictional labor reallocation. The model has an overlapping generations structure, with new cohorts replacing older ones which retire. In each period of their life, workers choose whether to work in agriculture or nonagriculture, subject to mobility frictions. Human capital varies across cohorts, due to both the endogenous response to changes in the relative returns from working in nonagriculture, where human capital is relatively more valued, and an exogenous shifter capturing all other forces that affect the costs and returns of human capital investments. The drivers of structural transformation traditionally emphasized in the literature (changes in relative agricultural productivity and in the demand for agricultural goods due to income effects) generate a decrease over time in the demand for agricultural labor. In addition, human capital growth contributes to labor reallocation by decreasing the agricultural labor supply.

The model provides a structural interpretation of the decomposition exercise. First, the theory guides us in the selection of the restriction that we need to separately identify cohort, year, and age effects. Mobility costs affect the level of agricultural employment for all cohorts, but—as long as they are not too large—not the rate of reallocation for relatively young cohorts; as a consequence, restricting the age effects to be identical in the first few periods that a cohort is active allows us to identify cohort and year effects. Moreover, the estimated cohort effects measure the cohort-level average human capital, and the change over time in the average of the estimated cohort effects captures the overall shift in agricultural labor supply driven by both endogenous and exogenous human capital growth. Year dummies absorb changes in the demand for agricultural labor, while age controls capture the effect of reallocation frictions.

Through the lens of the model, the decomposition results imply a sharp decline in the agricultural labor supply over time, which accounts for almost 40 percent of average labor reallocation out of agriculture, if we keep prices fixed. This accounting result implies that human capital growth, either as a mediating factor or as an independent driver, represents an important ingredient of structural transformation.

We then move beyond accounting and quantify how much labor reallocation could be generated by the shift in the supply of agricultural labor, once we take into account equilibrium adjustments in relative prices and wages. The general equilibrium impact of the supply shift is mediated by a general equilibrium multiplier, which is a combination of the model's parameters that controls the responsiveness of relative prices and wages. We consider two alternative approaches to quantify the multiplier: calibration and a regression-based exercise exploiting the cross-country variation in the estimated cohort effects and labor reallocation. In both cases, we conclude that general equilibrium forces roughly halve the partial equilibrium impact of human capital growth.

Finally, we rely on a natural experiment to discipline the relative role of endogenous and exogenous changes in human capital. Building on Gollin, Hansen, and Wingender (2021), we show that countries more exposed to the increase in agricultural productivity growth induced by the Green Revolution saw both faster labor reallocation out of agriculture and faster schooling growth for the affected cohorts, consistently with an endogenous adjustment of human capital growth to the demand forces behind structural transformation. Interpreted through our model, these estimates imply that approximately half of the observed global human capital growth was due to forces exogenous to structural transformation, while the rest was an endogenous response to the change in the relative value of human capital.

Overall, our results show that human capital accumulation dramatically transformed the labor force, shifting labor supply away from agriculture. This shift contributed in a quantitatively important way to the reallocation of employment across sectors. Based on this, we conclude that any credible quantitative analysis of structural transformation cannot fail to consider—as mostly done in the literature so far—its "human" side.

*Related Literature*.—We build on the work of Caselli and Coleman II (2001) and Acemoglu and Guerrieri (2008). To our knowledge, Caselli and Coleman II (2001) first argued that the supply of agricultural workers might be relevant to understand structural change. Acemoglu and Guerrieri (2008) build on an insight first proposed by Rybczynski (1955) and formalize the notion that changes in the supply of different inputs may lead to structural transformation if sectors vary in the intensity with which they use them. Our contribution is to develop and apply a methodology to measure changes in the supply of agricultural workers for many countries, link them to changes in schooling, and quantify their aggregate impact. In this sense, we add to a literature studying the quantitative role of changes in the demand for agricultural labor, driven by preferences or technology (Alvarez-Cuadrado and Poschke 2011; Boppart 2014; Comin, Lashkari, and Mestieri 2021).

More broadly, our work is related to a large literature on the contribution of human capital to industrialization and development. This literature argues that human capital is particularly useful in the modern and technologically advanced industrial sector, and that the expansion of formal education is a key ingredient of the transition from stagnation to growth (see Nelson and Phelps 1966; Galor 2005).[1] We contribute to this body of work by isolating empirically the effects of human capital growth across cohorts on the supply of agricultural workers and the reallocation of labor out of agriculture. Our cross-cohort analysis quantifies the role of human capital without relying on proxies based on wages or years of schooling, in line with growing evidence that these proxies miss a significant part of the variation in human capital across countries and over time (see Rossi 2020, for a review).

Our model combines elements and insights already present in Matsuyama (1992b); Lucas (2004); Galor, Moav, and Vollrath (2009); and more recently Herrendorf and Schoellman (2018); and Bryan and Morten (2019). We provide a tractable framework to analytically characterize labor reallocation by cohort in the presence of mobility frictions, which have been shown to affect significantly agricultural workers in developing countries (Ngai, Pissarides, and Wang 2019). Hsieh et al. (2019) also exploit year and cohort effects to calibrate a model of the allocation of talent;

[1] A large strand of this literature studies empirically the effect of educational attainment on income per capita; see for example Barro (1991); Mankiw, Romer, and Weil (1992); Glaeser et al. (2004); Gennaioli et al. (2013); and Valencia Caicedo (2018).

compared to their work, we focus on a simpler framework that allows us to analytically consider fixed-cost-type frictions, which turn out to be crucial to correctly identify the role of changes in the supply of agricultural workers. In emphasizing the importance of comparative advantage, our work also relates to Lagakos and Waugh (2013); Young (2013); and Nakamura, Sigurdsson, and Steinsson (2022).

Finally, with respect to the aim of separating the role of labor demand and supply as drivers of sectoral shifts, our paper is closely related to the work of Lee and Wolpin (2006), which devises and structurally estimates a rich model of the process of labor reallocation from manufacturing to services in the United States.[2] We study a conceptually similar question, though in a different context (the transition out of agriculture along the process of development). Moreover, we tackle it from a radically different perspective, imposing the minimal possible structure to interpret patterns of reallocation by cohort. The combination of cohort-level evidence and a model capturing general equilibrium effects makes our work related to a growing literature exploiting micro-level variation to discipline macroeconomic models (Nakamura and Steinsson 2018).

*Structure of the Paper*.—The paper is organized as follows. Section I describes the data, while Section II lays out the basic statistical decomposition of aggregate labor reallocation into cohort and year effects. In Section III we provide several pieces of evidence on the relationship between schooling and the estimated cohort effects. Section IV presents the model, which is used in Section V to unpack the aggregate labor reallocation into shifts of demand and supply of agricultural labor. Section VI includes further results on the equilibrium effects and sources of human capital growth. Section VII concludes.

## I. Data

We use data from the Integrated Public Use Microdata Series (IPUMS, see King et al. 2019). Our main source is IPUMS-International (henceforth, IPUMS-I), which includes censuses or labor force surveys that are representative of the entire population. To improve the coverage of the poorest countries in the world, we also use data from IPUMS-Global Health (henceforth, IPUMS-GH), which harmonizes the demographic health surveys (Boyle, King, and Sobek 2019), a collection of surveys focused on health variables that include information on agricultural employment.

For our benchmark analysis, we include all countries for which we have two or more repeated cross-sections spanning at least ten years, with available information on industry of employment for men aged 25 to 59.[3] We focus on this age range to capture working-age individuals with completed education, and exclude women from the analysis given that their low labor force participation in many countries

[2] Cociuba and MacGee (2018) also study frictional workers reallocation across sectors. However, they focus on business cycle frictions and on how demographics interact with sectoral reallocation. Their stationary framework would not be suitable to study the long-run trends in labor reallocation out of agricultural that we study.

[3] We exclude cross-sections for which information on industry is missing (which is always the case for the not employed) for more than 25 percent of men aged 35 to 45. Figure A.1 shows that this restriction excludes only very few cross-sections. All the figures and tables labeled "A" are included in the online Appendix.

makes it difficult to properly compute the cohort-level reallocation across sectors.[4] This gives us a sample of 58 countries and 241 cross-sections, covering more than two-thirds of the world population, five continents and most of the income distribution. The IPUMS-I data include 52 countries, of which 9 are high-income, 24 are middle-income and 19 are low-income; all the 12 countries in the IPUMS-GH data are low-income (some countries are in both IPUMS-I and IPUMS-GH).[5] On average, we observe countries over a period of 27 years in the IPUMS-I data and of 15 years in the PUMS-GH data. For robustness, we also consider an extended sample including all countries with cross-sections spanning at least five years and industry information for men aged 25 to 54; this gives us two more middle-income countries in the IPUMS-I data and 13 more low-income countries in the IPUMS-GH data, for a total of 69 countries and 285 cross-sections (see online Appendix A for a full list).

Our key variable of interest is agricultural employment at the cohort level. We use the variables *indgen* (IPUMS-I) and *wkcurrjob* (IPUMS-GH), which are harmonized across countries and time periods, to compute the share (properly weighted) of the male population employed in the industry "Agriculture, fishing and forestry."[6] Figure A.3a shows, for each country, the average number of observations at the cohort $\times$ year level. For almost all countries in IPUMS-I, we have at least 1000 observations per cell. Sample sizes in the IPUMS-GH data are much smaller. For this reason, we use the 52 countries in IPUMS-I as our core sample, and report results from the IPUMS-GH as robustness checks.[7]

We subject our data to three consistency checks. First, we inspect visually, for all countries, the growth rates in aggregate agricultural employment between cross-sections, searching for anomalies. This procedure leads us to exclude one observation from the IPUMS-I data, and two from the IPUMS-GH.[8] Second, we inspect visually the cross-sectional relationships between agricultural employment and birth year. We exclude ten cross-sections from the IPUMS-GH data that display very large swings across birth cohorts, casting doubts on data reliability.[9] Finally, we verify that the average agricultural employment computed in our final sample is comparable in magnitude with aggregate data from the World Development Indicators (see Figure A.5), a commonly used data source (Herrendorf, Rogerson, and Valentinyi 2014).[10]

[4] As Figures A.2a and A.2b show, the average employment rate of men aged 25–59 is high and constant.

[5] By high-income (low-income) countries we mean those with GDP per capita greater (smaller) than 45 percent (10 percent) of the one of the United States at purchasing power parity (PPP) in 2000. We use GDP per capita from the Maddison Project Database (Inklaar et al. 2018) to have a better coverage of historical data. Data for Fiji is missing; we assign it to the low-income countries. Puerto Rico is a territory, but we label it a country.

[6] In computing the agricultural employment share, we do not restrict the sample to individuals in the labor force. As we highlight in Figure A.2c, we do not want to confound entry into the labor force with reallocation out of agriculture. We consider alternatives for robustness.

[7] In several countries, we observe age heaping. We use a standard procedure, illustrated in Figure A.4, to get a smooth distribution of agricultural employment as a function of age.

[8] We drop the reallocation between 2000 and 2005 for the United States (which corresponds to a change in the sectoral classification) and the reallocation between 2015 and 2016 and between 2016 and 2017 for Senegal. The details are in Figure A.6 and notes.

[9] For most countries, the first available cross-section from the IPUMS-GH data is extremely noisy. Cote d'Ivoire has only two cross-sections, hence excluding the first one leads us to exclude Cote d'Ivoire altogether. The plots of all the omitted cross-sections are in Figure A.7.

[10] There are, nonetheless, a few discrepancies, which we discuss in Figure A.5.

## II. Decomposing Structural Change

We study patterns of labor reallocation out of agriculture by birth cohort. While most of the existing work focuses on aggregate rates of reallocation, we are among the first to systematically document micro-level evidence on the behavior of different cohorts in the process of structural transformation.[11]

### A. *Cohort and Year Components of Labor Reallocation*

In each country $j$, for each cross section $t$, and for each cohort $c$, we compute the share of the population in agriculture, $l_{A,t,c,j}$. We normalize $c$ to be equal to the birth year plus 25, so that a cohort first enters into our dataset when $c = t$ and is last in the dataset when $c = t + N$, where $N = 59 - 25 = 34$. The overall share of the population employed in agriculture is given by

$$L_{A,t,j} = \sum_{c=t-N}^{t} n_{t,c,j} l_{A,t,c,j},$$

where $n_{t,c,j}$ is the share of the overall male population aged 25 to 59 belonging to cohort $c$. Our first objective is to decompose changes over time in $L_{A,t,j}$ into a component that captures country-wide trends, and a component that captures changes in the composition of the active labor force.

*A Graphical Inspection*.—As an illustration, we regress, separately for high-, middle-, and low-income countries, $\log l_{A,t,c,j}$ on country fixed effects and dummies that take value one for each decade from 1960 to 2010. The dashed lines in panel A of Figure 1 plots the resulting decade effects, normalized to the average agricultural employment share in the sample. The figure shows two well-known facts: (i) high-income countries have lower agricultural employment, and (ii) labor has reallocated away from agriculture. It also shows that the share of agricultural employment declined at a log-linear rate, a feature of the data that we leverage in the model.

Next, we run the same specifications, but adding a full set of birth-year dummies. The solid lines in panel A of Figure 1 show that, when controlling for cohort effects, the estimated decade dummies decline at a much slower rate. The decline in agricultural employment obtained by following a given birth cohort over time is approximately half of the aggregate decline. This is because the aggregate decline is partly driven by compositional changes, as showed in panel B of Figure 1: younger birth cohorts have a lower share of agricultural employment in any given year.

These results highlight that aggregate structural transformation is due to two equally important mechanisms: (i) over time, individuals of all birth cohorts move away from agricultural employment—we call this the *year component* of labor reallocation, since it captures country-wide trends; (ii) younger cohorts that enter the labor market are less likely to be employed in agriculture—we call this the *cohort*

[11] Kim and Topel (1995); Lee and Wolpin (2006); and Perez (2017) document sectoral reallocation by cohort, but limit their focus to, respectively, South Korea, United States and Argentina. In ongoing work, Hobijn, Schoellman, and Vindas (2019) are also using the IPUMS dataset to document patterns on reallocation by cohort.

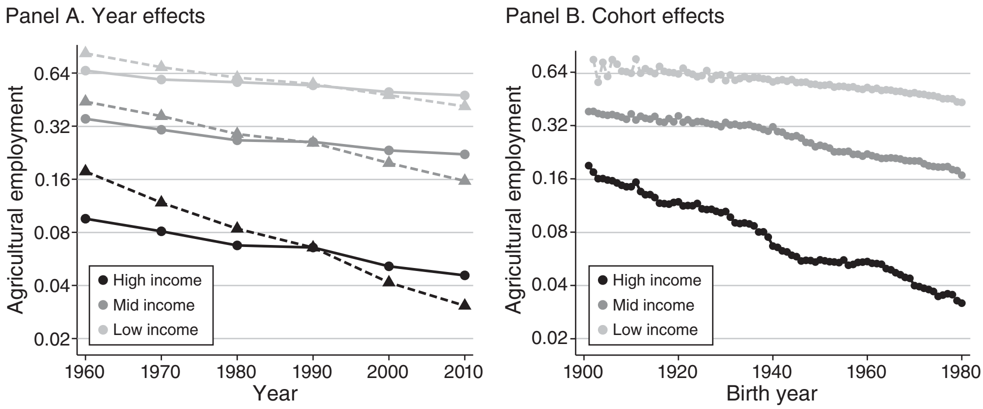


Figure 1. Decomposing Labor Reallocation

*Notes:* Panel A shows the estimates for the year effects conditional on country fixed effects only (triangles and dashed lines) and on country × cohort fixed effects (circles and solid lines). Panel B shows the estimates for the cohort effects, conditional on country × year fixed effects. All estimates are normalized to average to the overall average agricultural employment for each income group. The *y*-axis is on a log scale.

*component* of labor reallocation, since it captures changes in the composition of the active labor force.

*Two Examples*.—To further illustrate the role of year and cohort components in driving aggregate reallocation, panels A and B of Figure 2 plot agricultural employment by cohort for two countries. In Brazil, the year component largely drives aggregate reallocation: within each given cohort, a large share of individuals reallocates out of agriculture over time. In Indonesia, the cohort component plays a more important role: there is no systematic within-cohort time trend in agricultural employment, and, in any given year, younger cohorts are less likely to work in agriculture. As younger cohorts enter the labor market and older ones exit, aggregate agricultural employment decreases as a result.

*Formal Decomposition*.—We regress separately for each country the cohort-level agricultural employment on year and cohort effects,

$$\underbrace{\log l_{A,t,c,j}}_{\text{agr share of cohort } c \text{ at time } t} = \underbrace{\mathbf{Y}_{t,j}}_{\text{year effects}} + \underbrace{\mathbf{C}_{c,j}}_{\text{cohort effects}} + \varepsilon_{t,c,j},$$

and use the resulting estimates to unpack the aggregate rate of labor reallocation into year and cohort components.[12] The average yearly rate of labor reallocation between periods $t$ and $t + k_{t,j}$ for country $j$ is

$$\log g_{L_A,t,j} \equiv \frac{1}{k_{t,j}} \log \frac{L_{A,t+k_{t,j},j}}{L_{A,t,j}},$$

[12] We estimate equation (1) in first differences to provide a tight mapping with the model in Section IV.

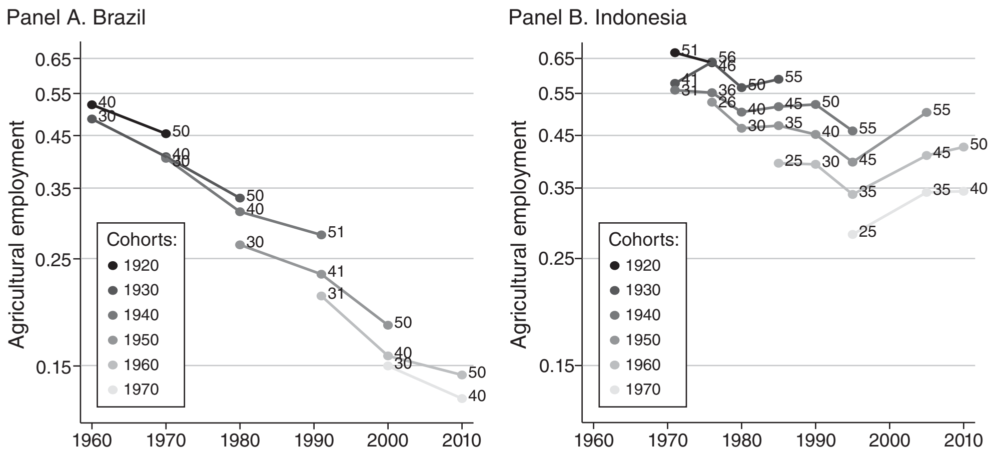


Figure 2. Labor Reallocation by Cohort, Two Examples

*Notes:* The panels plot agricultural employment by birth cohort in Brazil and Indonesia. We follow six birth cohorts aged 25–59, or as long as we observe them in our data. Dots highlight the years in which we observe agricultural employment. The ages of all cohorts in any observed year are reported.

where we define $k_{t,j}$ as the number of years between cross-section $t$ and the next cross-section in our data. We can write $\log g_{L_A,t,j}$ as

$$\underbrace{\log g_{L_A,t,j}}_{\text{rate of labor reallocation}} = \underbrace{\log \psi_{t,j}}_{\text{year component}} + \underbrace{\log \chi_{t,j}}_{\text{cohort component}}, \tag{2}$$

where

$$\log \psi_{t,j} \equiv \frac{1}{k_{t,j}}\left(\mathbf{Y}_{t+k,j} - \mathbf{Y}_{t,j}\right), \tag{3}$$

$$\log \chi_{t,j} \equiv \frac{1}{k_{t,j}} \log\left[\frac{\sum_{c=t+k_{t,j}-N}^{t+k_{t,j}} n_{t+k,c,j} \exp\left(\mathbf{C}_{c,j}\right)}{\sum_{c=t-N}^{t} n_{t,c,j} \exp\left(\mathbf{C}_{c,j}\right)}\right] = \log g_{L_A,t,j} - \log \psi_{t,j}. \tag{4}$$

The year component $\log \psi_{t,j}$ is the difference between the year effects at time $t$ and $t + k_{t,j}$, while the cohort component $\log \chi_{t,j}$ captures changes in the average cohort effects of the active cohorts. We compute $\log \psi_{t,j}$ and $\log \chi_{t,j}$ for each pair of cross-sections and calculate their average as

$$\log \psi_j = \frac{1}{|\mathbf{T}_j|} \sum_{t \in \mathbf{T}_j} \log \psi_{t,j},$$

$$\log \chi_j = \frac{1}{|\mathbf{T}_j|} \sum_{t \in \mathbf{T}_j} \log \chi_{t,j},$$

where $\mathbf{T}_j$ is the set of all cross-sections available for country $j$ excluding the most recent one, for which we cannot calculate the reallocation rate. The decomposition of the average reallocation rate between $\log \psi_j$ and $\log \chi_j$ summarizes the patterns of reallocation by cohort shown, for example, in Figure 2: the year component $\log \psi_j$

Table 1—Unpacking Structural Change

| Country type | $\log g_{L_A}$ (1) | $\log \psi$ (2) | $\frac{\log \chi}{\log g_{L_A}}$ (3) | $\log \tilde{\psi}$ (4) | $\frac{\log \tilde{\chi}}{\log g_{L_A}}$ (5) | $1 - \frac{\log \psi}{\log \tilde{\psi}}$ (6) | Obs. (7) |
|---|---|---|---|---|---|---|---|
| All | −2.11 | −0.92 | 0.56 | −1.32 | 0.38 | 0.30 | 52 |
| High income | −3.41 | −1.39 | 0.59 | −1.38 | 0.59 | -0.01 | 9 |
| Middle income | −2.19 | −0.97 | 0.56 | −1.60 | 0.27 | 0.40 | 24 |
| Low income | −1.41 | −0.64 | 0.55 | −0.93 | 0.34 | 0.31 | 19 |

is the average slope of the cohorts' paths, while the cohort component $\log \chi_j$ is the average vertical gap across cohorts, properly annualized.

Columns 1–3 of Table 1 summarize the decomposition results. Across all countries, the agricultural employment declines on average by 2.11 percent each year, of which 0.92 percent is due to the year component. Therefore, as showed in column 3, 56 percent of the aggregate reallocation is due to the cohort component; this contribution is similar across all income groups. Online Appendix B presents the results for each country in our sample, showing that the cohort component is negative in the vast majority of them.

## B. *Controlling for Age*

The statistical decomposition considered so far restricts age to have no effect on the cohort-level agricultural share. However, older workers plausibly face stronger barriers to reallocate across sectors, limiting their labor mobility over time. In the absence of age controls, this may contribute to a large role of cohort effects for the aggregate rate of labor reallocation. Therefore, we next include age controls in the previous decomposition.

It is well known that year, cohort and age are collinear, and can be separately identified only if an additional linear restriction is imposed.[13] Our restriction is that age has no effect in the first few years a cohort is employed. This choice is guided by theory, and will be fully motivated in the context of our model in Section V. Intuitively, this amounts to assuming that frictions to labor reallocation affect equally consecutive cohorts at the beginning of their working career.

In the implementation of this idea, we face a trade-off between the parametrization of age effects and the sample size for the identification of year effects. At one extreme, a specification including a full set of age dummies, with the coefficients on the first two restricted to be equal to each other, would identify year effects out of the reallocation behavior of one cohort only. To strike a balance between the two sides of this trade-off, we follow Card, Heining, and Kline (2013) and include quadratic and cubic terms for age, centered around a value $\bar{a}$ to be specified below. Separately for each country $j$, we run

$$\log l_{A,t,c,j} = \tilde{\mathbf{Y}}_{t,j} + \tilde{\mathbf{C}}_{c,j} + \beta_{1,j}(a_{c,t,j} - \bar{a})^2 + \beta_{2,j}(a_{c,t,j} - \bar{a})^3 + \varepsilon_{t,c,j}, \tag{5}$$

[13] See Deaton (1997), and more recently Lagakos et al. (2018).

where $\tilde{\mathbf{Y}}_{t,j}$ and $\tilde{\mathbf{C}}_{c,j}$ denote year and cohort dummies, and $a_{c,t,j}$ is the age of cohort $c$ at time $t$ (for country $j$). This specification restricts age effects to be 0 at age $\bar{a}$, both in levels and in changes.[14] Since our data come from repeated cross-sections that are on average (across all countries and time periods) 8.8 years apart, we set $\bar{a} = 29.4$, i.e., the average age of the youngest cohort that we observe for at least two successive cross-sections. We explore several alternative specifications for age effects in Table B.4, including country-specific values for $\bar{a}$, more flexible age dummies, and time-varying age controls; the results are similar to those shown below.

Given the estimates from specification (5), we compute, just as in Section IIA, the annualized year and cohort components

$$\log \widetilde{\psi}_{t,j} \equiv \frac{1}{k_{t,j}}\Big(\widetilde{\mathbf{Y}}_{t+k,j} - \widetilde{\mathbf{Y}}_{t,j}\Big),$$

$$\log \widetilde{\chi}_{t,j} \equiv \log g_{L_A,t,j} - \log \widetilde{\psi}_{t,j},$$

and take their average across all available cross-sections, $\log \widetilde{\psi}_j$ and $\log \widetilde{\chi}_j$.

Figure 3 plots $\log \widetilde{\psi}_j$ as a function of $\log \psi_j$. Controlling for age matters: almost all countries lie below the 45-degree line, which means that the year components estimated with age controls is larger (in absolute value). Younger birth cohorts, which are less likely to be constrained by mobility frictions, reallocate across sectors at a faster rate. However, even conditional on age, the cohort component of aggregate reallocation is still substantial. Column 4 of Table 1 shows that the average year component $\log \widetilde{\psi}_j$ is $-1.3\%$, which implies that the cohort component still explains almost 40 percent of the total reallocation out of agriculture, as shown in column 5.[15] Table 1 and Figure 3 also show that controlling for age effects has a larger impact on the estimated cohort components for middle- and low-income countries. We will return to this result in Section V, where we show that the ratio between $\log \widetilde{\psi}_j$ and $\log \psi_j$ directly maps into the structural parameters modulating mobility frictions.

## III. Understanding Cohort Effects: Evidence from Schooling Data

We have shown that cohort effects explain a large share of aggregate labor reallocation. Given the same aggregate conditions, younger birth cohorts are less likely to work in agriculture: the data reveal that they have a comparative advantage for nonagriculture. What determines the shift across cohorts in comparative advantage? This section provides several pieces of evidence to support the interpretation of cohort effects as shifts in human capital.

[14] In fact, the omission of a linear term for age is necessary to have the derivative of the age terms to be zero at $\bar{a}$, which is needed for identification of the year trend.

[15] Table B.4 shows that the results are similar when focusing on the IPUMS-GH sample. Moreover, online Appendix B shows that both specifications (1) and (5) capture the vast majority of the variation in the data.

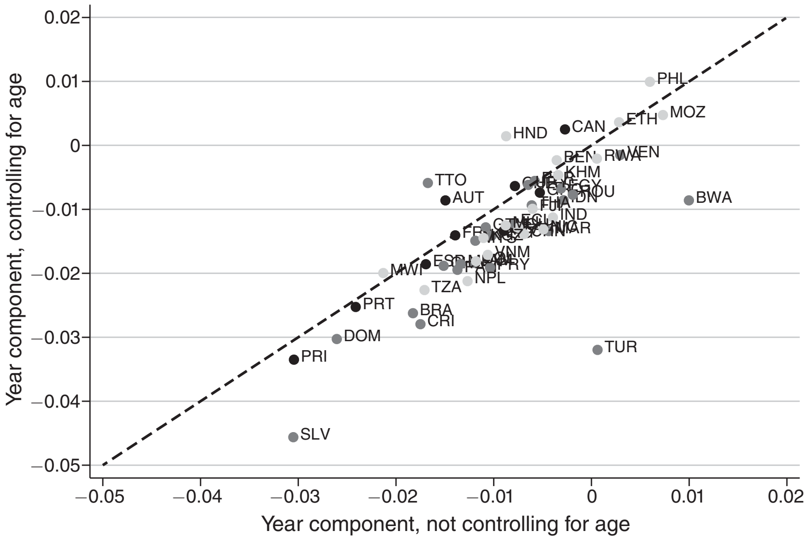


FIGURE 3. THE EFFECT OF AGE CONTROLS

*Notes:* The figure plots the year component estimated with age controls, $\log \widetilde{\psi}_j$, as a function of the year component estimated without age controls, $\log \psi_j$. The markers are black for high income countries, gray for middle income countries, and light gray for the low income countries. The 45-degree line shows that in most countries $\log \widetilde{\psi}_j$ is larger than $\log \psi_j$ in absolute value.

## A. *Correlation between Schooling and Cohort Effects*

We start by documenting that the cohort effects estimated in Section III are correlated with cohort-level educational attainment. We use individual-level educational attainment to compute the average schooling years for each cohort in our dataset. Since we observe cohorts in multiple cross-sections, we extract average schooling by cohort using, separately for each country, a procedure similar to the one used in DeLong, Goldin, and Katz (2003) for the United States. More specifically, we project the log of cohort-level average schooling years on a full set of cohort dummies and a cubic polynomial in age, which controls for late enrollment in school (i.e., after 25 years of age) and, especially, mortality differences by education groups. We transform the estimated cohort dummies in levels, and denote the resulting schooling level for cohort $c$ in country $j$ as $s_{c,j}$.

As a first step, panel A of Figure 4 replicates panel B of Figure 1, but using schooling rather than agricultural employment. The relationship between years of education and birth cohorts mirrors the one for agricultural employment, suggesting that the schooling increase might have played a role in shaping the comparative advantage of younger generations. At the same time, the comparison might be confounded by several factors; most obviously, similar time trends in both variables.

To make progress, we estimate specifications that control for quadratic time trends. We run, separately for each country,

$$\tilde{\mathbf{C}}_{c,j} = \alpha_j + \beta_j s_{c,j} + \delta_{1,j}(c - \bar{c}_j) + \delta_{2,j}(c - \bar{c}_j)^2 + \varepsilon_{c,j}, \tag{6}$$

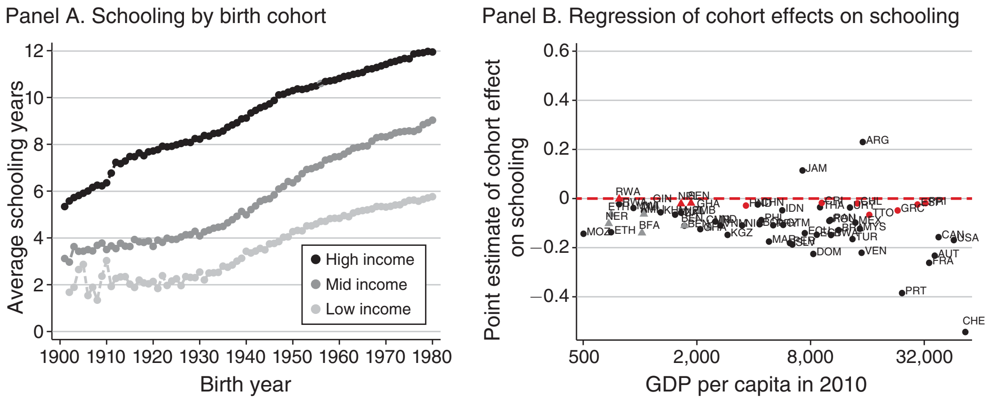


FIGURE 4. EDUCATIONAL ATTAINMENT AND AGRICULTURAL EMPLOYMENT

*Notes:* Panel A replicates Figure 1 (panel B) using cohort-level schooling rather than agricultural employment. Panel B plots for each country the point estimate of $\hat{\beta}_j$ from specification (6). Black circles and gray triangles are for IPUMS-I and IPUMS-GH countries for which $\hat{\beta}_j$ are negative and significant at 5 percent. Observations in red are not significantly different from 0.

where $\tilde{\mathbf{C}}_{c,j}$ is the cohort effect estimated in (5), and $\bar{c}_j$ is the first cohort that we observe in each country. The coefficient of interest is $\beta_j$; panel B of Figure 4 plots it as a function of GDP. For almost all countries, the coefficient is negative: cohorts that are more educated, relative to a country-specific quadratic trend, are less likely to work in agriculture. While the coefficient $\beta_j$ is negative and significant in almost all countries, there is some heterogeneity; in particular, one extra year of schooling in rich countries appears to have a larger effect on agricultural employment.[16] To focus on one magnitude, we run specification (6) pooling all countries together, allowing for country-specific time trends. Column 1 of Table 2 reports the results: one additional year of schooling decreases cohort-level agricultural employment by approximately 10 percent relative to what it would have been otherwise. The result is robust to the inclusion of decade of birth dummies interacted by income group (column 2).

We should be cautious in interpreting this relationship as causal. If educational decisions are forward-looking, changes in schooling might be ultimately driven by the anticipation of higher demand for nonagricultural labor (relative to the quadratic trend). Moreover, cohort-level average schooling might be correlated with other cohort-level characteristics that affect sectoral choices, such as average early-childhood human capital investments, ability or preferences.[17] While most of these possibilities are broadly consistent with the core thesis of this paper (i.e., human capital being an important driver of the supply of agricultural labor) establishing whether schooling plays an independent and direct role is important, as

[16] These patterns hold for IPUMS-GH countries as well, identified by triangles in Figure 4, panel B.

[17] Notice that these concerns are different (and, arguably, less severe) compared to those typically faced by individual-level analyses of returns to education. In particular, individual-level selection in terms of omitted characteristics is not problematic per se, as long as the cohort-level distribution of these characteristics does not vary over time. Moreover, recall that cohort effects are estimated conditional on year effects, therefore controlling for aggregate economic conditions.

Table 2—Role of Schooling

| | Cohort effect | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Cohort schooling | −0.100 | −0.108 | −0.169 | −0.157 |
| | (0.004) | (0.004) | (0.038) | (0.031) |
| Country trend | Yes | Yes | Yes | Yes |
| Country fixed effects | Yes | Yes | Yes | Yes |
| Birth-year controls | No | Yes | No | Yes |
| Method | OLS | OLS | IV | IV |
| Instrument | – | – | Ed. reforms | Ed. reforms |
| *F*-stat first stage | – | – | 2.93 | 2.40 |
| Observations | 3,238 | 3,238 | 907 | 907 |

*Notes:* Robust standard errors in parentheses. The country trends include both linear and quadratic terms. Birth-year controls are a full set of decade of birth dummies interacted with income group dummies.

formal educational attainment is, at least partially, under the direct influence of educational policy. In the rest of this section, we present several approaches to make progress in this direction.

### B. *Educational Reforms and Political Events*

As a first exercise, we focus on the cross-cohort variation in educational attainment induced by the timing of large country-wide shocks to the educational system. We compile a novel dataset of educational reforms and political events for the countries in our sample. For each country, we conduct an extensive search using all available online sources: academic papers, encyclopedias, newspaper articles, and blogs. Details of our approach to data construction are included in online Appendix A. We find a total of 39 policy reforms extending compulsory education, and 86 political events such as independence from colonial powers, transitions to democracy and wars that plausibly impacted (either positively or negatively) the working of the educational system or the costs of acquiring education. This new dataset covers the majority of our countries: we find at least one policy reform in 23 of our sample countries, and at least one political event in 38 of them. We further use historical sources to identify educational reforms that were either not fully implemented due to low state capacity, limited to some regions or phased in slowly over time; 15 out of 39 reforms fall within this category (we refer to them as "weakly implemented").

*An Example: Mozambique's Independence War.*—We use the independence war from Portugal fought by Mozambique between 1964 and 1975 to illustrate the type of empirical variation leveraged in this section. Figure 5 shows the evolution of the estimated cohort effects and schooling across the cohorts in our sample. There are stark trend breaks in both variables for the cohorts of schooling age at that time. The war disrupted the educational system, as confirmed by the stagnating educational attainment for the cohorts starting primary school in the 1964–1975 time window; when adults, the same cohorts were more likely to work in agriculture, relative to the trend. After independence, the Mozambique Liberation Front led extensive

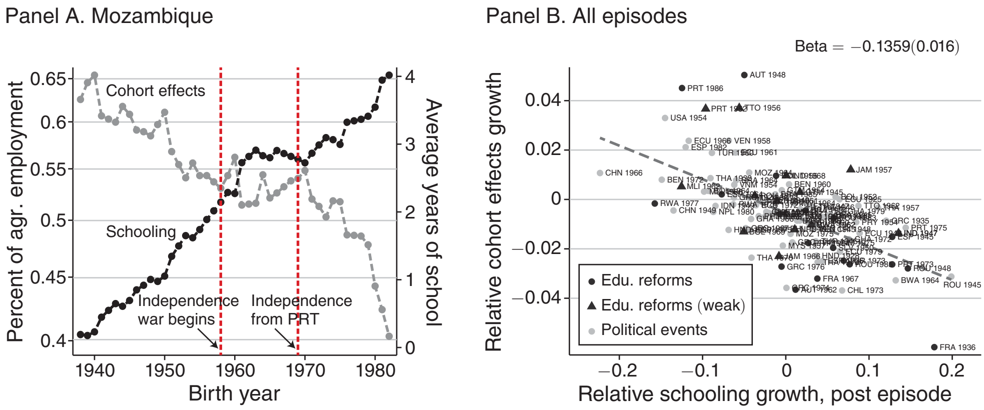


Figure 5. Trend Breaks around Education Reforms and Political Events

*Notes:* Panel A plots, for Mozambique, the cohort effects $\tilde{\mathbf{C}}_{c,j}$ estimated from specification (5) (left $y$-axis) and cohort schooling $s_{c,j}$ for all available birth cohorts (right $y$-axis). The red vertical lines identify the oldest cohorts not yet in school at the start and end of the independence war. Panel B plots the changes in the growth rate of cohort effects after a reform (in black) or political event (in gray), $\mathbf{A}_r$, against the corresponding changes for schooling, $\mathbf{S}_r$, as defined in equations (7) and (8). The black triangles identify the "weakly implemented" reforms. The gray dashed line shows the fit line, which has slope—0.1347 (0.016).

programs for economic development, including free healthcare and education, which are reflected in the faster schooling growth for cohorts born after 1970; once again, the cohort effects display a trend break in the opposite direction.

*All Reforms and Political Events*.—We formalize a similar analysis for the whole sample. For each policy reform or political event $r$, we denote as $\bar{c}_r$ the oldest cohort not yet in school at that time. We compute, for both cohort effects and schooling, the difference between the annualized growth across the cohorts born in ten-year windows before and after $\bar{c}_r$,

$$\mathbf{A}_r \equiv \frac{1}{10}\left(\tilde{\mathbf{C}}_{\bar{c}_r+10} - \tilde{\mathbf{C}}_{\bar{c}_r}\right) - \frac{1}{10}\left(\tilde{\mathbf{C}}_{\bar{c}_r-1} - \tilde{\mathbf{C}}_{\bar{c}_r-11}\right), \tag{7}$$

$$\mathbf{S}_r \equiv \frac{1}{10}\left(s_{\bar{c}_r+10} - s_{\bar{c}_r}\right) - \frac{1}{10}\left(s_{\bar{c}_r-1} - s_{\bar{c}_r-11}\right), \tag{8}$$

and plot $\mathbf{A}_r$ and $\mathbf{S}_r$ against each other.[18] Panel B of Figure 5 shows that when a reform or political event was followed by a positive trend break in schooling (i.e., by a faster increase in schooling for the affected cohorts) it was also followed by a negative trend break in cohort effects. In other words, the negative comovement that we have shown in panel A of Figure 5 for Mozambique generalizes to the whole dataset. To get a sense of the magnitude, we can compute the slope of the line, which would be, under a causal interpretation, the effect of schooling on agricultural

[18] We drop the episodes for which we have missing data for at least half of the cohorts in either of the two ten-year windows. Online Appendix C considers a variant to this exercise, where $\mathbf{A}_r$ and $\mathbf{S}_r$ are computed as the differences between the average annualized growth rates across all cohorts within the two ten-year windows; the results are very similar.

employment induced by the event. We find that one additional year of schooling decreases cohort-level agricultural employment by approximately 13.5 percent. Panel B of Figure 5 also highlights different patterns between fully implemented and weakly implemented educational reforms. Most of the fully implemented reforms lie on the bottom right quarter of the graph, meaning that they are associated with positive changes in schooling and negative changes in cohort effects; the pattern for the weakly implemented reforms (as well as for political events) is less clear: some of them are followed by positive and others by negative trend breaks.[19] In light of this, we next focus on the fully implemented reforms to quantify the impact of the associated schooling increase on the estimated cohort effects.

*An Event Study Design.*—We implement an event study design around the first cohort affected by the increase in compulsory education. For each policy reform $r$, we keep 10 cohorts older and 15 cohorts younger than $\bar{c}_r$. We detrend schooling and cohort effects using the growth across the cohorts born in a ten-year window before $\bar{c}_r$, and then regress each variable on a full set of dummies around $\bar{c}_r$. In particular, for schooling (and equivalently for the cohort effects), we estimate

$$\hat{s}_{c,r} = \delta_r + \sum_{x=-10}^{15} \mathrm{I}_{(c=\bar{c}_r+x)} + \varepsilon_{c,r}, \tag{9}$$

where $\delta_r$ is a reform fixed effect, $\mathrm{I}_{(c=\bar{c}_r+x)}$ is a dummy equal to 1 if cohort $c$ is born $x$ years after $\bar{c}_r$ and $\hat{s}_{c.r}$ is detrended schooling, constructed as

$$\hat{s}_{c,r} = s_{c,r} - \frac{c - \bar{c}_r + 10}{10}\left(s_{\bar{c}_r-1} - s_{\bar{c}_r-11}\right).$$

In panels A and B of Figure 6, we report the point estimates for the dummies $\mathrm{I}_x$. Consistently with the previous graphical analysis, we observe an increase in schooling and a decrease in cohort effects for cohorts born after $\bar{c}_r$. To have a sense of the implied magnitudes, we estimate specification (6), pooled across countries for which we have at least one education reform and using the dummies $\mathrm{I}_{(c=\bar{c}_r+x)}$ to instrument for schooling around the policy reforms. The results are shown in columns 3 and 4 of Table 2. The event study gives a negative, significant and large relationship between schooling and cohort effects: one extra year of schooling induced by the policy reforms led to a decline of cohort-level agricultural employment of almost 17 percent.

The event study makes progress relative to the correlations of Section IIIA. However, a few interpretation concerns remain. First, educational reforms are often part of broader policy interventions, which might affect the future returns from working in different sectors through other channels. Second, the reforms themselves might be implemented in anticipation of higher demand for nonagricultural labor. Third, large educational expansions may have general equilibrium effects that

[19] To formalize these points, online Appendix Figures C.2 and C.3 compare the distributions of trend breaks around the different types of reforms and political events with a placebo distribution of all the possible trend breaks in our data. As expected, the fully implemented reforms were followed by larger than average schooling increases, while the weakly implemented reforms are indistinguishable from the placebo trend breaks.

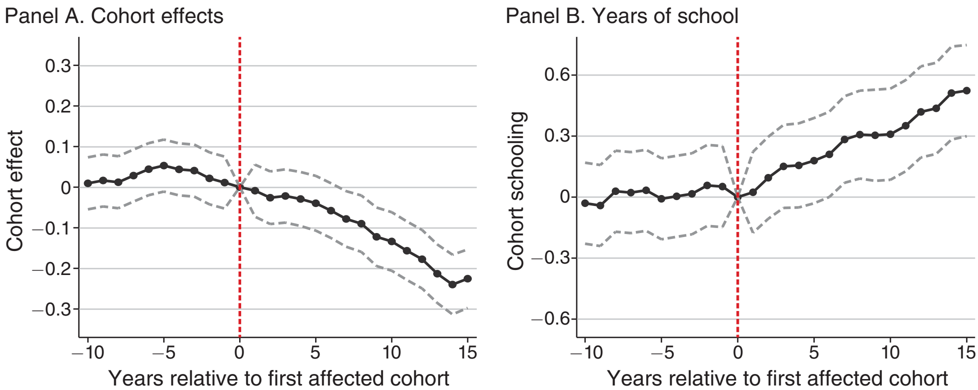


Figure 6. Event Study of Education Reforms on Agricultural Employment and Schooling

*Notes:* Panel A shows the point estimates and 90 percent confidence intervals for $I_x$ from specification (9). Panel B shows the point estimates for the same specification but using cohort effects on the left-hand side. The red line highlights the oldest cohort not yet in school at the time of the policy reform.

impact all birth cohorts, including those not directly exposed to the increase in compulsory education (Duflo 2004). While it is in general not possible to fully rule out these alternative interpretations, we notice that the fact that cohort effects are estimated controlling for year effects (absorbing any aggregate factor that affects the agricultural employment of all cohorts at a given point in time) helps to alleviate both the endogeneity and equilibrium concerns. For example, reverse causality is an issue only to the extent to which policy makers anticipate the timing of future changes in the demand for agricultural labor, that will affect some cohorts but not others. Similarly, general equilibrium forces contaminate our results only if they affect differentially young and old cohorts. Nonetheless, to further address these concerns, we next turn to a setting with more plausibly exogenous variation in cohort-level schooling.

## C. *School Construction in Indonesia*

Following the seminal work of Duflo (2001), we study the effects of the INPRES school construction program in Indonesia, which built 61,000 primary schools between 1974 and 1978. The identification exploits the facts that (i) the intensity of the program, as measured by the number of new schools per pupil, varied across districts, and (ii) only cohorts younger than six years old when the program started were fully exposed to it. We run a difference-in-difference exercise, comparing cohorts fully exposed to the treatment to those not exposed to it, in districts with higher and lower treatment intensity. The data (the 1995 intercensal survey of Indonesia) the identification strategy, and the specification closely follow Duflo (2001).[20] We refer the reader to that article for details.

[20] Karachiwalla and Palloni (2019) perform a very similar exercise using our same Indonesian data. While we both reach the same specification independently, the first version of our work was circulated in August 2017 (https://escholarship.org/uc/item/1ws4x2fg).

We restrict the sample to males born between 1950–1977. Consider the following specification:

$$(10)\qquad y_{icd} = \alpha_c + \eta_d + \sum_{k=1950}^{1977} (T_d \times \mathrm{I}_{ik})\delta_k + \sum_{k=1950}^{1977} (\xi_d \times \mathrm{I}_{ik})\varphi_k + \epsilon_{ijd},$$

where $(i, c, d)$ is an individual $i$, born in cohort $c$ and currently living in district $d$, $\alpha_c$ is a cohort fixed effect, $\eta_d$ is a district fixed effect, $T_d$ is the number of schools built per 1000 children in district $d$, $\mathrm{I}_c$ is a dummy equal to 1 if individual $i$ is born in cohort $c$, and $\xi_d$ is the school enrollment in 1972. The coefficients of interest are $\{\delta_c\}_{c=1950}^{1977}$, which capture the effects of program intensity on each cohort. We estimate (10) for three different outcome variables: (i) years of schooling; (ii) a dummy equal to 1 for agricultural employment; (iii) a dummy equal to 1 for nonagricultural employment.[21] We report the estimated coefficients and associated standard errors in Figure 7. As expected, the effect of the program was positive on education, negative on agricultural employment, and positive on nonagricultural employment.

In order to improve power, we follow Duflo (2001) and focus on the comparison of two cohorts: a treatment cohort of individuals that were between 2 and 6 years old at the time the program was implemented, and a control cohort of individuals that were between 12 and 17 years of age. The specification is the same as in (10), but with only one treatment cohort, and thus one coefficient of interest: the interaction between program intensity and the treatment cohort dummy.

Table 3 displays the results. Columns 1 and 2 show the reduced form specifications: the program is associated with a significant decrease in the probability of agricultural employment and an increase in the probability of nonagricultural employment; the latter is larger than the former, suggesting a significant flow from nonemployment to nonagricultural employment as well. Column 3 reports the first stage specification: one extra school per 1000 children increases schooling by $\sim$ 0.14, just as in Duflo (2001). Columns 4 and 5 show the IV results, where years of schooling are instrumented by the interaction between treatment intensity and the treated cohort dummy: one extra year of schooling reduces the probability of agricultural employment by 6.3 percentage points, and increases the probability of nonagricultural employment by 22.3 percentage points. This evidence shows that increases in schooling across cohorts led to lower propensities to work in agriculture.[22]

## D. *Summary and Discussion*

The results in this section support an interpretation of cohort effects as changes in human capital. We have considered several sources of cohort-level differences in educational attainment; in all cases, they are reflected in corresponding differences in agricultural employment, as captured by our cohort effects. The policy results are particular noteworthy, as they suggest that, at the micro level, governments can affect sectoral choices by increasing access to formal education.

[21] Nonemployment is the residual category with respect to (ii) and (iii); as a result, the effects on agricultural and nonagricultural employment are not necessarily symmetric.

[22] The results of this exercise are not directly comparable to the ones in Table 2, since here we are running a linear probability model. In online Appendix C, we run a cohort-level regression that allows us to compare the two, and find similar magnitudes.

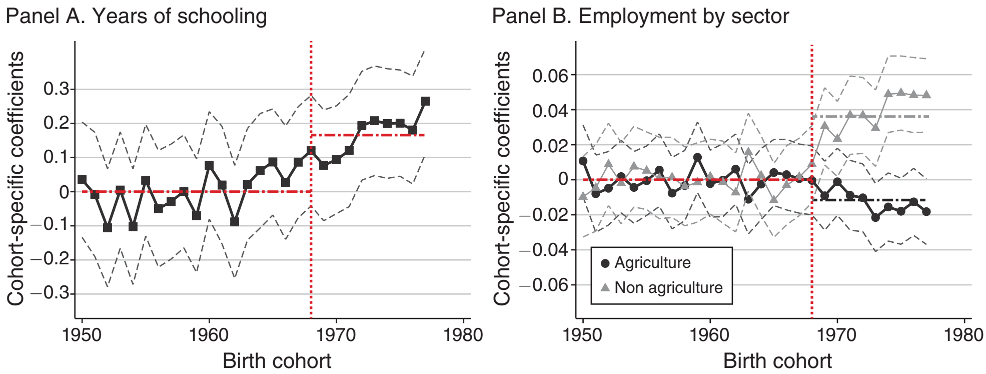


Figure 7. INPRES School Construction

*Notes:* Panel A shows the estimates of the cohort dummies from the first stage regression according to specification (10) when the left-hand side variable is years of schooling. Panel B shows the estimates for the reduced form results, from the same specification, with either agricultural or nonagricultural employment as left-hand side variables. The red dotted vertical line separates the treatment from the control cohorts. The coefficients are normalized to average zero for the control cohorts. Data for agricultural employment and schooling are from the 1995 intercensal survey of Indonesia (SUPAS); data for treatment intensity are from Duflo (2001).

Remarkably, the results are quantitatively similar across the different sources of empirical variation. Overall, we conclude that one extra year of school is associated with at least a 10 percent decline in agricultural employment. A simple back of the envelope calculation can shed light on the magnitude of this relationship. On average across our sample, the yearly increase in average schooling is 0.1057, which implies a decline in agricultural employment of a bit more than 1 percent. Abstracting from general equilibrium, we would conclude that the human capital deepening during the second half of the twentieth century accounts for about half of the overall labor reallocation from agriculture. In practice, equilibrium effects are likely to be important, and human capital growth might be endogenously related to other drivers of structural transformation. To address these points, we design an analytical model.

## IV. Model

This section develops a general equilibrium model of frictional labor reallocation out of agriculture by cohort. The model provides a structural interpretation of the cohort and year components estimated in Section II, and a framework to quantify their aggregate effects while taking into account equilibrium forces.

### A. *Environment*

We start by describing the economic environment. Time is discrete and runs infinitely.

*Demographics, Preferences, and Individual Traits*.—The economy is inhabited by $N + 1$ overlapping cohorts, indexed by $c$, each composed by a continuum of mass one of workers. Individuals of cohort $c$ enter the labor market at time $c$ and then work for a total of $N + 1$ periods; therefore, they work each period in $\{c, \ldots, c + N\}$.

Table 3—School Construction and Sectoral Employment: Evidence from Indonesia

| | Employed in agriculture (1) | Employed in nonagriculture (2) | Years of schooling (3) | Employed in agriculture (4) | Employed in nonagriculture (5) |
|---|---|---|---|---|---|
| Treated cohort × intensity | −0.009 | 0.031 | 0.137 | | |
| | (0.004) | (0.005) | (0.036) | | |
| Years of schooling | | | | −0.063 | 0.223 |
| | | | | (0.030) | (0.063) |
| Cohort fixed effects | Y | Y | Y | Y | Y |
| District fixed effects | Y | Y | Y | Y | Y |
| Method | OLS | OLS | OLS | IV | IV |
| *F*-stat first stage | – | – | – | 14.29 | 14.29 |
| Observations | 53,154 | 53,154 | 53,154 | 53,154 | 53,154 |

*Notes:* Robust standard errors in parentheses. Treated cohorts are those aged 2–6 when the school construction program was implemented. *Intensity* is the number of schools built per 1,000 children in each given district.

They derive an increasing and nonsatiated utility from the consumption of an agricultural and a nonagricultural good, and supply labor inelastically.

In each period, workers self-select in one of the two sectors of the economy, agriculture and nonagriculture. In agriculture, all workers have identical productivity. In nonagriculture, workers supply $h(c,\varepsilon)$ efficiency units, where $h(c,\varepsilon)$ depends on the birth cohort $c$ as well as on the individual-level (and time invariant) ability $\varepsilon$.[23] In particular, we assume the Cobb-Douglas aggregator

$$h(c,\varepsilon) = h_c^{\gamma}\varepsilon^{1-\gamma},$$

where $h_c$ captures a nonagricultural productivity shifter specific to cohort $c$, to which we return below, and $\gamma \geq 0$ is a parameter controlling the relative weight of the cohort- and individual-level components. In what follows, we refer to $h(c,\varepsilon)$ and $h_c$ as, respectively, individual-level and cohort-level human capital. We assume that $\varepsilon$ is distributed according to a $\beta(\nu,1)$ distribution, where $\nu$ is inversely related to the within-cohort variability. We define the aggregate stock of human capital at time $t$ as

$$H_t = \sum_{c=t-N}^{t} \int h(c,\varepsilon)dF(\varepsilon).$$

*Production*.—We index the agricultural sector by $A$ and the nonagricultural sector by $M$. The production of the agricultural good requires land $X$ and the labor input $L_{A,t}$ , while the production of the nonagricultural good only requires the labor input $L_{M,t}$ . Land is owned collectively by all individuals, who share the profits and use them to finance consumption. Productivity in agriculture, $Z_{A,t}$, may differ from productivity in nonagriculture, $Z_{M,t}$. The relative price of agricultural goods in equilibrium

[23] The assumption that nonagriculture is more human capital intensive than agriculture is consistent with widely documented patterns of sorting of high-skilled workers in nonagriculture (e.g., Gollin, Lagakos, and Waugh 2014; Young 2013; Porzio 2017), larger returns to skills in nonagriculture (see Herrendorf and Schoellman 2018) and skill-specific mobility across sectors (see Hicks et al. 2017).

is given by $p_t$, which we describe below. The revenue functions in agriculture and nonagriculture are

$$p_t Y_{A,t} = p_t Z_{A,t} X^{\alpha} L_{A,t}^{1-\alpha},$$
$$Y_{M,t} = Z_{M,t} L_{M,t}.$$

Individuals of different cohorts are perfect substitutes in both sectors. However, as discussed above, the efficiency units supplied to the nonagricultural sector are heterogeneous both across and within cohorts. Letting $\omega_t(c,\varepsilon)$ be the occupational choice function, taking value 1 if individual $(c,\varepsilon)$ at time $t$ works in agriculture and 0 otherwise, the agricultural and nonagricultural labor inputs are given by

$$L_{A,t} = \sum_{c=t-N}^{t} \int \omega_t(c,\varepsilon) dF(\varepsilon),$$
$$L_{M,t} = \sum_{c=t-N}^{t} \int h(c,\varepsilon)\left[1 - \omega_t(c,\varepsilon)\right] dF(\varepsilon),$$

where $F(\varepsilon)$ is the within-cohort distribution of $\varepsilon$.

Firms choose optimally how many workers to hire, and the labor market is competitive. As a result, workers are paid the marginal product of their labor: the individual-level earnings in the two sectors are given by

$$y_{A,t} = w_{A,t} = (1-\alpha) p_t Z_{A,t} X^{\alpha} L_{A,t}^{-\alpha}, \tag{11}$$
$$y_{M,t}(c,\varepsilon) = w_{M,t} h(c,\varepsilon) = Z_{M,t} h(c,\varepsilon),$$

where $w_{A,t}$ and $w_{M,t}$ denote wages per efficiency unit in agriculture and nonagriculture.

*Worker's Sectoral Choice Problem.*—Markets are complete and there is perfect foresight. As a result, we can think of individual $(c,\varepsilon)$ choosing at time $c$ a sequence of occupations $\{\omega_t\}_{t=c}^{N+c}$, one for each period in her life. This choice is made taking as given her income paths in agriculture, $\{y_{A,t}\}_{t=c}^{N+c}$, and nonagriculture, $\{y_{M,t}(c,\varepsilon)\}_{t=c}^{N+c}$, as defined in (11). Moreover, sectoral changes are associated with a cost $C(\omega_{t-1}, \omega_t, y_{A,t}, y_{M,t}(c,\varepsilon))$. Formally, individual $(c,\varepsilon)$ solves

$$\max_{\{\omega_t\}_{t=c}^{c+N}} \sum_{t=c}^{c+N} \beta^{t-c}\left[\omega_t y_{A,t} + (1-\omega_t) y_{M,t}(c,\varepsilon) - C(\omega_{t-1}, \omega_t, y_{A,t}, y_{M,t}(c,\varepsilon))\right],$$

subject to $\omega_{c-1} = 1$, where we are assuming that all individuals are born in agriculture.[24] The mobility friction takes the form

$$C(\omega_{t-1}, \omega_t, y_{A,t}, y_{M,t}(c,\varepsilon)) = \mathrm{I}(\omega_t = 0) i\, y_{M,t}(c,\varepsilon) + \mathrm{I}(\omega_t < \omega_{t-1}) f y_{M,t}(c,\varepsilon)$$
$$+ \mathrm{I}(\omega_t > \omega_{t-1}) f y_{A,t},$$

[24] This assumption keeps the problem symmetric for all ages, preserving the log-linearity of cohort-level agricultural employment both over time and across cohorts.

and includes (i) a per-period cost that reduces the nonagricultural wage by a fraction $i$ in each period, and (ii) a one-time fixed cost that reduces the wage in the destination sector by a fraction $f$ in periods when a change of sector takes place. The former can be interpreted as an amenity cost (as in Lagakos, Mobarak, and Waugh. 2019) or as any other flow cost associated with leaving the agricultural sector, as, for example, the exclusion from risk-sharing communities (Munshi and Rosenzweig 2016; Morten 2019). The fixed cost can be interpreted as a one-time mobility cost (as in Heise and Porzio 2021) which might be driven by actual moving expenses (if a geographical move is necessary to change sector) or any other associated cost, such as retraining, idle time in between jobs, or one-time emotional costs.

*Growth of Cohort-Level Human Capital*.—We postulate that the cohort-level human capital $h_c$ varies across cohorts according to

$$\log h_c^{\gamma} = \sigma \log \frac{V_{M,c}}{V_{A,c}} + \log \xi_c, \tag{12}$$

where $V_{x,c} = \sum_{t=c}^{c+N} \beta^t w_{x,t}$ denotes the present discounted value of wages in sector $x$ for cohort $c$ and $\log \xi_c$ is an exogenous shifter. Equation (12) allows human capital growth to be driven by both endogenous and exogenous forces to structural transformation. The first term captures in reduced form that cohorts facing a higher relative return from working in the human capital intensive sector might endogenously accumulate more human capital, with the parameter $\sigma$ modulating the strength of this effect. The exogenous shifter $\log \xi_c$ captures all other changes in the costs or returns of human capital accumulation, possibly driven by improvements in quality, advances in transportation technology or policies that extend access to public education.[25]

*Price of Agricultural Goods*.—Closing the model requires an equation determining the equilibrium agricultural price. We postulate the following log-linear equation, which allows to keep analytical tractability while encompassing, in reduced form, the main mechanisms suggested in the literature as possible drivers of structural change,

$$\underbrace{\log p_t}_{\text{agr price}} = \eta \Big( \underbrace{\log \theta_t}_{\text{demand}} + \underbrace{\eta_z \log z_t}_{\text{supply}} + \underbrace{\eta_H \log H_t}_{\text{human capital}} \Big), \tag{13}$$

where $\log \theta_t$ is a demand shifter that captures the relative demand for agricultural goods and $\log z_t$ is relative agricultural productivity, $z_t \equiv Z_{A,t}/Z_{M,t}$.

The parameters $\eta$, $\eta_z$, and $\eta_H$ modulate the role of each variable in determining the agricultural price and structural transformation. In particular, $\eta = 0$ corresponds to the case of a small open economy with no trade frictions— that is, an economy that takes the prices of agricultural and nonagricultural goods as given (we refer to this case as simply "small open economy"). When $\eta > 0$, a decrease over time in the relative demand for agricultural goods decreases the relative price, leading to labor

[25] In practice, education policies might themselves be endogenous responses to the growth in the human capital intensive sector, hence could also be captured, at least in part, by the first term of equation (12); see for example Galor and Mountford (2008) for a framework that includes this channel.

reallocation out of agriculture, as in Kongsamut, Rebelo, and Xie (2001) and Comin, Lashkari, and Mestieri (2021). With $\eta_z < 0$, faster productivity growth in agriculture leads to a decline in the relative price and labor reallocation our of agriculture if $\eta$ is large enough, as in as in Ngai and Pissarides (2007), while it pushes workers *into* agriculture when economy is sufficiently close to a small open economy, i.e., if $\eta$ is small, as in Matsuyama (1992a). The parameter $\eta_H$ governs the impact of human capital growth; its sign is priori ambiguous, given that an increase in $H_t$ should cause both (i) an income effect due to individuals becoming richer, decreasing the relative demand and the relative price of agricultural goods, and (ii) a decrease in relative agricultural labor productivity, possibly increasing the relative price.

*Role of log-Linearity*.—Equations (12) and (13) guarantee that cohort-level human capital and relative agricultural price are log-linear functions of the exogenous driving forces. Of course, these are a stark assumptions, yet ones that comes with large benefits in terms of tractability. Together with the assumption that $\varepsilon$ is distributed according to a $\beta(\nu, 1)$ distribution, they allow us, as we will show below, to derive analytical expressions for the aggregate demand and supply of agricultural workers.

One possibility would be to interpret equations (12) and (13) as log-linear approximations of the solution of a model with a fully specified demand system and human capital accumulation problem. In online Appendix E, we design, calibrate, and solve numerically a quantitative version of our model, and show that equations (12) and (13) indeed offer good approximations of its (nonlinear) solution.

DEFINITION OF EQUILIBRIUM: *An equilibrium in this economy is given by a sequence* $\left\{L_{A,t}, w_{A,t}, w_{M,t}(c,\varepsilon), \omega_t(c,\varepsilon) \forall c \in [t-N,t]\right\}_{t=0}^{\infty}$ *such that, given the paths of agricultural demand, sectoral productivities, and the cohort-level human capital shifter* $\{\theta_t, Z_{A,t}, Z_{M,t}, \xi_t\}_{t=0}^{\infty}$, *firms maximize profits taking wages as given, individuals choose optimally their occupation at each point in time taking wages as given, and the labor market clears in both the agriculture and nonagriculture sectors.*

## B. *Determinants of Labor Reallocation*

To simplify the analytical analysis, we impose four parametric assumptions, which, as we discuss below, lead to empirical predictions consistent with the evidence in Section III.

ASSUMPTION 1: *The demand shifter* $\theta_t$, *relative productivity* $z_t$, *and the human capital shifter* $\xi_t$ *change at constant rates, i.e., for each t and c:* $\log \frac{\theta_{t+1}}{\theta_t} = \log g_\theta$, $\log \frac{z_{t+1}}{z_t} = \log g_z$, *and* $\log \frac{\xi_{c+1}}{\xi_c} = \log g_\xi$.

ASSUMPTION 2: *The mobility costs are such that* $i \in [0, \bar{i}]$ *and* $f \in [0, \bar{f}]$, *where* $\bar{i}$ *and* $\bar{f}$ *are functions of the parameters reported in online Appendix D.*

ASSUMPTION 3: *The growth rates of the demand shifter* $g_\theta$ *and relative productivity* $g_z$ *satisfy* $\log g_{\theta z} \equiv \eta \log g_\theta + (1 - \eta\eta_z) \log g_z \leq \max\{0, -\Psi \log g_\xi\}$, *where* $\Psi \equiv \gamma\left(\eta\eta_H + \frac{\alpha\nu}{1-\gamma}\right)$.

ASSUMPTION 4: *The price effect of human capital satisfies* $\eta\eta_H \geq -\frac{\alpha\nu}{1-\gamma}\frac{1+\sigma}{\sigma} - \frac{1}{\sigma}$.

Assumption 1 allows us to focus on equilibria where agricultural employment decreases at a constant rate. Assumption 2 puts upper bounds on the mobility costs to guarantee that (i) a positive mass of individuals in each cohort works out of agriculture, and (ii) at least some workers switch sector within their lifetime. Assumption 3 ensures that the decline over time in the relative demand for agricultural labor is large enough to generate a negative year component (i.e., within cohort reallocation out of agriculture). All these predictions hold in the data. Assumption 4 rules out the extreme case where the price effect of human capital is so negative that an increase in the exogenous human capital shifter leads to a more than proportional decrease in endogenous human capital.[26]

*Labor Reallocation by Cohort.*—We characterize the solution of the sectoral choice problem. To illustrate the core mechanism, it is useful to first consider the model's solution with no fixed cost to change sectors, i.e., with $f = 0$. In this case, the sectoral choice problem becomes a repeated static choice: an individual $(c,\varepsilon)$ works in the nonagricultural sector if her net income is higher there, hence if $(1-i)w_{M,t}h(c,\varepsilon) \geq w_{A,t}$. Given that $h(c,\varepsilon) = h_c^{\gamma}\varepsilon^{1-\gamma}$, the inequality holds if her ability satisfies $\varepsilon \geq \hat{\varepsilon}_t(c) \equiv \left[\frac{w_{A,t}}{(1-i)w_{M,t}}h_c^{-\gamma}\right]^{1/(1-\gamma)}$. The agricultural employment for cohort $c$ at time $t$ is thus given by the share of workers with ability smaller than $\hat{\varepsilon}_t(c)$:

$$\log l_{A,t,c} = \log F(\hat{\varepsilon}_t(c)) = \frac{\nu}{1-\gamma}\log\frac{w_{A,t}}{(1-i)w_{M,t}} - \frac{\gamma\nu}{1-\gamma}\log h_c. \tag{14}$$

Labor reallocates out of agriculture for a given cohort as long as the relative agricultural wage, $\log\frac{w_{A,t}}{w_{M,t}}$, declines over time, and between successive cohorts as long as human capital, $\log h_c$, grows across cohorts. Moreover, there is sorting based on comparative advantage, both within and across cohorts. Within any cohort, individuals with higher ability work out of agriculture. Across cohorts, the ones with a higher cohort-level human capital have a larger share of individuals out of agriculture.

These same forces are present in the general case with $f > 0$, but need to be refined to take into account the possibility that some individuals are constrained by the fixed cost $f$. For young individuals with sufficiently high ability the fixed cost does not bind, as it is discounted over their (long) remaining working life. The rate of labor reallocation of young birth cohorts is thus identical to the case with $f = 0$ (we refer to these cohorts as "unconstrained" at time $t$).[27] Workers that are still in agriculture when old, instead, may be trapped there by the fixed cost, given that

[26] Assumption 4 is not strictly necessary for our purposes, and the mapping between the model and the empirical decomposition would be unaffected even if this Assumption does not hold. However, in that case the model would give predictions both counterintuitive and at odds with the empirical evidence; for example, labor reallocation out of agriculture could be generated by an increase in the demand for agricultural goods, and overall human capital growth by a decrease in the exogenous human capital shifter. We impose this Assumption here to abstract from these degenerate cases in the following discussion. Moreover, our estimates in Section VIB suggest that the endogenous and exogenous components of human capital are both growing over time, consistently with Assumption 4.

[27] Their level of agricultural employment, however, is still declining in $f$, as a larger moving cost makes the nonagriculture less appealing.

even the highest ability among them might not be willing to switch sector with only a few periods left to work (we refer to these cohorts as "constrained" at time $t$). The agricultural employment of constrained cohorts remains constant for the rest of their working life, and is proportional to the relative wage in the last period where they were unconstrained. We formalize these results in Lemma 1; we focus here on the case where the relative agricultural wage decreases at a constant rate, which, as we show later, holds in equilibrium.

LEMMA 1 (Labor Reallocation by Cohort): *Let* $a_t(c) = t - c$ *be the age of cohort* $c$ *at time* $t$*. If the relative agricultural wage* $w_{A,t}/w_{M,t}$ *decreases at a constant rate, there exists a threshold* $\hat{a}(f)$, *with* $1 \leq a(f) \leq N$, *such that for any* $c$ *and* $t$,

$$\log l_{A,t,c} = \begin{cases} \lambda(i,f) + \frac{\nu}{1-\gamma}\log\frac{w_{A,t}}{w_{M,t}} - \frac{\nu\gamma}{1-\gamma}\log h_c & \text{if } a_{t+1}(c) \leq \hat{a}(f); \\ \lambda(i,f) + \frac{\nu}{1-\gamma}\log\frac{w_{A,c+\hat{a}}}{w_{M,c+\hat{a}}} - \frac{\nu\gamma}{1-\gamma}\log h_c & \text{if } a_{t+1}(c) > \hat{a}(f); \end{cases}$$

*where* $\lambda(i,f)$ *and* $\hat{a}(f)$ *are constant over time and across cohorts, and satisfy* $\lambda(0,0) = 0$ *and* $\hat{a}(0) = N$.

PROOF:

See online Appendix D.

Figure 8 visualizes the patterns of labor reallocation implied by Lemma 1 for two successive cohorts, $c$ and $c+1$. While younger than $\hat{a}$, agricultural employment declines at a common rate across cohorts, proportional to the decline in the relative agricultural wage. At a given point in time, the agricultural employment gap between two unconstrained cohorts is proportional to the human capital gap between the two generations. For constrained cohorts, instead, there is an additional gap due to the shorter horizon of the older cohort.

*Aggregate Labor Reallocation.*—Aggregating up cohort-level agricultural employment from Lemma 1, we can write the overall agricultural labor supply at time $t$ as

$$(S_t) \qquad \log L_{A,t} = \lambda_S - \frac{\nu}{1-\gamma}\log H_t + \frac{\nu}{1-\gamma}\log\frac{w_{A,t}}{w_{M,t}},$$

where $\lambda_S$ is a time-invariant term. The supply is upward sloping with respect to the relative wage, as a higher relative wage induces more individuals to stay in agriculture. Increases in human capital lead to a downward shift of the agricultural labor supply, as human capital is more valued outside of agriculture. Mobility frictions are subsumed into the $\lambda_S$ term, and do not affect the slope or the magnitude of the shift associated with changes in $H_t$.

In equilibrium, agricultural employment is given by the intersection between $(S_t)$ and agricultural labor demand, which comes from the firms' optimality conditions and it is given by

$$(D_t) \qquad \log L_{A,t} = \lambda_D + \frac{1}{\alpha}(\log p_t + \log z_t) - \frac{1}{\alpha}\log\frac{w_{A,t}}{w_{M,t}}.$$

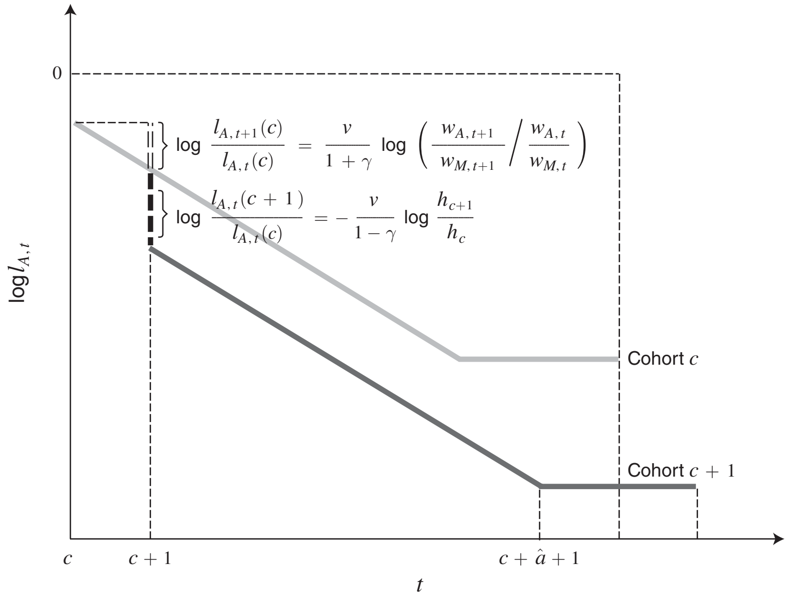


Figure 8. Labor Reallocation by Cohort

Aggregate labor reallocation out of agriculture occurs if either the demand or the supply curve (or both) shift downward. The two main mechanisms behind structural change commonly emphasized in the literature, i.e., changes in the relative demand for agricultural goods ($\theta_t$) and in relative agricultural productivity ($z_t$), lead to downward shifts in the relative demand for agricultural labor. The focus of this paper, instead, is to study the role of downward shifts in the relative supply curve, which is usually assumed constant in the literature.

Panel A of Figure 9 illustrates how an increase in $H$ between any two periods $t$ and $t+1$ can lead to aggregate labor reallocation, even keeping $\theta$ and $z$ constant. As discussed, human capital growth generates a downward shift of the supply curve. In partial equilibrium, i.e., if the relative wage and price are constant, this shift would result at $t+1$ in a level of agricultural employment of $L_{A,t+1}^{PE}$. When wages are allowed to adjust but prices are kept fixed (the case of a small open economy) the resulting agricultural employment is $L_{A,t+1}^{SOE}$, which is larger than $L_{A,t+1}^{PE}$ since the adjustment in relative wages attenuates the employment effect of the supply shift. Finally, if the relative price of the agricultural good adjusts as well (if $\eta > 0$, as in a closed economy), the increase in $H_t$ leads additionally to a downward or upward shift in the demand curve, depending on the sign of $\eta_H$. The resulting agricultural employment $L_{A,t+1}^{CE}$ can be higher or lower than $L_{A,t+1}^{SOE}$, and can in principle even be higher than $L_t$: if the price elasticity is high enough, a decrease in the supply of agricultural labor could increase the relative price sufficiently to pull workers into agriculture.[28] Panel A of Figure 9 shows the case where $L_{A,t+1}^{CE}$ is in between $L_{A,t+1}^{SOE}$ and $L_t$.

[28] This result is reminiscent of Matsuyama (1992b), which shows that agricultural productivity growth has opposite implications on agricultural employment in a closed and open economy. The same is potentially true for

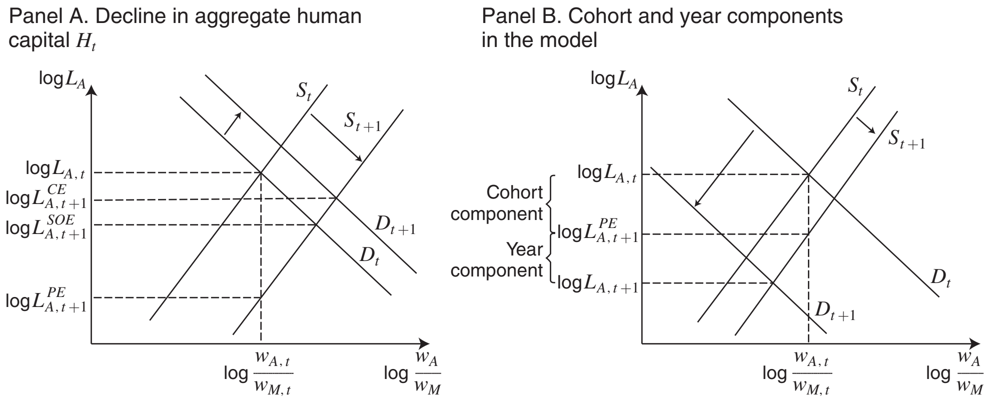


Figure 9. Aggregate Labor Reallocation

Proposition 1 generalizes this discussion, characterizing the forces driving labor reallocation. First, equation (15) establishes the existence of an equilibrium where all endogenous variables change at constant rates.[29] Second, equation (16) shows that the aggregate rate of labor reallocation is determined by shifts over time of the $(S_t)$ and $(D_t)$ curves. Third, substituting equilibrium prices, equation (17) highlights that labor reallocation out of agriculture $(\log g_{L_A} < 0)$ is triggered by demand forces $(\log g_{\theta z} < 0)$ and human capital growth $(\log g_h > 0)$, with the direct effect of each mediated by the within-cohort ability distribution (which determines the mass of workers leaving agriculture for a given change in relative wages) and by general equilibrium effects. The impact of demand forces is unambiguous, since $\Theta_D \in [0,1)$. The effect of human capital growth can be amplified or attenuated by general equilibrium forces; in a small open economy $\Theta_S = \Theta_D \in [0,1)$, while in a closed economy $\Theta_S$ can be positive or negative depending on the parameters' values. We refer to $1 - \Theta_S$ as the *general equilibrium* (*GE*) *multiplier* of human capital growth. Finally, equation (18) shows that human capital growth is itself a combination of the exogenous human capital shifter and the endogenous response to the demand forces behind structural transformation. As a result, as long as $\sigma > 0$, a decline in demand for agricultural labor may be amplified by an induced decline in the supply of agricultural workers due to individuals accumulating more human capital.[30]

PROPOSITION 1 (Aggregate Labor Reallocation): *There exists an equilibrium where agricultural employment, cohort-level human capital, the relative agricultural*

changes in human capital; however, for those to increase agricultural employment, it needs to be the case that both the economy is sufficiently closed $(\eta > 0)$ and the productivity effect of human capital on prices dominates the income effect $(\eta_H > 0)$.

[29] The fact that the "initial old" born at time 0 have a shorter life span than other cohorts makes their dynamic problem artificially different and might imply, depending on the initial conditions, that the economy reaches this "constant reallocation path" after a transition. To simplify the exposition, we set initial conditions for those cohorts such that the economy is on the constant reallocation path right away; see online Appendix D for the details.

[30] In online Appendix D, we solve for the independent role of the exogenous human capital shifter in labor reallocation, combining equations (17) and (18).

*price, and the relative agricultural wage change at constant rates, i.e., for each t and c,*

$$\begin{aligned}\log\frac{L_{A,t+1}}{L_{A,t}} &= \log g_{L_A},\\ \log\frac{h_{c+1}}{h_c} &= \log g_h,\\ \log\frac{p_{t+1}}{p_t} &= \log g_p,\\ \log\left(\frac{w_{A,t+1}}{w_{A,t}}/\frac{w_{M,t+1}}{w_{M,t}}\right) &= \log g_w.\end{aligned} \tag{15}$$

*Labor reallocation out of agriculture is driven by demand and supply shifts,*

$$\log g_{L_A} = \frac{\nu\alpha}{1-\gamma+\alpha\nu}\underbrace{\left[\frac{1}{\alpha}(\log g_p + \log g_z)\right]}_{\text{Demand shift }(\mathbf{D})} + \frac{(1-\gamma)}{1-\gamma+\alpha\nu}\underbrace{\left[-\frac{\nu\gamma}{1-\gamma}\log g_h\right]}_{\text{Supply shift }(\mathbf{S})}, \tag{16}$$

*and can be written as*

$$\log g_{L_A} = \underbrace{\left(\frac{v}{1-\gamma}\right)}_{\text{Skill distr.}}\left[\left(\underbrace{1-\Theta_D}_{\text{GE}}\right)\log g_{\theta z} - \left(\underbrace{1-\Theta_S}_{\text{GE}}\right)\gamma\log g_h\right], \tag{17}$$

*where* $\Theta_D \equiv \frac{\alpha v}{1-\gamma+\alpha v}$, *and* $\Theta_S \equiv \frac{\alpha v + (1-\gamma)\eta\eta_H}{1-\gamma+\alpha v}$. *The growth rate of human capital satisfies*

$$\gamma\log g_h = \frac{(1-\gamma+\alpha\nu)\log g_\xi - \sigma(1-\gamma)\log g_{\theta z}}{(1-\gamma)(1+\eta\eta_H\sigma)+\alpha\nu(1+\sigma)}. \tag{18}$$

PROOF:

See online Appendix D.

### C. *Summary and Discussion*

We developed an analytical model where labor reallocation is driven by income effects, unbalanced productivity growth, and human capital growth across cohorts. While the first two forces—traditionally emphasized in the literature—work through a downward shift in the agricultural labor demand curve, human capital growth leads to an additional shift in the supply curve. The model highlights how shifts in supply and demand are intertwined due to price effects and endogenous skill accumulation. Moreover, it provides a useful mapping from the three exogenous driving forces (sectoral productivities, the relative demand for agricultural goods, and the cost of human capital acquisition) into shifts of demand and supply and the aggregate rate of labor reallocation.

## V. Interpreting the Empirical Decomposition

We now use the model developed in Section IV to interpret the decomposition results from Section II. This leads us to a core result of the paper: shifts in the relative supply of agricultural labor account for a large share of labor reallocation out of agriculture.

### A. *Accounting for Supply and Demand Shifts*

Through the lens of the model, the empirical decomposition pins down the relative role of supply and demand shifters for aggregate labor reallocation. Lemma 1 shows that, for unconstrained cohorts, agricultural employment can be written as the combination of year and cohort effects. The year effects are proportional to the relative agricultural wage, which depends on aggregate prices and quantities, while the cohort effects capture the cohort-level human capital. In the presence of mobility frictions, the cross-cohort agricultural employment gaps for constrained cohorts are larger than what implied by the human capital difference alone, as can be visualized in Figure 8. This is where the age controls introduced in specification (5) become important. Under the identification restriction of a zero age effect for any young (unconstrained) cohort, age controls capture the different reallocation behavior of old (constrained) cohorts, so that the estimated cohort effects retain their structural interpretation as measures of human capital.

Consider the year and cohort components, as defined in Section II, of the rate of labor reallocation between $t$ and $t+1$. The year component captures the difference between the year effects associated to $t$ and $t+1$, which, once age effects are controlled for, is proportional to the change in the relative agricultural wage per efficiency unit,

$$\log \widetilde{\psi}_t = \frac{\nu}{1-\gamma} \log g_w.$$

The cohort component captures the change over time in the average cohort effects for the active cohorts. Given that in our model cohort effects change across cohorts by a constant amount, this corresponds to the difference between the cohort effects of any two consecutive cohorts, which in turn is proportional to the growth rate of human capital

$$\log \widetilde{\chi}_t = -\frac{\nu\gamma}{1-\gamma} \log g_h. \tag{19}$$

These two quantities correspond to different aspects of the process of labor reallocation. Notice from equation (16) that $-\frac{\nu\gamma}{1-\gamma}\log g_h$ represents the shift in the agricultural labor supply driven by human capital growth. As a result, the cohort component captures the partial equilibrium effect of the change in supply, i.e., the decrease from $\log L_{A,t}$ to $\log L^{PE}_{A,t+1}$, as displayed in panel B of Figure 9 . The year component instead captures the residual part of reallocation, i.e., the difference between $\log L^{PE}_{A,t+1}$ and $\log L_{A,t+1}$. Intuitively, gaps in agricultural employment between different cohorts at a given point in time (i.e., the cohort component) identify the extent to which changes in human capital shift the supply curve, keeping wages fixed; on the other hand, changes over time for a given cohort (i.e., the year component) identify the movement along a given supply curve driven by changes in relative wages. The following Proposition formalizes these results.

PROPOSITION 2 (Decomposition of Labor Reallocation): *Consider the specification*

$$\underbrace{\log l_{A,t,c}}_{\text{agr share of cohort } c \text{ at time } t} = \underbrace{\tilde{\mathbf{Y}}_t}_{\text{year effects}} + \underbrace{\tilde{\mathbf{C}}_c}_{\text{cohort effects}} + \underbrace{\tilde{\mathbf{A}}_{t-c}}_{\text{age dummies}} + \varepsilon_{t,c},$$

*estimated with model-generated data under the restriction that* $\tilde{\mathbf{A}}_a = \tilde{\mathbf{A}}_{a-1}$, *where* $a \in [1, \hat{a}]$. *Define the year and cohort components of labor reallocation between* $t$ *and* $t+1$ *as*

$$\log \tilde{\psi}_t \equiv \tilde{\mathbf{Y}}_{t+1} - \tilde{\mathbf{Y}}_t,$$

$$\log \tilde{\chi}_t = \log L_{A,t+1} - \log L_{A,t} - \log \tilde{\psi}_{t+1}.$$

*Then, for all* $t$,

$$\log \tilde{\psi}_t = \log \tilde{\psi} = \frac{\nu}{1-\gamma} \log g_w = \frac{\nu\alpha}{1-\gamma+\alpha\nu}(\mathbf{D}-\mathbf{S}),$$
$$\log \tilde{\chi}_t = \log \tilde{\chi} = -\frac{\nu\gamma}{1-\gamma} \log g_h = \mathbf{S},$$

*where* $\mathbf{D}$ *and* $\mathbf{S}$ *are the demand and supply shifts, as defined in Proposition 1.*

PROOF:

See online Appendix D.

*Large Decline in the Supply of Agricultural Labor*.—As shown in Table 1, the cohort component is on average $-0.79$ percent, corresponding to 38 percent of the rate of labor reallocation. Proposition 2 tells us how to read these figures in the context of the model: human capital growth has led, on average across countries, to a downward shift of the agricultural labor supply at annual rate of $-0.79$ percent. If wages are kept fixed (i.e., if the labor demand is vertical) this coincides with the labor reallocation induced by human capital growth, representing 38 percent of the overall reallocation; in general equilibrium, the impact of the supply shift is mediated by the multiplier $1-\Theta_s$ (as defined in Proposition 1), which we quantify in Section VIA. Moreover, cross-country differences in the cohort component can be interpreted as reflecting different magnitudes of the supply shift. Revisiting for example Figure 2, the larger role of the cohort component in Indonesia implies, through the lens of the model, that human capital growth played a relatively more important role in that country compared to Brazil.[31]

These results highlight a key take-away of the paper: changes over time in agricultural labor supply (both endogenous and exogenous to demand forces) represent an important feature of the process of structural transformation, which is missed by models that abstract from workers' skills and their differential use across sectors.

[31] Consistently with this conclusion, the schooling increase has been in fact steeper in Indonesia over the sample period, while Brazil has seen faster growth in agricultural productivity.

### B. *Quantifying Mobility Frictions*

The following corollary shows that the comparison between the decomposition results with and without age controls is directly informative on the role of reallocation frictions. Since mobility costs limit the reallocation of older workers, not controlling for age results in an overstatement of the cohort component and an understatement of the year component. The difference between the year components estimated with and without age controls is proportional to the share of constrained cohorts, $\lambda(f)$, a natural measure of the severity of reallocation frictions.[32]

COROLLARY 1 (Bias in the Basic Decomposition): *Consider specification (1) estimated with model-generated data. The estimated year and cohort components would be*

$$\log\psi = \left[1 - \lambda(f)\right]\log\tilde{\psi},$$

$$\log\chi = \log\tilde{\chi} + \lambda(f)\log\tilde{\psi},$$

*where* $\lambda(f) = \frac{N - \hat{a}(f)}{N+1} \in [0,1)$ *is the share of constrained cohorts which is increasing in the fixed cost f and does not depend on the per-period cost i.*

PROOF:
See online Appendix D.

Building on Corollary 1, we compute for each country $j$ the implied value of $\lambda(f_j)$ as $\lambda(f_j) = 1 - \log\psi_j/\log\tilde{\psi}_j$, and display summary statistics in column 6 of Table 1. On average across countries, $\lambda(f_j)$ is approximately 30 percent, which means individuals' reallocation decision is constrained by the fixed cost in the last 30 percent of their working life, or in our sample, after they turn 49 years old. Frictions are virtually nonexisting in high-income countries, and considerably more severe in poorer countries.[33]

### C. *Summary and Discussion*

Through the lens of the model, we interpret the cohort component as measuring the decline in agricultural labor supply associated with human capital growth. We

[32] Human capital growth and reallocation frictions can be quantified without relying on the measurement of wages, which is notoriously difficult for developing countries and the agricultural sector. The model does have predictions on wages that are in line with the limited available evidence; in particular, we show in online Appendix D that it is consistent with the observational wage gains for workers moving from agriculture to nonagriculture being smaller than the corresponding cross-sectional gaps (Hicks et al. 2017; Herrendorf and Schoellman 2018; Alvarez 2020). However, we also show that even a panel of wage data would not be enough to infer the magnitude of the frictions; the fixed cost makes the sectoral decision dynamic, and to estimate mobility costs one would need the hypothetical wage paths in agriculture in nonagriculture for both movers and nonmovers. Corollary 1 gives an alternative way of quantifying these costs.

[33] The fact that fictions appear to be larger in middle-income countries than low-income ones is largely driven by the three outliers in terms of the effect of age controls displayed in Figure 3. Indeed, Table B.5 shows that when comparing median values across income groups (as opposed to using averages), the magnitude of frictions is strictly decreasing with development.

conclude that this supply shift, whether by itself an independent driver or simply a mediating factor, is a key proximate cause of labor reallocation, in the language of the development accounting literature (Caselli 2005). As a result, ignoring it, as done by most of the literature on structural change, leads to an incomplete understanding of the process through which structural transformation unfolds.

This result, while important, has two apparent shortcomings. First, the exact quantification hinges on the structure of the model. Second, it is an accounting result, in that it does not tell us the primitive drivers or the equilibrium effect of human capital growth. To make progress, we conclude this section by discussing how variations to some of the assumptions would affect the mapping between the model and the empirical decomposition, while in the next section we bring additional evidence to go beyond the accounting nature of the results.

First, the model maps cohort-level differences in agricultural employment into changes in an attribute that makes individuals more productive in nonagriculture. In practice, the nonmonetary value of working in nonagriculture might vary across cohorts as well, perhaps as a result of changes in the quantity, quality and content of their education. As we show in online Appendix D, in such setting the cohort component captures the supply shift induced by both productivity and nonmonetary factors; while their relative importance cannot be separately identified, our results give their joint effect. Second, the model abstracts from human capital accumulation over the life-cycle, so that the reallocation behavior of young cohorts only reflects changes in aggregate conditions. Experience human capital would have opposite effects on our results depending on whether it is general or sector-specific. The accumulation of general human capital would increase the reallocation rate of young cohorts, leading us to overestimate the year component and underestimate the cohort component; sector-specific human capital would have the opposite effect, acting effectively as a mobility barrier, and hence being partially absorbed in our moving cost $f$.[34]

## VI. Beyond Accounting

This section presents additional evidence to go beyond the accounting results and shed light on two questions. First, how large is the equilibrium impact of human capital growth? Second, how important is the endogenous adjustment of human capital growth to the demand drivers of structural transformation?

### A. *General Equilibrium Effects*

The cohort component captures the magnitude of the supply shift associated with (both endogenous and exogenous) human capital growth. Quantifying the equilibrium impact of such shift requires going beyond the empirical decomposition and taking a stand on the parameters mediating the adjustment of relative prices. Combining Propositions 1 and 2, the overall impact of human capital growth on

[34] While our empirical approach is silent on the relative role of these two forces, we notice that estimates based on US data from Altonji, Smith, and Vidangos (2013) suggest that most experience human capital is general; similarly, Lee and Wolpin (2006) find that the degree of sectoral specificity of experience does not appear to be an important determinant of the relative size or growth of sectors. These results suggest that the cohort component might be a conservative estimate of the role of human capital growth.

labor reallocation is the product of the cohort component and the general equilibrium multiplier, as stated in the following Lemma.

LEMMA 3 (Equilibrium Impact of Human Capital Growth): *The equilibrium reallocation rate is given by*

$$\log g_{L_A} = (1 - \Theta_D)\log g_{\theta z} + \underbrace{(1 - \Theta_S)}_{\text{GE multiplier}} \times \underbrace{\log \tilde{\chi}}_{\text{cohort component}}, \tag{20}$$

*where* $(1 - \Theta_S) = \frac{1 - \eta\eta_H}{1 + \frac{\alpha\nu}{1-\gamma}}$.

PROOF:

See online Appendix D.

The multiplier depends on two sets of parameters. First, the parameters modulating general equilibrium adjustments in the labor market: the land share in agricultural production, $\alpha$, and the distributional parameter $\nu/(1 - \gamma)$, which is the elasticity of the agricultural labor supply to the relative wage, as showed in Lemma 1. The multiplier is decreasing in both; intuitively, a higher $\alpha$ implies a larger change in agricultural wages following a given shift in relative labor supply, while a higher $\nu/(1 - \gamma)$ implies a larger reallocation of labor following a given change in the relative wage. Second, the larger the elasticity of the agricultural price with respect to the human capital stock, $\eta\eta_H$, the more human capital growth is reflected in higher agricultural prices rather than lower agricultural employment. The GE multiplier is likely to vary across countries, as a function of their stage of development and openness to trade. While a country-specific quantification of the multiplier is beyond the scope of the paper, the next subsections propose two illustrative calculations under different sets of assumptions.

*Calibration with No Price Effects*.—Consider first an economy for which either $\eta = 0$ (i.e., a small open economy) or $\eta_H = 0$. In both cases, human capital growth does not have a direct impact on the relative agriculture price. For brevity, we will refer to both as "small open economy," although $\eta_H = 0$ would also apply to an economy in which the income and relative productivity effects of human capital growth roughly cancel out.

In this case, the GE multiplier only depends on $\alpha$ and $\nu/(1 - \gamma)$, which can be mapped into observables as follows. First, $\alpha$ corresponds to the land income share in agriculture. Valentinyi and Herrendorf (2008) estimate a land share of 18 percent in the United States. Land may play a larger role in low-income countries, where agricultural production is less capital intensive; for example, Gollin and Udry (2021) estimate production functions for micro plots in Uganda and Ghana and find land shares of 40–50 percent. We therefore consider values of $\alpha$ in the 0.18–0.5 range.

Second, we use information on wage dispersion in nonagriculture to bound $\nu/(1 - \gamma)$. The within-cohort variance of log nonagricultural wages implied by the model is

$$\mathrm{var}\left[\log w_{M,t}(c,\varepsilon)\right] = (1-\gamma)^2 \mathrm{var}\left[\log\varepsilon \,|\, \log\varepsilon \geq \log\hat{\varepsilon}_t(c)\right] \leq \left(\frac{1-\gamma}{\nu}\right)^2, \tag{21}$$

where the equality uses the equilibrium wage, and the inequality is due to the properties of the $\beta$ distribution.[35] The within-cohort standard deviation gives an upper bound for $v/(1-\gamma)$, which we can use to compute a lower bound for the GE multiplier (which, as discussed above, is decreasing in $v/(1-\gamma)$). While our dataset does not include wages for most countries, Lagakos et al. (2018) provided us with the value of the within-cohort standard deviation for each of the 18 countries in their sample, spanning the income distribution from Bangladesh to the United States.[36] The average standard deviation across these countries is 0.67, with no systematic correlation with GDP per capita. We therefore use $v/(1-\gamma) = 1/0.67 = 1.5$.

Combining the values for the two parameters, we find a GE multiplier ranging between 0.57 and 0.79, with low values in this range more likely to apply to low-income countries. This range implies that the inferred downward shift of the agricultural labor supply can account for between 20 percent and 30 percent of the observed rate of labor reallocation.

*The General Case: A Regression Approach*.—The calibration above leaves open the possibility that price effects might make the multiplier smaller in closed economies. As an alternative approach, we exploit the cross-country variation in the inferred cohort component to estimate the GE multiplier directly from equation (20). Intuitively, the larger the GE multiplier, the more larger cross-cohort gaps in agricultural employment (i.e., a more negative cohort component) should be reflected in a faster reallocation out of agriculture; at the extreme, if the GE multiplier is equal to 0 (i.e., $\Theta_s = 1$), larger cohort components would be compensated exactly by smaller year components, with no impact on the reallocation rate.

The key difficulty with this approach is the potential correlation between the cohort component and the $\log g_{\theta z}$ term. Indeed, the model does imply such correlation, given that, unless $\sigma$ is equal to 0, human capital growth responds endogenously to the growth in relative wages, which in turn depends on the growth in agricultural demand and relative productivity. To make progress, we control for observable proxies of $\log g_\theta$ and $\log g_z$, namely the growth rate in GDP per capita and in relative agricultural value added per worker.[37] In addition, we also control for the initial level of log GDP per capita to make sure that our results are not driven by the systematic differences between high- and low-income countries reported in Table 1.[38]

Column 1 of Table 4 shows that the estimated multiplier is 0.51, at the lower end of the range obtained from the calibration exercise. This value implies that

[35] If $\varepsilon \sim \beta(v,1)$, then $-\log\varepsilon \sim \mathrm{Exp}(v)$. Also, the variance of a truncated exponential is smaller than the unrestricted variance, which is $v^{-2}$.

[36] Refer to Lagakos et al. (2018) for data description and details. Wages are constructed as earnings divided by total hours of work in the period of observation, which is either weekly, monthly, or yearly. We drop the top and bottom 1 percent of wages to check that the variance estimates are not driven by outliers. For each country, we keep the most recent available cross-section.

[37] We use data on GDP per capita from the Maddison Project Database (Inklaar et al. 2018), and on real value added per worker by sector from the GGDC 10-Sector Database (Timmer, de Vries, and de Vries 2015), the Economic Transformation Database (de Vries et al. 2021), and the World Development Indicators (World Bank 2017). See online Appendix A for more details on the data construction. The results are very similar when we further correct the relative agricultural value added per worker measure to control for the role of land and human capital growth, as proxied by schooling; see online Appendix F for this and other robustness checks.

[38] This is a conservative specification, as high-income countries tend to have both faster reallocation and larger cohort components relative to low-income countries. Indeed, the specification without the control for initial GDP per capita, reported in online Appendix F, gives a somewhat larger GE multiplier.

human capital growth accounts for about 20 percent of labor reallocation. Column 2 replaces the cohort component with a direct measure of human capital growth, i.e., the change in average years of schooling of the working age population. Increases in schooling are negatively associated with growth in agricultural employment; interestingly, the coefficient is about half in magnitude compared to the one from the cohort-level regression in Table 2, consistently with the result that general equilibrium effects approximately halve the partial equilibrium effect of human capital on agricultural employment. Finally, to avoid the concern that noise or measurement error might generate a spurious correlation between the reallocation rate and the cohort component, column 3 uses as alternative regressor the cohort component predicted by cohort-level schooling and the empirical relationship between schooling and cohort effects. In particular, we compute

$$\text{Pred Cohort Component} \ = \ \frac{1}{k}\log\left[\frac{\sum_{c=t+k-N}^{t+k} n_{t+k,c,j}\exp\left(\hat{\mathbf{C}}_{c,j}^{S}\right)}{\sum_{c=t-N}^{t} n_{t,c,j}\exp\left(\hat{\mathbf{C}}_{c,j}^{S}\right)}\right],$$

where $\hat{\mathbf{C}}_{c,j}^{S} = \hat{\beta} s_{c,j}$, and $\hat{\beta} = -0.108$ is taken from Table 2; the resulting multiplier is very similar to the one in column 1.

*Evidence from Price Data*.—The similarity between the GE multiplier inferred from the small open economy calibration and the regression approach suggests limited price effects of human capital growth. To validate this result, we compute the growth in the relative agricultural price and examine its correlation with the cohort component.[39] As shown graphically in online Appendix F, the relative price has declined since 1960 in most countries, consistently with a secular decrease in the demand for agricultural goods and an increase in relative agricultural productivity; however, this decline has not been slower in countries with faster human capital growth (i.e., a more negative cohort component), as it would be implied by the model if $\eta\eta_H > 0$. Column 4 of Table 4 reports the regression results, including the controls for the growth in agricultural demand and relative productivity discussed above. The coefficient on the cohort component is positive, small and statistically insignificant.[40] This result is consistent with faster human capital growth not being associated with large and counteracting adjustments in the relative agricultural price.

[39] We compute sectoral prices as the ratios between nominal and real value added per worker. See online Appendix A for more details on the data construction, and online Appendix F for several robustness checks.

[40] Substituting the cohort component from Proposition 2 in (13), the growth in the relative agricultural price can be written as

$$g_p = \eta\log g_\theta + \eta\eta_z\log g_z - \eta\eta_H\frac{1-\gamma}{\nu}\log\tilde{\chi},$$

so that the model counterpart of the coefficient of the price regression is $-\eta\eta_H\frac{1-\gamma}{\nu}$. Given the calibration for $\nu/(1-\gamma)$ and $\alpha$ discussed above, a point estimate of 0.09 would imply that price adjustments would make the GE multiplier larger than the small open economy one by between 0.08 and 0.11. Considering a coefficient one standard deviation lower than the point estimate in Table 4 $(-0.23)$ would still imply a relatively small decline in the GE multiplier due to price effects, between 0.20 and 0.27 (compared to a small open economy multiplier in the 0.57–0.79 range).

Table 4—Estimating the GE Multiplier

| | $\Delta$ log agr employment | | | $\Delta$ log relative agr price |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Cohort component | 0.508 | | | 0.091 |
| | (0.210) | | | (0.322) |
| $\Delta$ log GDP p.c. | −0.109 | −0.091 | −0.090 | 0.158 |
| | (0.067) | (0.077) | (0.077) | (0.157) |
| $\Delta$ log relative agr prod | −0.087 | −0.065 | −0.066 | −0.158 |
| | (0.062) | (0.070) | (0.071) | (0.137) |
| log initial GDP p.c. | −0.004 | −0.005 | −0.005 | −0.009 |
| | (0.001) | (0.001) | (0.001) | (0.003) |
| $\Delta$ Avg yrs school | | −0.053 | | |
| | | (0.025) | | |
| Pred cohort component | | | 0.520 | |
| | | | (0.246) | |
| Observations | 46 | 46 | 46 | 46 |

*Note:* Robust standard errors in parentheses.

## B. *Endogenous Human Capital: Evidence from the Green Revolution*

The cohort component captures overall human capital growth, both exogenous and endogenous with respect to the demand forces behind structural transformation. As shown in Lemma 4, the exogenous and endogenous parts can be additively decomposed, with the latter being proportional to the year component. Intuitively, a negative year component reflects declining relative returns from working in agriculture over time; as long as $\sigma > 0$, this generates faster human capital growth, reflected in a more negative cohort component.

LEMMA 4 (Exogenous and Endogenous Human Capital Growth): *The cohort component can be written as*

$$\log\tilde{\chi} = -\frac{\gamma\nu}{1-\gamma}\log g_h = \underbrace{-\frac{\nu}{1-\gamma}\log g_\xi}_{\text{exogenous growth}} + \underbrace{\sigma\log\tilde{\psi}}_{\text{endogenous growth}}. \tag{22}$$

PROOF:

See online Appendix D.

What is the relative importance of the exogenous and endogenous parts of $\log\tilde{\chi}$? To make progress, we leverage the agricultural employment and schooling effects of the Green Revolution, as recently studied in Gollin, Hansen, and Wingender (2021). Starting in the 1960s, the introduction of modern crop breeding techniques led to a large and persistent increase in agricultural productivity growth. As shown in Gollin, Hansen, and Wingender (2021), countries were differentially exposed to this wave of innovation depending on the preexisting crop composition, and more exposed countries saw lower agricultural employment growth and larger income and schooling growth in the following decades. Under the assumption that the preexisting crop composition is uncorrelated with other drivers of human capital growth, we can obtain an estimate of $\sigma$ based on the schooling response to this decrease in agricultural labor demand.

More specifically, we follow Gollin, Hansen, and Wingender (2021) in constructing the predicted extra agricultural yields due to the Green Revolution as the sum of the crop-specific productivity increase induced by the use of high-yielding varieties, weighted by the preexisting crop shares. We then regress this country $\times$ year measure of agricultural productivity on aggregate agricultural employment on one hand, and lifetime years of schooling of the cohorts that started schooling at that time on the other.[41] Columns 1 and 2 of Table 5 display the results: countries with faster yield growth due to the Green Revolution saw faster decline in agricultural employment and growth in schooling of the affected cohorts.[42]

These results suggest that human capital does endogenously increase in response to a decline in the demand for agricultural labor; i.e., $\sigma > 0$ in our model. We can compute the implied magnitude for $\sigma$ by noticing from (22) and the identity $\log g_{L_A} = \log\tilde{\chi} + \log\tilde{\psi}$ that the elasticities of agricultural employment and the cohort component with respect to predicted yields satisfy

$$\frac{\partial \log\tilde{\chi}}{\partial \log \text{Predicted Yields}} = \frac{\sigma}{1+\sigma}\frac{\partial \log g_{L_A}}{\partial \log \text{Predicted Yields}}.$$

Intuitively, the larger the effect on cohort-level human capital relative to the aggregate effect on agricultural employment, the larger $\sigma$ must be. To implement this calculation, we convert the estimated schooling effect into the implied cohort component by multiplying the former by $-0.10$, the estimated coefficient in Table 2. We find $\sigma = 0.24$; plugging the average cohort and year components from Table 1 into (22), this value of $\sigma$ implies $\frac{\gamma\nu}{1-\gamma}\log g_{\xi} = 0.47$. Endogenous human capital growth can therefore account for about 40 percent of the cohort component, with the remaining 60 percent $(0.47/0.79)$ being attributed to the exogenous human capital shifter. While the magnitudes are suggestive, we conclude that both margins are likely to be quantitatively important.[43]

### C. *Summary and Discussion*

The analysis in this section gives two new insights. First, general equilibrium adjustments in relative agricultural prices and wages plausibly halve the partial equilibrium effect of the supply shift identified by the cohort component. We thus

[41] Given the limited overlap with the IPUMS sample used in the rest of the paper, we rely here on the sample and data sources in Gollin, Hansen, and Wingender (2021). Agricultural employment comes from Wingender (2014), while the schooling variable is the average lifetime years of schooling of the individuals aged five to ten from Barro and Lee (2013). We use the 1955–1995 sample period, and include all countries for which the required data are available (86 countries for the agricultural employment regression, 77 for the schooling one).

[42] Online Appendix F provides a graphical visualization of this result by plotting the estimates from event study specifications around the start of the Green Revolution.

[43] In online Appendix E, we build a quantitative model with a micro-founded cohort-level human capital choice. We estimate the model to target the effect of the green revolution on schooling and GDP and we reach similar conclusions: changes in the exogenous component of human capital account for at least half of the overall increase across cohorts.

TABLE 5—THE GREEN REVOLUTION AND ENDOGENOUS HUMAN CAPITAL

| | log agr employment (1) | Years school (2) |
|---|---|---|
| log predicted yields | −1.314 | 2.569 |
| | (0.337) | (1.468) |
| Country fixed effects | Yes | Yes |
| Year fixed effects | Yes | Yes |
| Observations | 3,471 | 3,147 |

*Note:* Robust standard errors in parentheses.

conclude that if human capital growth (either endogenous or exogenous) did not take place, we would have observed 20 percent less labor reallocation out of agriculture. These numbers should be interpreted with caution: as well known, obtaining reliable estimates for general equilibrium effects is challenging since it usually requires to exploit the scarce cross-country empirical variation. In light of this inherent uncertainty, in online Appendix E, we report results from alternative calibrations of a quantitative model. Our benchmark results are consistent with a broad range of parameters.

Second, the endogenous response to the demand forces behind structural transformation and other exogenous factors both contributed significantly to human capital growth and the resulting shift in the supply of agricultural labor. Overall, we conclude that human capital growth is both an independent driver of structural transformation as well as a mediating factor of the demand forces emphasized in the literature.

## VII. Conclusion

This paper explores the hypothesis that the steep increase in human capital during the 20th century contributed to the process of structural transformation, by equipping the new generations of workers with skills more useful outside of the agricultural sector.

We use theory and evidence to support this hypothesis. Drawing on micro data from many countries at different levels of development, we document that a large part of the aggregate rate of labor reallocation out of agriculture was driven by new cohorts entering the labor market, as opposed to movements across sectors for given cohorts. Using information on cohort-specific educational attainment and a newly compiled dataset on educational reforms and other relevant political events, we show that the increase in schooling for more recent cohorts contributed to their lower agricultural employment. A model of frictional labor reallocation out of agriculture provides a structural interpretation of the empirical results, suggesting that human capital growth led to a sharp downward shift in the agricultural labor supply, accounting for about 40 percent of labor reallocation when keeping prices fixed. Guided by the model, we estimate that general equilibrium forces roughly halve the partial equilibrium effect of human capital growth. Importantly, we show that exogenous factors and the endogenous response to the demand forces behind structural transformation contributed in similar magnitude to the overall increase in human capital.

We emphasize two important implications of these results. First, while theories of structural change typically focus on factors decreasing the demand for agricultural labor, supply-side changes in the workforce composition and skills—what we call the “human side” of structural transformation—are quantitatively important. Second, to the extent that human capital growth can be promoted by increased access to schooling and educational reforms, these policies should be considered potential tools to accelerate the process of structural transformation.

## REFERENCES

**Acemoglu, Daron, and Veronica Guerrieri**. 2008. “Capital Deepening and Nonbalanced Economic Growth.” *Journal of Political Economy* 116 (3): 467–98.

**Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James A. Robinson**. 2019. “Democracy Does Cause Growth.” *Journal of Political Economy* 127 (1): 47–100.

**Aghion, Philippe, Fabrizio Zilibotti, Michael Peters, and Robin Burgess**. 2019. “Structural Change and Service-led Growth in India.” *Society for Economic Dynamics 2019 Meeting Papers* 393.

**Altonji, Joseph G., Anthony A. Smith Jr, and Ivan Vidangos**. 2013. “Modeling Earnings Dynamics.” *Econometrica* 81 (4): 1395–1454.

**Alvarez-Cuadrado, Francisco, and Markus Poschke**. 2011. “Structural Change Out of Agriculture: Labor Push versus Labor Pull.” *American Economic Journal: Macroeconomics* 3 (3): 127–58.

**Alvarez, Jorge A**. 2020. “The Agricultural Wage Gap: Evidence from Brazilian Micro-Data.” *American Economic Journal: Macroeconomics* 12 (1): 153–73.

**Barro, Robert J**. 1991. “Economic Growth in a Cross Section of Countries.” *Quarterly Journal of Economics* 106 (2): 407–43.

**Barro, Robert J., and Jong Wha Lee**. 2013. “A New Data Set of Educational Attainment in the World, 1950-2010.” *Journal of Development Economics* 104 (C): 184–98.

**Boppart, Timo**. 2014. “Structural Change and The Kaldor Facts in a Growth Model with Relative Price Effects and Non-Gorman Preferences.” *Econometrica* 82 (6): 2167–96.

**Boyle, Elizabeth Heger, Miriam King, and Matthew Sobek**. 2019. “IPUMS-Demographic and Health Surveys: Version 7 [dataset].” Minnesota Population Center and ICF International https://doi.org/10.18128/D080.V7.

**Bryan, Gharad, and Melanie Morten**. 2019. “The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia.” *Journal of Political Economy* 127 (5): 2229–68.

**Card, David, Jörg Heining, and Patrick Kline**. 2013. “Workplace Heterogeneity and the Rise of West German Wage Inequality.” *Quarterly Journal of Economics* 128 (3): 967–1015.

**Caselli, Francesco**. 2005. “Accounting for Cross-Country Income Differences.” In *Handbook of Economic Growth*, edited by Philippe Aghion and Steven N. Durlauf, 679–741. Amsterdam: Elsevier.

**Caselli, Francesco, and Wilbur John Coleman II**. 2001. “The U.S. Structural Transformation and Regional Convergence: A Reinterpretation.” *Journal of Political Economy* 109 (3): 584–616.

**Chernozhukov, Victor, and Han Hong**. 2003. “An MCMC Approach to Classical Estimation.” *Journal of Econometrics* 115 (2): 293–346.

**Cociuba, Simona E., and James MacGee**. 2018. *Demographics and Sectoral Reallocations: A Search Theory with Immobile Worker*s. London, Ontario, Canada: Western University.

**Comin, Diego A., Danial Lashkari, and Martí Mestieri**. 2015. “Structural Change with Long-run Income and Price Effects.” NBER Working Paper 21595.

**Deaton, Angus**. 1997. *The Analysis of Household Surveys.* Baltimore, MD: Johns Hopkins University Press.

**DeLong, J. Bradford, Claudia Goldin, and Lawrence F. Katz**. 2003. “Sustaining US Economic Growth.” In *Agenda for the Nation,* edited by Henry J. Aaron, James M. Lindsay, and Pietro S. Nivola, 17–60. Washington, DC: Brookings Institution Press.

**de Vries, Gaaitzen, Linda Arfelt, Dorothea Drees, Mareike Godemann, Calumn Hamilton, Bente Jessen-Thiesen, Ahmet Ihsan Kaya, Hagen Kruse, Emmanuel Mensah, and Pieter Woltjer**. 2021. *The Economic Transformation Database* (*ETD*)*: Content, Sources, and Methods*. Tokyo: United Nations University.

**Duflo, Esther**. 2001. “Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment.” *American Economic Review* 91 (4): 795–813.

**Duflo, Esther**. 2004. “The Medium Run Effects of Educational Expansion: Evidence from a Large School Construction Program in Indonesia.” *Journal of Development Economics* 74 (1): 163–97.

**Eckert, Fabian, and Michael Peters**. 2018. "Spatial Structural Change." Unpublished.
**Galor, Oded**. 2005. "From Stagnation to Growth: Unified Growth Theory." In *Handbook of Economic Growth,* edited by Philippe Aghion and Steven N. Durlauf, 171–293. Amsterdam: Elsevier.
**Galor, Oded, and Andrew Mountford**. 2008. "Trading Population for Productivity: Theory and Evidence." *Review of Economic Studies* 75 (4): 1143–79.
**Galor, Oded, Omer Moav, and Dietrich Vollrath**. 2009. "Inequality in Landownership, the Emergence of Human-Capital Promoting Institutions, and the Great Divergence." *Review of Economic Studies* 76 (1): 143–79.
**Gennaioli, Nicola, Rafael La Porta, Florencio Lopez de Silanes, and Andrei Shleifer**. 2013. "Human Capital and Regional Development." *Quarterly Journal of Economics* 128 (1): 105–64.
**Glaeser, Edward L., Rafael La Porta, Andrei Lopez-de-Silanes, and Florencioand Shleifer**. 2004. "Do Institutions Cause Growth?" *Journal of Economic Growth* 9 (3): 271–303.
**Gollin, Douglas, and Christopher Udry**. 2021. "Heterogeneity, Measurement Error, and Misallocation: Evidence from African Agriculture." *Journal of Political Economy* 129 (1): 1–80.
**Gollin, Douglas, Casper Worm Hansen, and Asger Mose Wingender**. 2021. "Two Blades of Grass: The Impact of the Green Revolution." *Journal of Political Economy* 129 (8): 2344–84.
**Gollin, Douglas, David Lagakos, and Michael E. Waugh**. 2014. "The Agricultural Productivity Gap." *Quarterly Journal of Economics* 129 (2): 939–93.
**Heise, Sebastian, and Tommaso Porzio**. 2021. "The Aggregate and Distributional Effects of Spatial Frictions." NBER Working Paper 28792.
**Herrendorf, Berthold, Richard Rogerson, and Ákos Valentinyi**. 2014. "Growth and Structural Transformation." In *Handbook of Economic Growth,* Vol. 2, edited by Philippe Aghion and Steven N. Durlauf, 855–941. Amsterdam: Elsevier.
**Herrendorf, Berthold, and Todd Schoellman**. 2015. "Why is Measured Productivity so Low in Agriculture?" *Review of Economic Dynamics* 18 (4): 1003–22.
**Herrendorf, Berthold, and Todd Schoellman**. 2018. "Wages, Human Capital, and Barriers to Structural Transformation." *American Economic Journal: Macroeconomics* 10 (2): 1–23.
**Hicks, Joan Hamory, Marieke Kleemans, Nicholas Y. Li, and Edward Miguel**. 2017. "Reevaluating Agricultural Productivity Gaps with Longitudinal Microdata." NBER Working Paper 23253.
**Hobijn, Bart, Todd Schoellman, and Alberto Vindas**. 2019. "Structural Transformation by Cohort." Unpublished.
**Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow**. 2019. "The Allocation of Talent and U.S. Economic Growth." *Econometrica* 87 (5): 1439–74.
**Inklaar, Robert, Harmen de Jong, Jutta Bolt, and Jan Van Zanden**. 2018. "Rebasing 'Maddison': The Shape of Long-Run Economic Development." *VoxEU,* January 25.
**Karachiwalla, Naureen, and Giordano Palloni**. 2019. "Human Capital and Structural Transformation: Quasi-Experimental Evidence from Indonesia." International Food Policy Research Institute Discussion Paper 01836.
**Kim, Dae-II, and Robert H. Topel**. 1995. "Labor Markets and Economic Growth: Lessons from Korea's industrialization, 1970-1990." In *Differences and Changes in Wage Structures*, edited by Richard B. Freeman and Lawrence F. Katz, 227–64. Chicago: University of Chicago Press.
**King, Miriam, Steven Ruggles, J. Trent Alexander, Sarah Flood, Katie Genadek, Matthew B. Schroeder, Brandon Trampe, and Rebecca Vick.** 2019. "Integrated Public Use Microdata Series, International: Version 7.2 [dataset]." Ipums International, https://doi.org/10.18128/D020.V7.2.
**Kongsamut, Piyabha, Sergio Rebelo, and Danyang Xie**. 2001. "Beyond Balanced Growth." *Review of Economic Studies* 68 (4): 869–82.
**Lagakos, David, Mushfiq Mobarak, and Michael E. Waugh**. 2019. "The Welfare Effects of Encouraging Rural-Urban Migration." Unpublished.
**Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman**. 2018. "Life-Cycle Wage Growth Across Countries." *Journal of Political Economy* 126 (2): 797–849.
**Lagakos, David, and Michael E. Waugh**. 2013. "Selection, Agriculture, and Cross-Country Productivity Differences." *American Economic Review* 103 (2): 948–80.
**Lee, Donghoon, and Kenneth I. Wolpin**. 2006. "Intersectoral Labor Mobility and the Growth of the Service Sector." *Econometrica* 74 (1): 1–46.
**Lucas, Jr, Robert E.** 2004. "Life Earnings and Rural-Urban Migration." *Journal of Political Economy* 112 (1): S29–59.
**Mankiw, N. Gregory, David Romer, and David N. Weil**. 1992. "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics* 107 (2): 407–37.
**Matsuyama, Kiminori**. 1992a. "Agricultural Productivity, Comparative Advantage, and Economic Growth." *Journal of Economic Theory* 58 (2): 317–34.

**Matsuyama, Kiminori**. 1992b. "A Simple Model of Sectoral Adjustment." *Review of Economic Studies* 59 (2): 375–88.
**Morten, Melanie**. 2019. "Temporary Migration and Endogenous Risk Sharing in Village India." *Journal of Political Economy* 127 (1): 1–46
**Munshi, Kaivan, and Mark Rosenzweig**. 2016. "Networks and Misallocation: Insurance, Migration, and the Rural-urban Wage Gap." *American Economic Review* 106 (1): 46–98.
**Nakamura, Emi, Jósef Sigurdsson, and Jón Steinsson**. 2022. "The Gift of Moving: Intergenerational Consequences of a Mobility Shock." *Review of Economic Studies* 89 (3): 1557–92.
**Nakamura, Emi, and Jón Steinsson**. 2018. "Identification in Macroeconomics." *Journal of Economic Perspectives* 32 (3): 59–86.
**Nelson, Richard R., and Edmund S. Phelps**. 1966. "Investment in Humans, Technological Diffusion, and Economic Growth." *American Economic Review* 56: 69–75.
**Ngai, L. Rachel, and Christopher A. Pissarides**. 2007. "Structural Change in a Multisector Model of Growth." *American Economic Review* 97 (1): 429–43
**Ngai, L. Rachel, Christopher A. Pissarides, and Jin Wang**. 2019. "China's Mobility Barriers and Employment Allocations." *Journal of the European Economic Association* 17 (5): 1617–53.
**Perez, Santiago**. 2017. "Railroads and the Rural to Urban Transition: Evidence from 19th-Century Argentina." Unpublished.
**Porzio, Tommaso**. 2017. "Cross-Country Differences in the Optimal Allocation of Talent and Technology." Unpublished.
**Porzio, Tommaso, Federico Rossi, and Gabriella Santangelo.** 2022. "Replication data for: The Human Side of Structural Transformation."American Economic Association [publisher], Inter-university Consortium for Political and Social Research [Distributor]. https://doi.org/10.3886/E165042V1.
**Rossi, Federico**. 2018. "Human Capital and Macro-Economic Development: A Review of the Evidence."*World Bank Research Observer* 35 (2): 227–62.
**Rybczynski, Tadeusz M**. 1955. "Factor Endowment and Relative Commodity Prices." *Economica* 22 (88): 336–41.
**Storesletten, Kjetil, Bo Zhao, and Fabrizio Zilibotti**. 2019. "Business Cycle during Structural Change: Arthur Lewis' Theory from a Neoclassical Perspective." NBER Working Paper 26181.
**Timmer, Marcel, Gaaitzen J. de Vries, and Klaas de Vries**. 2015. "Patterns of Structural Change in Developing Countries." In *Routledge Handbook of Industry and Development*, edited by John Weiss and Michael Tribe, 65–83. New York: Routledge.
**Valencia Caicedo, Felipe**. 2018. "The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America." *Quarterly Journal of Economics* 134 (1): 507–56.
**Valentinyi, Akos, and Berthold Herrendorf**. 2008. "Measuring Factor Income Shares at the Sectoral Level." *Review of Economic Dynamics* 11 (4): 820–35
**Wingender, Asger Moll**. 2014. "Structural Transformation in the 20th century: A New Database on Agricultural Employment Around the World." University of Copenhagen Department of Economics Discussion Papers 14-28.
**World Bank**. 2017. *World Development Indicators 2017*. Washington, DC: World Bank Group.
**Young, Alwyn**. 2013. "Inequality, the Urban-Rural Gap, and Migration." *Quarterly Journal of Economics* 128 (4): 1727–85.