

NBER WORKING PAPER SERIES

THE LONG RUN EFFECTS OF LABOR MIGRATION ON HUMAN CAPITAL FORMATION
IN COMMUNITIES OF ORIGIN

Taryn Dinkelman
Martine Mariotti

Working Paper 22049
<http://www.nber.org/papers/w22049>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2016

Freed Kumchulesi, Grace Kumchulesi, Ashley Wong, Lucy Xie and Zheng-Yi Yang provided excellent research assistance for this project. We also thank officials at the Malawi National Archives and Lucy McCann at Rhodes House Library for their invaluable assistance in data collection. We thank Michael Clemens, Elizabeth Cascio, Alan de Brauw, Eric Edmonds, James Fenske, David McKenzie, Caroline Theoharides, Rebecca Thornton, Steve Stillman, Dean Yang, and seminar participants at the Australian National University, the 2014 Australian and Pacific Economic and Business History Meetings, the 2014 Australasian Development Economics Workshop, the 2014 Barcelona GSE Summer Forum (Development and Migration Workshops), the 2014 Economic Demography Workshop at the Population Association of America meetings, Georgetown University, the 2015 NBER Summer Institute (Children/Labor Studies), The 2015 World Economic History Congress, Oxford University, the Paris and Toulouse Schools of Economics, the University of Namur and Williams College for comments. This document is an output from a project funded by the UK Department for International Development (DFID) and the Institute for the Study of Labor (IZA) for the benefit of developing countries (GA-C2-RA4-181). The views expressed are not necessarily those of DFID, IZA, or the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by Taryn Dinkelman and Martine Mariotti. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Long Run Effects of Labor Migration on Human Capital Formation in Communities of Origin

Taryn Dinkelman and Martine Mariotti

NBER Working Paper No. 22049

February 2016

JEL No. F22,F24,N37,O12,O15,O55

ABSTRACT

We provide new evidence of one channel through which circular labor migration has long run effects on origin communities: by raising completed human capital of the next generation. We estimate the net effects of migration from Malawi to South African mines using newly digitized Census and administrative data on access to mine jobs, a difference-in-differences strategy and two opposite-signed and plausibly exogenous shocks to the option to migrate. Twenty years after these shocks, human capital is 4.8-6.9% higher among cohorts who were eligible for schooling in communities with the easiest access to migrant jobs.

Taryn Dinkelman

Department of Economics

H. Box 6106

Dartmouth College

Hanover, NH 03755

and NBER

taryn.l.dinkelman@dartmouth.edu

Martine Mariotti

Research School of Economics

H W Arndt Building 25a

Australian National University

Acton

ACT 2601

Australia

martine.mariotti@anu.edu.au

1. Introduction

Many economists recognize that migration in some form or other is conceptually one of the most direct strategies that the poor can use to improve their living standards (Pritchett 2006, Banerjee and Duflo 2007, Clemens 2011). Indeed, temporary, or circular, migration across borders has long been a feature of labor markets in developing countries, particularly in sub-Saharan Africa. And, although a large body of work documents the short run impacts of international migration and access to migrant remittances on individual and household-level outcomes,² considerable debate remains over whether and how migration affects communities of origin over the long run (Lucas, 2005, Constant and Zimmerman 2013).

This paper provides the first evidence that temporary labor migration raises human capital formation in origin communities over the long run. In order to do this, we address two empirical issues that have constrained what researchers have been able to learn about the lasting impacts of labor migration on growth and development outcomes in origin communities. First, there is a general lack of high quality data on migrant and remittance flows in developing countries. Second, even with the right data in hand, researchers confront difficult identification issues associated with comparing outcomes across high and low migration areas (Clemens and McKenzie 2014). For example, high migration regions are likely to differ from low migration regions in ways that are highly correlated with economic outcomes of interest (Gibson, McKenzie and Stillman 2013). Researchers have cleverly used visa lottery designs and quasi-experimental variation in the value of remittances deriving from unanticipated exchange rate shocks to estimate how labor migration affects sending households in the short and medium-runs. But these designs are not well suited to estimating long run effects of migration on entire communities of origin.

We study the long run effects of differential labor migration flows from Malawi to South Africa's gold mines, leveraging new data on plausibly exogenous shocks to historical international migration between these two countries. During the twentieth century, Malawians faced historical differences in the costs of accessing migrant jobs across space within their country, and encountered time-varying exogenous shocks to the option to migrate for these jobs. We implement a difference-in-differences analysis by combining newly digitized Census and administrative data that capture migrant flows at sub-national level over time with pre-existing spatial variation in migration costs, exogenous shocks to labor migration and migrant earnings received by sending communities, and measures of long run educational attainment in these

² Important recent work in this area has been done by Yang (2008), Antman (2011, 2012), Beegle, De Weerd and Dercon (2011), Gibson, McKenzie and Stillman (2011, 2013), Theoharides (2014) and Kosack (2015). These papers pay careful attention to identification concerns in settings where migrants are a selected group of individuals.

communities. We are not aware of any other dataset that would allow a similar exploration of the impact of migration on such long run community-level outcomes.³

Malawi provides a uniquely advantageous setting for answering our research question. Throughout the twentieth century, it was one of the primary suppliers of unskilled male workers for the South African Chamber of Mines. By the late 1970s, one in five Malawian men had worked abroad, typically on a series of two-year labor contracts that diverted up to two thirds of earnings – called deferred pay – to be paid upon return to Malawi. These men could not choose when to return, nor could they choose the amount of deferred pay to send. By entering into an institutionalized system of circular labor migration, the Malawian government could be certain that a large share of incomes earned abroad would return home, and that labor resources would not be forever lost.

Our approach exploits two sides of a natural experiment that generated exogenous variation in the opportunity to migrate internationally. In 1967, a new labor treaty eliminated historical recruiting quotas in Malawi, leading to an immediate three-fold increase in migration to South African mines. Seven years later, Malawi's president banned new recruiting and recalled all mineworkers, after a mining plane crash in April 1974 killed a group of Malawian workers. The ban only remained in place until 1977, but labor migration never recovered to pre-1974 levels. The timing of these two events interacted with pre-existing differences in migration costs at the district level generate differential childhood exposure to labor migration shocks and associated migrant income. This exogenous and differential exposure forms the basis of our identification strategy.

The first comparison in our difference-in-differences strategy estimates gaps in the human capital of cohorts who were age-eligible for primary school after the migration expansion, between districts with and without access to historically-placed mine recruiting stations. Our second difference uses control cohorts who are too old to be in primary school after the migration expansion to net out any pre-existing differences in educational attainment across recruiting and non-recruiting districts. We use newly digitized Census data from 1977 and 1998 along with administrative data on the historical location of recruiting stations to estimate these models, controlling for cohort and district fixed effects, and for linear trends interacted with historical district-level covariates.

³ It is typically impossible to capture migration prevalence at district-level in general household surveys (De Brauw, Mueller and Lee 2013). Woodruff (2007) discusses studies that use household survey or Census data to estimate the causal effects of current labor migration on contemporaneous outcomes in origin communities in Mexico, using historical migration networks to instrument for current migration.

Our focus on long run outcomes for the community is important. Most existing research focuses on measuring the causal impacts of migration or remittances on contemporaneous school outcomes, like enrolment, or grade attained by the end of a specific year, for specific individuals and families, often in households with a migrant worker. We have two distinct factors in mind regarding the long run impacts of migration on the education profiles of communities. First, impacts of labor migration on child enrolment in or access to school in the short run could generate long run differences in total human capital within a community, long after the end of the migration episodes, as long as effects are large enough for specific cohorts within the community. Second, migration shocks may have persistent impacts on origin communities if households continue to make different schooling choices even after these shocks have subsided. For example, families may continue to invest more in their children, and in new cohorts of children, even after the end of labor migration. School enrolment at critical ages may also make it more worthwhile to continue to higher levels of education. In these cases, persistence could generate larger long run impacts of migration on education in sending communities.

We find some evidence for both types of long run impacts in Malawi. Twenty years after the migration shocks subsided, exposed cohorts in recruiting station districts have between 0.12 and 0.18 more years of schooling. Given low average levels of completed education in this population – 2.5 years – the impacts we estimate represent between 4.8 and 6.9% more education. In addition, we show that out of every 100 adults in the most exposed districts and cohorts, between 1 and 2 more adults attain any primary schooling, off a base of about 40 adults with any primary education.⁴ Consistently, we find that the effect of the migration shock in the labor ban years, after all migrants return and collect their deferred pay, is significantly larger than the effect of the migration shock during the labor expansion period, when fewer migrants are returning to Malawi each year. The impact of the migration shock persists beyond the end of the labor ban period, although with smaller positive impacts on the human capital attainment of the youngest cohorts. We also find some evidence that long run impacts are largest in areas where the opportunity cost of going to school is lowest.

Both theory and existing empirical evidence on the short run impacts of migration suggest that our results are not obvious (Antman 2011, 2012, 2013). Researchers have found both positive and negative impacts of labor migration on the next generation's human capital formation, although the vast majority of this evidence is from non-African settings.⁵ For example: access to remittance income from migration could

⁴ Importantly, we find no differences in the size of these education impacts between males and females (results not shown). Because mining jobs were restricted to men, it seems unlikely aspirations about earning a higher return to this education as a future labor migrant explain these effects.

⁵ For example, Cox-Edwards and Ureta (2003) and Yang (2008) show that migrant remittances increase education spending and contemporaneous outcomes among children in El Salvador and the Philippines. Antman (2012) and

relax credit constraints, thereby directly reducing barriers to schooling (e.g. Yang 2008, Hanson and Woodruff 2003). Expected wage differentials between migrant and non-migrant jobs may incentivize kids to stay in school as the “brain gain” hypothesis suggests (Mountford 1997; for an empirical test see Batista, Lacueste and Vicente 2012) or may undermine human capital accumulation by encouraging school drop-out (De Brauw and Giles, 2006; Gibson, McKenzie and Stillman 2011; McKenzie and Rapoport 2011). Losing a prime-aged adult, even temporarily, may crowd out schooling and cause families to use child labor instead; it may also have direct negative impacts on child performance in school. At the same time, depending on who migrates, changing bargaining power within the household could increase or decrease investments in education (Antman 2012). At an aggregate level, large labor migration flows may also have important effects on wages and prices in general equilibrium, generating positive or negative spillovers for education decisions in non-migrant households in the same communities. And while short run gains in enrolment or attendance cannot be undone, long run effects could end up being larger, or could end up affecting new cohorts, if attainment gains cumulate and school investments persist beyond the end of migration episodes.

In our setting, the positive impacts of labor migration on human capital investments appear to dominate any negative impacts, and the aggregate impact of migration and migrant earnings is large enough to alter human capital profiles of exposed cohorts in origin communities over the long run. The elasticities we estimate for cohorts exposed during the migration shocks are quantitatively smaller than the most credible estimates in the literature on the impact of remittances and labor demand shocks on contemporaneous school enrolments (e.g. Yang 2008; Theoharides 2014). However, they suggest that we can generalize the patterns of positive short run results to the long run in an African setting, despite the different context for schooling and child labor decisions.

To motivate our empirical strategy, we first discuss the background to mine migration from Malawi and local alternatives to migration. In Section 3, we describe the context of education in Malawi during the pre- and post-colonial period. Section 4 describes our data while Section 5 presents the empirical strategy and addresses potential threats to validity. Section 6 presents our results and Section 7 concludes.

2. Labor Migration from Malawi: Background and Context

Malawi has a long tradition of cross-border labor migration, facilitated by its central location in Africa, high population densities, lack of natural resources, and few non-agricultural economic opportunities.

Hanson and Woodruff (2003) use different empirical strategies to show that parental migration to the US increases schooling among Mexican girls. In contrast, Antman (2011) and McKenzie and Rapoport (2011) respectively show that parental migration increases work and reduces schooling in the short run among Mexican boys, and provides an option to migrate that reduces secondary schooling among male youth in Mexico.

Early in the colonial period, a bureaucratic infrastructure for labor exports was established. Our analysis takes advantage of massive and largely unexpected fluctuations in labor exports and reverse flows of money in the post-colonial period. This section describes the early establishment of the system of mine labor recruitment in Malawi and the sources of exogenous shocks to migration between 1950 and 1990, and then compares earnings from mine work with domestic earnings opportunities.

Circular labor migration from Malawi and labor migration shocks

Throughout the twentieth century, Malawians took advantage of employment opportunities in South Africa and Rhodesia (Zimbabwe and Zambia) to boost local incomes, although demand for workers in Rhodesia had slowed significantly by the early 1960s. To facilitate access to cheap sources of foreign labor, the South African mining industry established a physical presence in Colonial Malawi (then Nyasaland), opening and operating a large network of recruiting stations by the 1930s. The Witwatersrand Native Labour Agency (WNLA, or Wenela), the South African Chamber of Mines' centralized labor recruitment organization, was responsible for coordinating all recruitment activities through recruiting stations outside of South Africa (Crush et al. 1991, pg 40; Jeeves 1987).⁶

We collected and digitized administrative data on the location of these Wenela stations as of 1937 and show their prevalence across the country in Figure 1. The thick black borders in the figure represent district boundaries, the red hatched areas indicate districts with a Wenela station and unshaded areas show sub-districts without a Wenela station.

Between 1950 and 1990, the number of Malawians working on South African gold mines rose from just over 10,000 men per year to a high of 120,000 men per year, and back down to almost zero (see Figure 2). To put these numbers in perspective: by 1977, one in five adult males had ever worked abroad. Most of these mineworkers were engaged on two-year contracts, after which they had to return home for some time before being allowed to reengage for subsequent contracts (Wilson 1972, pg 68, Prothero 1974 and Lucas 1985). The nature of these contracts meant that mine migration was longer term and circular, rather than seasonal in nature.

⁶ Decisions about where and when to set up recruiting practices in southern Africa in the first half of the 1900s were spearheaded by the mining industry's "labor czar", William Gemmill (Jeeves 1987). Gemmill presided over the expansion of Wenela into colonial Malawi during the 1920s and 1930s at a time when local agriculture was struggling with the Great Depression and the collapse of tobacco prices. He looked towards "Tropical Areas" to recruit labor, knowing that local employers could not compete with mine wages. (McCracken 2012, Jeeves 1987). Wilson (1972) and Lucas (1985) provide accounts of how Wenela and the local Native Recruitment Corporation (NRC) operated as labor monopsonists in the colonies, keeping wages low by hiring workers from Mozambique, Northern and Southern Rhodesia, Nyasaland, Lesotho, Swaziland, Tanzania, Angola, Botswana and South Africa.

The costs associated with signing up for mine work were not negligible, and the complex logistics of getting access to a mining job in South Africa involved many steps.⁷ A potential migrant needed to obtain official verification of no outstanding tax obligations from the local chief; then he needed to get similar approval from the local tax authority; following which he had to travel to a Wenela station. At the station, he had to pass a medical examination (mainly regarding a minimum weight requirement) and get ‘attested’ (approved for travel), after which he delivered the attestation documents back to the local district officer for processing of his foreign travel documents. The final step involved returning to the Wenela station to await transportation to a main Wenela depot for transfer to South Africa.⁸

In our empirical strategy, we use spatial variation in the number of Wenela stations in a district as a proxy for (lower) labor migration costs that are exogenous to individuals and plausibly exogenous to the district, conditional on a set of baseline control variables. Our reading of the historical literature suggests that decisions about location placement were made after simple visual inspection of the potential of these areas for recruiting, without much regard to local economic conditions other than the labor supply potential of districts. Practically, there would have been little available hard data to influence these location decisions made in the late 1920s and early 1930s. We explore the correlates of recruiting station placement further in Section 4.

Interacting with the spatial distribution of migration costs were two important exogenous shocks to the opportunity to migrate to South Africa. Until 1967, national labor quotas restricted the number of workers Wenela could recruit from Nyasaland to a few thousand workers per year. European plantation owners in Nyasaland and in Southern Rhodesia (Zimbabwe) lobbied for these restrictions as a way to protect access to a cheap source of labor (Paton 1995, pg 46; Jeeves 1987). Between 1946 and 1959, Wenela quota increased from 8,500 men to 20,000 men, which represented roughly 2% of the target population of working age males (Coleman 1972, Chirwa, 1992). However, by the 1960s opportunities for employment in Southern Rhodesia began to decline due to circumstances in that country (Clarke 1977, p. 31-32, Paton 1995, p. 47-48) and in 1967, Malawi’s President Banda signed a new agreement with Wenela removing all quotas on recruitment of Malawian workers (Treaty Series No. 10/1967). Figure 2 shows the 200% increase in the number of Wenela workers in the seven years following the abolition of quotas.

⁷ This section draws on original colonial documents retrieved at Malawi’s National Archives, including Governor’s Memorandum on Labor Migrancy in Malawi (1956) and Provincial Office Memo (December 7 1961) and from Prothero (1974).

⁸ Sources suggest that health and ability of measures we only used to sort workers into jobs, and not used to screen Malawians out of mine work entirely. Health restrictions on migrant workers were minimal and workers were recruited for many types of jobs. Once on the mines, a simple colored shape sorting test sorted workers into mechanical or non-mechanical jobs; a physical examination of eyesight, weight, height and stress-testing further sorted workers into very heavy, heavy or light work (Weyl 1981: p. 16).

The 1967 labor agreement required Wenela to withhold two thirds of miner wages until the miners' return to Malawi (called deferred pay), implying that the massive increase in migration constituted a large positive income shock to the country. This fact will be important for interpreting our results for a few reasons. At the individual level, deferred pay, which forces a migrant to wait until returning home to collect most of their earnings, could reduce the impact of remittances, at least during the time the migrant is away. However, the forced savings aspect of deferred pay could also increase the total amount of money returning to the sending region, thereby increasing the impacts of remittances. At a practical level, we can be sure that at least two thirds of earnings returned to origin communities. Using our administrative data, we will be able to show large increases in money inflows towards districts with Wenela stations in both the early period (as migrants leave and return on rotating contracts), as well as when migrants return *en masse* in the later period. And, since miners themselves did not choose the value of their deferred pay, we do not need to be as concerned with the usual endogeneity problems associated with migrant motivations about whether to and how much to remit (Lucas and Stark 1985; Yang 2011).

To get a sense of the lower bound of how much an individual returning miner could bring home with him, we apply deferred payment shares to mine wage data (from Wilson 1972: 46 and Crush et al., 1991: 19) and conservatively assume that voluntary remittances are zero. The total amount of income that a family could have received from a miner returning from a two-year contract in 1969 would have been about K276 (Malawian Kwacha), which is two-thirds of two years' worth of mining wages. By 1973, this had increased to K437 and by 1974 this number was K691. Converting to USD (the exchange rate during the period was 0.83USD to K1), these amounts lie between 230USD and 832USD per trip. We compare deferred pay earnings with local agricultural wages below.

The post-1967 labor expansion came to an abrupt halt in April 1974, when a Wenela plane transporting miners back to Malawi crashed, killing 74 Malawian miners. In response, President Banda rescinded the labor agreement, banned Wenela recruiting, and recalled all Malawian migrant workers (Lucas 1985; Chirwa 1996). Between 1974 and 1977, mine employment fell dramatically (see Figure 2) from a high of over 120,000 men to zero by 1975. The initiation of the ban entailed a lump sum payout of deferred pay earnings for all returning miners. Wenela recruiters were able to restart operations in Malawi in 1977, and although there was a small increase in migration after this, employment levels for Malawians never returned to 1970 levels (see Figure 2). The South African mines had turned their strategy of recruitment inwards, substituting local labor for what they saw as unreliable foreign supplies (Crush 1986; Crush et al 1991, pg 129, Mariotti, 2015).

Alternatives to labor migration: Employment in Malawi

As early as the 1940s and 1950s, the attraction of migrating to work on the South African mines was strong. The domestic economy had always been small, with few perceived opportunities for growth. Despite several attempts to develop large estates producing cash crops for export (first by the colonial government, and then taken over by the post-independence Banda government, see Haviland 1953, Green 2007 and 2013 and McCracken 2012), wages in the local economy were always far below what workers could earn on the mines.

Malawians remaining at home had essentially two, not necessarily mutually exclusive, options for work in the agricultural sector. Workers could work for wages or as visiting tenants on the large tea, tobacco, cotton and sugar estates. Alternatively, they could work in the peasant smallholding sector (Kydd and Christiansen 1982) growing cash crops for export, or food crops to sell to estate laborers and visiting tenants who had no time or land to cultivate their own food.

Under Banda's leadership, the commercial agricultural estate sector grew substantially, underpinning Malawi's 6% annual growth rate from the late 1960s to the late 1970s. Employment on these estates also increased, and since estate wage-workers typically earned more than visiting tenants and smallholders (Chirwa 1997), this sector offered the most lucrative local alternative to mining. Yet, because estate wage earnings were always substantially lower than potential mine earnings (Chirwa 1997), it is not clear how the growth of opportunities for local employment and estate wages would have affected labor migration to South Africa (Weyl 1981, Chirwa 1997). Average annual earnings on local estates were K94.4 in 1968, about 70% of what a returning migrant received in 1969. These wages rose to K112 in 1973 and to K126.80 in 1974, but still represented only 37% of a migrant's deferred pay in that year (Pryor and Chipeta 1990) due to faster growth in mine wages. Despite the growth of the estate sector immediately after independence, many men continued to seek employment with Wenela because of low relative wages in agriculture, challenges associated with being a visiting tenant, and the decline of the smallholding sector.⁹ The absence of these mineworkers from the Malawian economy, their abrupt return in 1974, and the substantial migrant earnings associated with each of these flows is likely to have impacted investments in education of the next generation through several channels which we discuss next.

3. Education in Malawi in historical context

⁹ Although there is debate about whether Banda wanted to stem the tide of migration to support the growth of the estate sector, the timing of the labor recall initiated by the plane crash, was clearly unexpected. Certainly the South African Chamber of Mines Annual Report of 1973 did not note any concerns with regard to existing labor recruiting practices (Chamber of Mines Annual Report 1973; Paton 1995, pg 54).

Levels of human capital accumulation in Malawi are low, although they have been increasing over time. Prior to independence in 1964, missionaries were responsible for education and emphasized vocational training above literacy. Consequently, less than 6% of the population was literate in 1945. By 1964, less than 35% of school-aged children were enrolled in primary school. Between 1959 and 1978 however, total student enrollment in primary school increased by 30% (Heyneman 1980). These increases occurred despite a lack of investment in primary school construction. The number of primary schools did not increase between 1960 and 1992 (Ministry of Education Annual Report 1960: 37, Table 1; Malawi Statistical Yearbook 1995: 57). By the early 1990s, primary school enrollment had grown to 50% of the relevant age range (http://www.childinfo.org/files/ESAR_Malawi.pdf).

Before the mid-1990s, cost was an important constraint to attending school in Malawi. Average annual school fees were the same across the country: around K2.75 for the first four years of primary school and K5.75 for the next four years. We estimate that the tuition cost of sending three children to school was around 12 % of an agricultural worker's annual wage.¹⁰ On top of tuition, parents were responsible for expenditures like textbooks, exercise books, writing materials and school uniforms (Heyneman 1980). Following the introduction of universal free primary education in 1994, primary school enrollment increased by 50%, indicating that school fees were a substantial impediment to enrollment even as late as the 1990s (World Bank and UNICEF 2009).

Apart from the direct costs of education, outside options to work and obligations within the family raised the opportunity costs of sending children to school. Children could work in home production, on family farms doing subsistence farming (*mlimi*), or for other farmers or landlords, including owners of estates producing tea and tobacco for export. Chirwa (1993) provides a few reasons to expect that outside options for child workers might have differed across space in Malawi, particularly along the dimensions of access to agricultural estate work. On agricultural estates that paid wages, children could expect to earn higher wages than in household production or agricultural production on the family plot, where wages were often zero. On agricultural estates with visiting tenant systems, tenant families may have required their children's assistance without pay, to meet landlord-specified production quotas. The shadow value of child labor on estates would therefore have been high, especially when coupled with absentee adult males.

Both the direct and indirect costs of education played a role in the demand for education between 1950 and 1990. Table 1 uses Census data to provide some summary statistics on enrollment patterns among children age 10 to 14 just prior to the labor migration shock period. Only 34% of these children were enrolled in school that year. There is interesting variation across districts, though. First, enrollment is

¹⁰ Average fertility during this period was seven to eight children per woman.

significantly higher in Wenela (39%) relative to non-Wenela districts (24%). These pre-existing differences will be controlled for using district fixed effects in our empirical specification. Second, school enrollment is lower in districts with agricultural estates (Panel B, 27%) than in districts without estates (Panel C, 37%). These patterns are consistent with children having more alternatives to attending school in estate than in non-estate districts.

The means in Table 1 also reflect high prevalence of child labor for this young age group in the late 1960s. Around one in five children aged 10 to 14 were working for someone else or on family farms, with or without pay (Panel A, column 1). These shares probably underestimate the fraction of children working since the definition of work used here excludes home production.¹¹ The patterns of child labor across districts with and without estates are striking. In districts with agricultural estates (Panel B), child labor rates are high and approximately the same across Wenela and non-Wenela areas (18-19%). On the other hand, in districts without estates (Panel C), there is a significant 5 percentage point gap in child labor shares, with lower shares of children working in non-Wenela relative to Wenela districts (16% versus 21%). Before the labor migration shocks of the 1960s and 1970s, child labor was lowest in districts where the opportunity cost of going to school would have been lowest: in non-Wenela, non-estate districts. These differences in the patterns of enrollment and child labor will lead us to explore heterogeneity in the long run effects of labor migration shocks among districts with and without agricultural estates.

4. Data and descriptive statistics

We collect and assemble data from several sources to analyze the impacts of labor migration on long run human capital attainment. We digitize and use both district level and individual level census data on human capital attainment and migration as well as archival records on the location of Wenela stations. We supplement these with data from various years of the Malawian Statistical Yearbooks. This section outlines key features of our data and highlights some of the variation in migration used in our analysis. The Data Appendix contains further details on sample and variable construction.

Census data are from the complete 1977 and 1998 population Census⁷ of Malawi and cover 24 districts across all three regions (North, Central and Southern regions). The 1998 Census is comprised of individual level data while the 1977 data are only available at more aggregated district-five year cohort-gender level. We collapse the 1998 data to the same district-cohort-gender groups as in the 1977 data and match these to the 1977 data, restricting the sample to adults ages 20 to 44 (in five-year age groups) in

¹¹ About 11-12% of children report working in home production in the 1977 Census, the first year for which these data exist. There are no significant differences in the prevalence of this home production work across Wenela and non-Wenela districts, or across estate and non-estate districts.

1998 data. Because life expectancy in Malawi was only 46 years in the late 1990s (http://www.theglobaleconomy.com/Malawi/Life_expectancy/), we are concerned about mortality selection at ages over 40 affecting the composition of our sample. In Appendix 2, we use data from both Census years to show that less educated individuals in the older cohorts attrite at higher rates from Wenela districts relative to non-Wenela districts. Using only the 1998 data in our analysis would lead us to construct a biased estimate of the education gap between Wenela and non-Wenela districts in these older cohorts (which we show in that Appendix). Instead, we use the 1977 Census data to construct a synthetic older cohort of adults, using those aged 20 to 44 in 1977 to represent the 41 to 64 year old cohorts in 1998. Education questions are identical in both Census waves. In section 5, we describe how we define which cohorts were age-eligible for primary school during the labor expansion years and during the labor ban years.

To proxy for the pre-existing costs of getting access to mine work in South Africa, we collect administrative data on the number of Wenela recruiting stations established in each district prior to 1937. As shown in Figure 1, there are recruiting stations across the length of the country, so we can make within-region comparisons across cohorts facing higher versus lower migration costs.

It is useful to get some insight into factors that might have driven initial station placement. Table 2 presents correlations between the number of Wenela stations at the district level and district-level historical and geographic variables.¹² First, log population density measured in 1931 is negatively correlated with Wenela stations, although not significantly so in most cases. This could be because mine recruiters were unwilling to compete for male labor in areas where agricultural opportunities offered good outside options (i.e. where there were initially high densities of population on fertile land). Second, districts at higher altitude (i.e. more likely to be free of malaria), districts without estates, and districts in the Central region are more likely to have a Wenela station. Notably, the level of historical literacy does not predict the number of Wenela stations in a district. The implication is that the mining industry did not place additional value on local levels of human capital. Once region fixed effects are included, there is very little correlation between historical and geographic variables and the historical placement of the Wenela stations.

Our identification strategy rests on comparing outcomes across different cohorts from locations facing substantially different costs of signing up for mine work, proxied by the presence of a local Wenela station. Table 3 provides direct evidence that the number of Wenela stations in a district predicts substantially higher labor outmigration and higher inflows of money in response to exogenous shocks to

¹² Results are similar using “any Wenela station” in the district as an outcome.

labor demand between 1966 and 1977. The first four columns show results from regressions of the total number of circular migrants between 1967 and 1977 on the number of Wenela stations in the district, region fixed effects, historical literacy and population density, an indicator for the presence of an agricultural estate in the district, and the interaction of estate districts with Wenela districts. The outcome is the difference between the number of men who had ever migrated from a district and who returned from working abroad in the past 10 years, recorded in the 1977 Census, and the district-level stock of male migrants captured in the 1966 Census. Since the labor ban was still in place at the time of the 1977 Census, this variable gives us a complete picture of the total number of circular migrants at the district level between 1966 and 1977 (that is, it is a change in the stock of migrant workers).

Exposure to one more Wenela station in the district significantly raises the number of labor migrants by 1,532 men. Relative to the average change in migrants over the period (3,445) (Table 3, column 4), this is a 44% increase in migration for each additional station. This is a very large effect: relative to the initial mean stock of migrants in a district in 1966 (about 10,000), one additional recruiting station raises migration by 15 percentage points. Note that education levels measured by the local literacy rate in 1945 do not predict outmigration. If anything, higher historical literacy rates suggest less migration.¹³

The second set of columns confirms this relationship for money flows. The outcome in columns (5) to (8) is total deferred pay per person received by each district (using 1966 population totals) between 1966 and 1977. Deferred pay shares are set by contract, and constitute about 88% of total miner money flows back to Malawi (own calculations using administrative data on money flows). For each additional recruiting station, there is an additional K24 per person (enough to pay for about 8.6 years of primary school tuition) received over the ten-year period. Our measure of deferred pay per person smooths these receipts over the entire district; miner households would have experienced larger positive shocks to income because of the changing opportunities to work in South Africa.

Table 4 presents summary statistics for the full sample and for districts with and without a Wenela recruiting station along with the *p* value of the difference in means. In Panel A, we show means of treatment, treatment exposure, and outcome variables at the district-cohort-gender level. 63% of districts have at least one Wenela station and the average number of stations per district is 2.79. The fraction of cohorts eligible for primary school in either period, and in each of the younger and older cohorts, is balanced across Wenela and non-Wenela areas. Our sample of adults (aged 20 to 64) has a very low mean

¹³ Results are similar using a modified outcome, the migrant growth rate between 1966 and 1977, which takes into account initial migration levels in 1966. A district with one more recruiting station has 20 percentage points faster migration growth than the average district (Table 3, column 8), or an overall migrant growth rate of almost 50% (mean rate of migrant growth across all districts over these years 30%).

level of education: 2.56 years. Average education is 2.85 years in Wenela areas and only 2.07 years in non-Wenela areas; the share of adults who have ever been to primary school is 45% in Wenela areas and only 35% in non-Wenela areas. Both of these differences are strongly significantly different from zero and echo the differences in child enrollment patterns seen in Table 1.

In Panel B, we show district-level summary statistics for migration and money flows, along with historical and geographic variables from Census data from 1931, 1945, 1966 and 1977. An average district contains about 225,000 individuals. The table shows that while high shares of working age men report ever working aboard by 1977 (20%) in both Wenela and non-Wenela districts. However, although migration was widespread across the country, the incidence of the migration shocks and related money flows between 1967 and 1977 was concentrated in Wenela districts. The change in the number of men working as migrants is much higher in Wenela districts (5,060) relative to non-Wenela districts (752), and the migrant growth rate in Wenela areas is 46%, relative to a migrant growth rate of just 3% in non-Wenela areas between 1966 and 1977 (p value of this difference is 0.04). Wenela districts received on average K56 per person in deferred pay across the period, while non-Wenela districts received only 28% of that amount. The final part of the table shows no statistically significant differences in local labor market opportunities across Wenela and non-Wenela areas just before the labor treaty was signed in 1967. Just over 60% of men were engaged in some form of wage work in both areas, and almost 40% were not earning any wage. Our analysis will account for all pre-existing differences between Wenela and non-Wenela districts by using district fixed effects.

5. Empirical Strategy

Our aim is to estimate the long run impacts of migration on the intensive margin of the next generations' educational attainment (total years of schooling attained) and on the extensive margin (any primary school attained). These are relevant margins of adjustment for Malawi during this period, when there were only 50 government secondary schools across the country. Furthermore, focusing on primary education is important since functional literacy – usually gained after four years of education – is deemed necessary for positive returns to education in agriculture (Foster and Rosenzweig 1996; Appleton 2000).

Our identification strategy exploits variation in childhood exposure to mine labor migration shocks induced by external labor shocks interacting with district-specific migration costs.¹⁴ We use access to Wenela stations as shown in Figure 1 as a proxy for migration costs. Figure 3 illustrates the variation in

¹⁴ Our use of migration shocks measured at the level of the origin community is similar to the shock used in Theoharides (2014). In that paper, Theoharides estimates the short run impacts of labor migration demand shocks in the Philippines on enrollments in private secondary schools, also at district-level.

childhood exposure to the external labor shocks using the definition of eligibility for primary school enrollment across cohorts. We construct five-year birth cohorts for adults who are ages 20 to 64 in 1998 and have therefore completed whatever primary schooling they are likely to get.

We group cohorts into four broad categories, defining pre-treatment, early and late treatment, and post-treatment cohorts. The oldest cohorts born between 1933 and 1953 are too old for primary school by the start of the labor shock period, 1967. These cohorts range in age from 14 to 34 in 1967, and constitute our natural pre-treatment (control) group. The two middle groups in Figure 3 are eligible for primary school at some point during the labor migration shock. Cohorts born between 1954 and 1963 are eligible for primary school during the early labor expansion period: they are aged 4 to 13 in 1967 and are the *Early Treated* cohorts. Cohorts born between 1964 and 1973 are eligible for primary school during the labor ban period, as they are all under the age of 10 in 1974. They are the *Late Treated* cohorts. The youngest adult cohorts who are born between 1974 and 1978 are ineligible for primary school between 1967 and 1977: the eldest in this cohort is three years old in 1977. These youngest cohorts may experience lingering effects of the labor migration shock, which is why we refer to this group as the *Post Treatment* cohort, rather than a control cohort.

Although our definition of cohort eligibility using five-year age bins is coarse, we have good reasons for preferring this definition relative to one that specifies a treatment assignment for each individual year of age. In low literacy populations such as Malawi's, age misreporting makes it difficult to measure accurately a person's exposure to any age-specific treatments. Using the five-year age bins softens the impact of this misreporting on our estimates. Using these five-year bins also helps us get around the issue that age eligibility rules for enrollment in Malawi are only loosely enforced: in 1977, 20% of those aged 15 were enrolled in some lower primary school (grades 1-5). Finally, and more practically, the 1977 Census data are only available for five-year cohorts, restricting the way we can define treatment assignment.

An implication of using broad cohorts to define treatment is that treatment is somewhat fuzzy. This likely attenuates our estimates of the effects of migration on education. To be specific: not everyone who is defined to be eligible for primary school was eligible for the same length of time during the labor shock years. For example, the oldest individuals in the early treatment cohorts would have spent less total time exposed to the labor expansion shock than the youngest in the same cohort. The five-year age bins average effects across these different treatment intensities.

We estimate the education impacts of the labor migration shock using the following empirical model:

$$Y_{cgd} = \sum_{c=2}^9 \gamma_c Cohort_c * Wenela_d + \delta_c + \delta_g + \delta_d + (G_d * Trend)' \pi + \varepsilon_{cgd} \quad (1)$$

where Y_{cgd} is either the average years of schooling attained or the share of adults with any primary school in each cohort-gender-district (cgd) cell. δ_c , δ_g and δ_d are sets of five-year cohort, gender, and district fixed effects. G_d is a set of district-specific historical and geographic variables that might influence changes in education over time: the log of historical population density measured in 1931, the share of literate adults in 1945, and indicators for the region in which the district is located. In our full specification, we interact each of these baseline covariates with a linear trend ($Trend$) (results are almost identical when we interact these with cohort dummies instead of a linear trend, but the trend specification uses up fewer degrees of freedom). $Wenela_d$ is a count variable of the number of historical recruiting stations in the district (results are similar if we use an indicator for at least one Wenela station in the district). $Cohort_c$ is a set of dummy variables for each five-year cohort for the sample aged 20 to 64 in 1998; the omitted cohort is the oldest age group, 60-64. And ε_{cgd} is an idiosyncratic error term. Robust standard errors are clustered at the district level.

To trace out the differential effects of the labor migration shock across cohorts in districts with fewer versus more Wenela stations, we plot estimates of γ_c on the interaction terms $Cohort_c * Wenela_d$. Estimates of γ_{45-49} , γ_{50-54} , and γ_{55-59} measure the gap in education outcomes across Wenela and non-Wenela districts among the older control cohorts relative to the oldest cohort ineligible for primary schooling during the labor shock years. Ideally, we want these estimates to be zero, which would tell us there is no difference in education levels across Wenela and non-Wenela districts among cohorts eligible for primary schooling before the labor migration shocks begin.

The main parameters of interest are the remaining γ_c 's. Estimates of γ_{35-39} and γ_{40-44} measure the gap in educational attainment across Wenela and non-Wenela districts in cohorts exposed to the labor expansion (the *Early Treated* cohorts) compared with the same gap among the oldest ineligible cohorts. Estimates of γ_{25-29} and γ_{30-34} measure the gap in educational attainment across Wenela and non-Wenela districts in cohorts exposed to the labor contraction (the *Late Treated* cohorts), compared with the same education gap among the oldest ineligible cohorts. Finally, the estimate of γ_{20-24} captures any persistence in education gaps between Wenela and non-Wenela districts among *Post Treatment* cohorts, relative to the same gap amongst oldest ineligible cohorts. As noted in Section 4, because the definition of eligibility for primary school is necessarily coarse, we will not be able to capture more subtle variation in treatment exposure within a five-year cohort.

To condense results, we also present estimates from a related regression that groups five-year cohorts into the *Early Treated_c*, the *Late Treated_c*, and the *PostTreated_c* cohorts:

$$Y_{cga} = \beta_1 \text{EarlyTreated}_c * \text{Wenela}_d + \beta_2 \text{LateTreated}_c * \text{Wenela}_d + \beta_3 \text{PostTreated}_c * \text{Wenela}_d + \delta_c + \delta_g + \delta_d + (G_d * \text{Trend})' \kappa + \varphi_{cga} \quad (2)$$

In this specification, differences between β_1 and β_2 tell us about the relative education impacts of the labor expansion shock versus the labor ban, while β_3 tells us whether there is any persistence in the effects of the labor migration shocks among the youngest post-treatment cohorts.

Identification of the difference-in-differences parameters in specifications (1) and (2) relies on exploiting district-level variation in the costs of accessing mine jobs proxied by the presence of historical Wenela recruiting stations combined with within-district level variation in age eligibility of different cohorts. In most studies of migration, drawing causal conclusions from comparing migrant and non-migrant families or individuals is difficult because the decision to become a migrant is seldom random. Differences in educational outcomes of children of migrants cannot necessarily be attributed to the impact of migration, if migrants take more risks, are poorer or more able, or simply more interested in investing in human capital. There are several reasons why, in our setting, these sources of selection bias are less of a concern.

First, we measure outcomes at the district level and include a host of controls in equation (1) that limit concerns about potential confounding from underlying unobservable or observable differences between districts, from short, sharp, differential shocks to the local economy, and from differential trends across districts.¹⁵ Region-trend interaction terms also address concerns about mean reversion in education across districts. Most importantly, any concerns about differential education trends across Wenela and non-Wenela districts would need to account for the inverted U-shaped pattern of coefficients that we find. It is difficult to think of any reason for the endogenous placement of recruiting stations that would account for this changing pattern of coefficients.

Second, we use the older pre-treatment cohorts to control for counterfactual differences in educational attainment across Wenela and non-Wenela districts. Parallel education pre-trends in this older comparison group (that is, estimates of γ_{45-49} , γ_{50-54} , and γ_{55-59} that are zero) support our identification assumption that, conditional on controls, non-Wenela districts provide a valid counterfactual for Wenela districts. Third, migration to the mines was highly prevalent and there were migrants in every district and across all age

¹⁵ For example, the Northern Region was the birthplace of formal schooling in Malawi (Heyneman 1980), originally established by missionaries. Including region-specific trends helps us to account for any differences in education trends across regions that stem from different initial conditions.

groups. Differences across districts in migrant flows during the labor shock years had less to do with the characteristics of people who wanted to migrate, and more to do with how easy it was to get recruits from areas with pre-existing Wenela stations. Fourth, Wenela imposed few selection criteria on mine recruits, beyond requiring a physical fitness test, making it very unlikely that migrant selection on cognitive ability drives our results. Finally, our identification strategy relies on assumptions of no differential changes in school supply across Wenela districts: all of the effect comes from the changes in the demand-side. Since there was no massive school expansion program established in the post-independence period (see the discussion in Section 3), it is unlikely that changes in schooling access could confound the interpretation of our results.

One potential threat to validity of our results arises because proximity to Wenela stations in childhood (and hence $Wenela_{it}$) could be mismeasured. Our data do not contain information on birth district at the individual level, so we cannot be sure that a person's current district of residence is the district in which she went to school. This may pose problems if internal migrants have substantially more, or less, education than the average level of education of their origin and destination districts. To address this concern, we use information from the 1977 Census on the district, age and gender-specific prevalence of internal migration between 1967 and 1977 to create bounds for our education estimates in equation (1). In Appendix 2, we discuss how we implement this bounds analysis and show that our results are robust to accounting for the two extreme types of composition effects driven by internal migration.

Conditional on our identification assumptions, β_1 and β_2 capture the causal effect of exposure to the migration expansion and subsequent contraction on long run average educational attainment. We treat the expansion and contraction separately because they entail a different combination of shocks to the local district economy. In the earlier period between 1967 and 1974, men migrate to South Africa in increasing numbers, send back money in the form of remittances and collect cash via deferred payments when they circle back to Malawi after their two-year contracts. If children are required to substitute for missing male labor during the labor expansion, and if this effect is larger than the positive income effect (IE) from increased migrant earnings, we might see $\beta_1 < 0$. Alternatively, $\beta_1 > 0$ implies that the positive IE on educational investments dominates any crowd out of education during the labor expansion years.

In the later period between April 1974 and 1977, migrants repatriate to rural Malawi and bring home accumulated deferred pay and savings. With returning male migrants, children would no longer need to substitute for missing male labor. And, given the size of the earnings payouts to returning migrants, the labor ban also implied a large positive shock to current income, which should serve to increase contemporaneous investments in education, as long as schooling is a normal good. Since there is no

substitution effect in operation during the later period when men are back, we expect to see $\beta_2 > 0$. Since migrant labor flows never recovered to pre-1974 levels (see Figure 2), a finding of $\beta_3 > 0$ is consistent with at least some of the overall education response to the combined labor migration shocks being a response to a net positive income shock.

To compare with other estimates in the literature, we pursue an instrumental variables strategy (IV) to scale our reduced form estimates in two different ways: using the total number of circular migrants from each district between 1967 and 1977 and using the amount of mining money received by each district between 1967 and 1977. Each measure operationalizes the idea of the migration shock in a slightly different way. We estimate two versions of the following regression:

$$Y_{cgd} = \alpha_1 \text{EarlyTreated}_c * M_d + \alpha_2 \text{LateTreated}_c * M_d + \alpha_3 \text{PostTreated}_c * M_d + \delta_c + \delta_g + \delta_d + (G_d * \text{Trend})' \mu + \xi_{cgd} \quad (3)$$

where M_d is either the total number of circular migrants from the district between 1967 and 1977 (the same outcome used in the first four columns of Table 3), or the total deferred pay per capita flowing back to each district over these ten years (the same outcome used in the last four columns of Table 3).¹⁶ We use deferred pay rather than total remittances, because deferred pay shares were set by contract, and make up the majority (over 88%) of total money flows back to Malawi. We then instrument for each of $\text{EarlyTreated}_c * M_d$, $\text{LateTreated}_c * M_d$, and $\text{Post-Treated}_c * M_d$ using $\text{EarlyTreated}_c * \text{Wenela}_d$, $\text{LateTreated}_c * \text{Wenela}_d$ and $\text{Post-Treated}_c * \text{Wenela}_d$. All other controls are the same. While the difference-in-differences design already provides an estimate of the causal effects of exposure to the migration shocks, the instrumental variables strategy is a useful a re-scaling exercise. The estimates of α_1 , α_2 and α_3 tell us how much more education is gained by each of the treatment cohorts across high and low migration shock districts, or across high and low migrant money shock districts, controlling for differences in education outcomes across these districts using the older ineligible cohorts.

6. Empirical Results

Figures 4 and 5 show the main results for the long run effects of labor migration shocks on education. We plot the difference-in-differences coefficients γ_c from specification (1) on the y-axis for each cohort on the x-axis, for each outcome: total years of education and the share of adults with any primary schooling. The solid line connects our point estimates, the dashed lines represent the 95% confidence intervals, and the

¹⁶ We use the 1967 and 1977 Census to construct a measure of the number of migrants who left and returned to the district during this ten-year period. Without data on year-by-year flows of migrants at district level, we cannot separate out the effects of currently missing migrants from total ever migrants in this setting.

omitted category is the oldest cohort, age 60 to 64 in 1998. In each figure, the black line represents the gap in educational attainment (in Figure 4) or in primary school access (in Figure 5) across Wenela and non-Wenela districts, for each of the five-year age cohorts. The older pre-treatment cohorts include those aged 45 to 60 in 1998 (plus the omitted group); the labor expansion cohorts are those aged 35 to 44 in 1998; the labor contraction cohorts are aged 25 to 34 in 1998; and the youngest post-treatment cohorts are aged 20 to 24 in 1998.

Several features of these figures stand out. First, educational attainment and access to primary schooling both jump up, starting with the 40 to 44 year old cohort comprised of children eligible for schooling at the start of the labor expansion period. Positive gaps in educational attainment and access to primary schooling remain large and statistically different than zero for all younger cohorts down to those aged 25 to 29, the cohort eligible for schooling at the very end of the labor contraction period. There is also a positive difference in educational attainment and primary school access between Wenela and non-Wenela districts among the youngest, post-treatment cohorts, although the coefficients are smaller than during the labor ban years. The effects of the migration shock persist, although appear to wear off somewhat in the youngest cohorts.

Second, we see that the difference in total educational attainment between Wenela and non-Wenela districts (Figure 4) was essentially constant and zero among all of the pre-treatment cohorts. This parallel pre-trend among the oldest control groups is reassuring, as it gives us no indication that Wenela districts were improving before those children eligible during the labor expansion period were able to be in school. In Figure 5, we see that Wenela districts actually had declining entry into primary school among the oldest control cohorts, relative to non-Wenela districts. There is no reason to think that Wenela districts were improving educational outcomes in advance of the migration shocks.

Overall, both figures show that in the space of only ten years, exposure to mining employment shocks and concomitant migrant remittances enabled Wenela districts to overtake non-Wenela districts in their total amount of human capital, with long-lasting effect. Two decades after mine labor opportunities disappeared in Malawi, districts that had been most exposed to the labor shocks continued to have higher stocks of human capital among cohorts most exposed to these shocks. Additionally, districts most exposed to the labor shocks had persistently higher levels of human capital even among cohorts in primary school after the end of the migration shock, although these positive education impacts shrink over time. Adjustments occurred on both extensive and intensive margins: children going to school stayed in school for longer, and there were more kids going to school overall, as indicated by the increase in share

of adults with any primary schooling. These positive primary school enrollment effects account for between 7 and 10% of the total increase in enrollment rates between 1967 and 1978.¹⁷

Table 5 puts some numbers on these figures using the condensed specification in equation (2). The first two columns present results for years of education attained and the next two columns show results for the share of adults with any primary schooling. For each outcome, we show results from specifications including district fixed effects, female and age group dummies, and then including a set of trend interactions with baseline covariates (region fixed effects, the log of historical population density in 1931 and the share of the adult population literate in 1931). The estimates in columns 2 and 4 of Table 5 are weighted averages of the cohort-specific effects shown in Figures 4 and 5, with the weights being the shares of the sample in each cohort. Robust standard errors are clustered at the district level. Since there are only 24 districts, we report significance levels throughout using the small sample t distribution.¹⁸ The significance of our results is also robust to using randomization inference procedures to construct exact p values (see Appendix 3 for a description and the results of this procedure). The last three lines of Table 5 present p values for tests of equality between each of the difference-in-differences parameters.

Relative to the older pre-treatment cohorts, all of the post-1967 cohorts have higher educational attainment and are more likely to have ever been in primary school in Wenela compared with non-Wenela districts. These effects are large and statistically significant, and follow an inverted U-shaped pattern. For each additional Wenela station, cohorts eligible during the labor expansion gain 0.13 more years of education and are 1.4 percentage points more likely to have ever been to school (Table 5 columns 1 and 3). For each additional Wenela station, cohorts eligible during the labor contraction have 0.2 more years of education and are 2 percentage points more likely to have ever been to school. Among the youngest, post-treatment cohorts, for each additional Wenela station, education is 0.17 years higher and enrollment is 1.6 percentage points higher. Adding in trend interaction terms (Table 5 columns 2 and 4) reduces the size of the education impacts somewhat, but the effects remain positive, robust, and significant. For each additional Wenela station in the district, cohorts eligible for school during the labor expansion and labor ban periods gain an additional 0.12 to 0.18 years of education and are between 1.1 and 1.6 percentage points more likely to have ever been in primary school. These estimates represent a 4.8 to 6.9% gain in total years of education and a 2.6 to 3.9% gain in the share with any primary school.

¹⁷ Heyneman (1980, Table 3) provides national enrollment numbers for Malawi in 1967 and 1978. We use Census 1977 data (Table 1, Population counts) to construct the total number of children ages 5 to 19 inclusive in each of the 1966 and 1977 Census years. We estimate the primary school enrollment rate in 1967 was approximately 20%, rising to about 35% in 1978.

¹⁸ Relevant p values are taken from the t -distribution with degrees of freedom ranging from 23 (when there are no other controls in the regression) to 6 (when all controls are included and the sample is restricted to estate districts).

The two sides of the migration shock had different impacts on educational attainment and access to primary schooling, with the largest impacts felt by cohorts in Wenela districts who were eligible for schooling during the labor ban period: that is, the years in which miners were forced to return to Malawi and receive their lump sum payouts of deferred pay. In addition, the effects of the shock lingered (although with a smaller effect) as is evidenced by the higher educational attainment among the post-treatment cohorts.¹⁹ The last part of the table tests for the equality of outcomes across the three difference-in-differences parameters. Regardless of the specification, we can reject the null that the effects of the migration shock are the same among the labor ban cohorts and each of the labor expansion and post-treatment cohorts at the 1%, 5% or 10% level. The labor ban cohorts in Wenela districts gain significantly more education than cohorts exposed to the labor expansion and cohorts eligible for schooling after the end of the labor ban, controlling for educational gains of similar cohorts in non-Wenela districts. This pattern of coefficients is difficult to reconcile with any concern about placement of Wenela recruiting stations being correlated with positive (or negative) education trends.

Our positive estimates for each of the treated cohorts (early, late and post) are consistent with labor migration having large, positive, long run impacts on human capital profiles of sending communities. The immediate and persistent impacts of migration on school enrollment, attendance, and grade progression among exposed cohorts in sending communities manifest in higher overall human capital attainment among exposed cohorts in treated communities. Whether this accumulation of human capital ultimately translated into growth effects at district level is a natural next question which we leave for future research. At the most basic level, we observe strong cross sectional correlations between education and measures of household well-being. We show in Appendix 5 that Malawian households with more educated heads accumulate significantly more assets than those with lower levels of education, regardless of the measure of education used and controlling for a host of demographics, household characteristics, and district fixed effects. While usual concerns about selection and measurement error caution us against interpreting the point estimates as causal, the strong positive education gradient in household assets suggests that human capital formation and skills learned at school are important for household well-being in Malawi.

Discussion of magnitudes

To get a better sense of how many additional migrants left Malawi during the labor expansion years, how much money flowed back, and how these numbers translate into education impacts, we scale the

¹⁹ Various reasons in the literature exist for this persistence. For example, increased nutrition in childhood could have increased the returns to investing in the youngest children (Hoddinott et al 2011); or increases in savings facilitated by remittance income could continue to relax credit constraints and finance schooling beyond the end of the labor ban years (Ashraf, Aycinena, Martinez and Yang, forthcoming).

difference-in-differences results using estimates from Table 3. From this table, we know how the presence of a recruiting station induced migration flows out of districts and money flows into districts between 1966 and 1977. Without any other controls, each Wenela station in the district induces an additional 1,138 more labor migrants and an additional K13.5 per person. From Table 5, we know that for each additional Wenela station, eligible cohorts have between 0.13 and 0.2 years more schooling, excluding other controls (column 1). Combining these pieces of information, we can construct an estimate of the impact of labor migration on education. At the mean level of labor migration shock in a district (3,445), total educational attainment increases by between 0.4 years ($0.13/1,138*3,445$) and 0.6 years ($0.2/1,138*3,445$). Alternately, at the mean level of deferred pay per person in the district (K40), total educational attainment increases by between 0.39 years ($0.13/13.5*40$) and 0.59 years ($0.2/13.5*40$).²⁰

Table 6 does this re-scaling more formally, using the instrumental variables strategy in equation (3) that controls for other potential confounders in a regression framework. For each outcome, we present IV results including demographic and district level controls, and region-specific linear trends and trend terms interacted with baseline population density and literacy rates. First stage regression estimates appear in Appendix Table 1. Coefficients are presented for the mean district-level migration shock (3,445 additional men) and the mean district-level deferred pay shock per person (K40 per person).

Estimates in Table 6 indicate that at the mean level of migration in a district, human capital formation among eligible cohorts is higher by between 0.34 and 0.48 years (column 1), and these cohorts are between 3.2 and 4.3 percentage points more likely to have ever been to primary school (column 3). Compared with average levels of education and enrollment, this implies an increase in education of between 13 and 18% for the mean migration shock in a district, and an increase in any primary schooling of between 7 and 10%. Alternatively, at the mean level of deferred pay inflows in a district, human capital formation among eligible cohorts is higher by between 0.25 and 0.42 years (column 2), and these cohorts are between 1.6 and 3.8 percentage points more likely to have ever been to primary school (column 4). All results for the early and late treatment cohorts are significant at the 1% or 5% level.

There are reasons to expect that the return of remittances and receipt of deferred pay in lump sums several times larger than annual earnings would likely have had direct impacts on the local economy outside of migrant households. For example, families sharing resources within an extended family network could generate positive impacts for children of non-migrant households. Spending of this additional income could have generated local multiplier effects. Another potential source of spillovers could be driven by

²⁰ This is similar to, although not exactly, the Wald estimate. The unit of observation in Table 3 is the district, while in Table 5 it is the district-gender-cohort group.

equilibrium impacts on wages of left behind workers. Both the reduced form estimates of the impact of labor migration in Table 5 as well as the IV scaled results in Table 6 tell us how differential growth in migrant flows due to recruiting station placement affected the education profile of cohorts across the entire district. These estimates include any spillovers of labor migration to non-migrant households. Theoharides (2014) is the only other paper we are aware of that takes into account potential spillovers from migrant to non-migrant households by estimating the impacts of labor migration demand shocks measured at the local labor market level. Given her estimates, we calculate an elasticity of high school enrollment with respect to average migrant labor demand shocks in the Philippines of 0.17. Our estimates of the elasticity of primary school enrollment in Malawi are substantially lower: 0.07 to 0.1 (using the IV estimates in Table 6 column 3 and dividing by the average value of the education outcome: $0.032/0.41$ and $0.043/0.41$).²¹

We can also use our re-scaled IV estimates in Table 6 to compare our results with estimates of the elasticity of school enrollment to migrant income. In our context, the income elasticity of education is the percent change in enrollment divided by the percent change in migrant income at district level. Dividing the point estimates in Table 6 (column 4) by the mean of primary school enrollment variables turns these into elasticities. The elasticity of school enrollment with respect to deferred pay income is therefore between 0.07 and 0.09 ($0.029/0.41$ and $0.038/0.41$) for the early and late treatment cohorts respectively.²² This elasticity is far below the most credible estimate in the literature: Yang (2008) reports that the elasticity of school spending with respect to remittance income within migrant households in the Philippines is 0.55. Using these estimates, we compute the elasticity of school enrollment with respect to remittance income in that setting at 0.44. One reason for these differences could be the very different context of outside options for child work in Malawi. In the final section, we discuss our evidence on how these outside options may have muted the effects of labor migration on investments in human capital of the next generation.

Exploring Mechanisms

For both education outcomes, we reject the null that β_1 and β_2 in specification (2) are the same at the 1% or 5% levels (p values range from 0.01 to 0.03 in Table 5). Relative to both the *Early Treated* cohorts and the *Post Treated* cohorts, education impacts are always statistically larger for cohorts' eligible during the labor contraction period, when men have returned home and immediately after they collect their lump

²¹ The elasticity of 0.17 comes from an average migrant labor demand shock of 0.12 percentage points and an enrollment effect of 1.2 percentage points off a base of 56.8 percentage points: $(1.2/56.8)/0.12$.

²² Estimating a similar elasticity for the post-treatment cohorts is difficult, since it requires us to know the savings rate for remittance income.

sum payouts. One interpretation of these differences is that during the labor ban period, children no longer need to substitute for missing male labor, but the income effect still operates since men return home with deferred earnings and savings. Under this interpretation, we could use the difference between our estimates of β_1 and β_2 to back out the size of the substitution effect when labor migration is expanding. A second interpretation is that less money returned to each district in Malawi in each year of the expansion period than in the labor contraction period, generating a smaller impact on education outcomes. Using our disaggregated deferred pay data, we do find that the average amount returned to each district in each year, per capita, is around K3 in the early period and around K3.8 in the late period.

We offer two additional pieces of evidence to support the idea that there was some substitution between male and child labor in the early labor expansion period, which was not present in the later labor ban period. First, we examine how our difference-in-differences estimates from Table 5 vary across districts with greater versus fewer outside opportunities for child labor. Specifically, we look for treatment effect heterogeneity across districts with and without agricultural estates. Second, we ask directly whether there are contemporaneous impacts of the labor migration shock on the prevalence of child labor that differs across Wenela and non-Wenela districts, using data from 1968 and 1977 that bracket the massive shocks to migration. Because splitting the sample into estate and non-estate districts generates small subsamples for testing, these results are necessarily suggestive.

As Section 3 outlines, children in rural Malawi may have worked in the home, on the family farm, or on estates. Especially in the case of estates, where tenant farmers were required to satisfy annual quotas of output to protect their land rights, the value of additional child labor would have been higher than in areas without estates throughout the twentieth century. And, since the estate sector was booming in the late 1960s and early 1970s, the value of child labor in these districts would have been even higher during the time of the two migration shocks. Thus, we expect the effects of labor migration on human capital to depend on these differences in local agricultural production technology.

Table 7 shows just these relationships. We estimate versions of equation (2) using the district level dataset from the combined 1977 and 1998 Census. We split the sample into districts with and without large tea, tobacco, sugar or cotton estates. Although all areas experience the positive impacts of the migration shocks on educational attainment, effects are largest among eligible cohorts in high exposure (Wenela) districts without large agricultural estates. In these districts, cohorts eligible for primary school between 1967 and 1977 have between 0.15 and 0.22 more years of education and are between 1.4 and 2.1 percentage points more likely to be in school. In contrast, point estimates for the sample of estate districts are almost always smaller than for the non-estate sample.

Table 8 presents direct evidence on the contemporaneous effects of these labor migration shocks on the prevalence of child labor. We regress the share of children aged 10 to 14 years working in 1968 or in 1977 on a 1977 year dummy, the number of Wenela stations in the district, the interaction of the 1977 dummy with number of stations, district fixed effects, and a range of demographic characteristics. These regressions allow us to ask: did child labor among children aged 10 to 14 change differentially across districts with and without Wenela stations between 1968 and 1977? The results indicate that it did. For the full sample, districts with Wenela stations experienced a 1 percentage point decline in child labor among children aged 10 to 14 relative to districts without Wenela stations (Table 8 column 1). This decline only happens in the non-estate districts, where child labor declined by 1.3 percentage points. Among estate districts, the picture is very different. The share of children working rose by almost 20 percentage points in Wenela relative to non-Wenela districts. These effects make sense in the wake of the massive migration shocks of the 1960s and 1970s, and simultaneous growth in the estate sector.²³

Putting the results of Tables 7 and 8 together suggests that the local technology of production influenced the way in which circular labor migration affected families left behind. Exposure to mine employment shocks affected long run educational attainment differently in estate and non-estate districts. At the same time, exposure to these labor migration shocks differentially raised child labor in Wenela relative to non-Wenela districts, but only within estate areas. In non-estate areas, exposure to the labor migration shocks lowered rates of child labor over the 10-year period. We interpret this as evidence that where child labor was less prevalent and arguably less valuable, (i.e. in non-estate districts), circular labor migration and concomitant money flows substantially increased investments in schooling. The effects of this migration on education of treated cohorts were muted in estate districts, where child labor was a potentially valuable input into production, and even possibly a substitute for missing male labor.

7. Conclusions

Circular migration has long been considered one of the most immediate ways out of poverty for families, and potentially for whole sending communities. Despite its prevalence in most African labor markets, there is scant empirical evidence on this type of migration due to challenging identification concerns and data limitations. Our paper uses newly collected administrative data on exogenously determined access to migration opportunities and two waves of newly digitized Census data to provide direct evidence on one of the channels through which this migration can positively affect the lives of those left behind.

²³ We are unaware of any historical data that would allow us to test the direct impact of the labor expansion and contraction on child wages. However, such a test would likely be uninformative: the historical literature suggests that children were not earning market wages when employed in home production or on estates.

We show that the massive and unanticipated expansion and contraction of employment of Malawian men on South African mines over a ten-year period had lasting effects on human capital formation in rural communities of origin. This new evidence from Africa shows that the positive income effects of labor migration on the demand for education outweigh any potential negative substitution effects of this migration on the demand for child time in the labor market. However, there is some indication these long run impacts of labor migration on human capital accumulation of the next generation are muted, the higher the shadow value of child work is in a region. There is also some evidence that the effects persist beyond the end of the shock, although at a declining rate.

Our findings have broad relevance. The increase in educational attainment between 1967 and 1977 reflected in long run adult education outcomes twenty years later shows the extent to which a colonial era institution continues to play a role in Malawi's economy, long after independence and the end of formal recruiting. These lessons are relevant for many countries with historical, bilateral guest worker programs including Germany and Turkey (the *Gastarbeiterprogramm*), the US and Mexico (the *Bracero* program) and all Southern African countries that were part of the pre-and post-colonial labor reservoir cultivated by the South African mines. Given the scale of the circular migration we document in Malawi, it is surprising that there is so little empirical evidence on the long run effects on sending economies of these labor recruitment institutions.²⁴ The lessons we learn from this important yet little documented period in the economic history of Africa also have considerable relevance for existing and future planned guest worker programs in many low-income countries (Gibson, McKenzie and Rohorua 2013). Although some details of these modern programs may differ (e.g. deferred pay in worker contracts), modern guest worker programs also have limited-time work contracts, in-built circularity in migration flows and periodic labor bans that mirror the Malawian experience.²⁵ The evidence in this paper suggests that there is scope for modern guest worker programs to have positive, and importantly, long-lasting impacts on human capital formation in communities of origin.

²⁴ Lucas (1985) and Lucas (1987) are exceptions focusing on cross-country comparisons to get at aggregate effects.

²⁵ Malawi has developed a new guest worker program that has drawn significant controversy (Mponda 2013). Labor bans have recently occurred in Nepal <http://www.cnn.com/2012/08/09/world/meast/nepal-migrant-workers/index.html>.

References

- Antman, Francisca M., 2011 “The intergenerational effects of paternal migration on schooling and work: What can we learn from children’s time allocations?”, *Journal of Development Economics*, No. 96: 200-208.
- Antman, Francisca M., 2012 “Gender, Educational Attainment, and the Impact of Parental Migration on Children Left Behind”, *Journal of Population Economics*, Vol. 25 (4): 1187-1214.
- Antman, Francisca M., 2013 “The impact of migration on family left behind”, in *International Handbook of Migration*, edited by Amelie Constant and Klaus Zimmerman. Edward Elgar. Ch. 16: 293-308
- Appleton, Simon, 2000 “Education and Health at the household level in sub-Saharan Africa” *Explaining African Economic Growth Performance Conference Series*, CID Working Paper 33
- Ashraf, Nava, Diego Aycinena, Claudia Martínez A., and Dean Yang, forthcoming “Savings in Transnational Households: A Field Experiment Among Migrants from El Salvador” *Review of Economics and Statistics*
- Batista, Catia, Aitor Lacueste and Pedro C. Vicente, 2012 “Testing the ‘brain gain’ hypothesis: Micro evidence from Cape Verde”, *Journal of Development Economics*, Vol. 97: 32-45
- Beegle, Kathleen, Joachim De Weerd, Stefan Dercon, 2011 “Migration and economic mobility in Tanzania: Evidence from a tracking survey” *Review of Economics and Statistics*, Vol. 9 (3): 1010-1033
- Chamber of Mines of South Africa, 1973 *Annual Report*, Johannesburg.
- Chirwa, Ephraim W. and Mirriam M. Matita, 2009 “The rate of return on education in Malawi”, *Department of Economics Working Paper Number 2009/01*, University of Malawi, Chancellor College
- Chirwa, Wiseman Chijere, 1992 “‘TEBA is power’: Rural labour, migrancy and fishing in Malawi, 1890s – 1985” Ph.D thesis, Queen’s University, Kingston, Ontario
- Chirwa, Wiseman Chijere, 1993 “Child and youth labour on the Nyasaland Plantations 1890-1953”, *Journal of Southern African Studies*, Vol. 19 (4): 662-680
- Chirwa, Wiseman Chijere, 1996 “The Malawi Government and South African Labour Recruiters, 1974-1992”, *Journal of Modern African Studies* Vol. 34 (4): 623-642
- Chirwa, Wiseman Chijere, 1997 “‘No TEBA...Forget TEBA’: The Plight of Malawian Ex-migrant Workers to South Africa, 1988-1994”, *International Migration Review* Vol. 31 (3): 628 -654
- Clarke, Duncan. G., 1977 *Agricultural and Plantation Workers in Rhodesia: A Report on conditions of labour and subsistence*. Mambo Press.
- Clemens, Michael A., 2011 “Economics and Emigration: Trillion-Dollar Bills on the Sidewalk?” *Journal of Economic Perspectives*, Vol. 25 (3): 83-106
- Clemens, Michael and David McKenzie, 2014 “Why don’t remittances appear to affect growth?” *World Bank Policy Research Working Paper No. 6856*
- Coleman, G., 1972 “International Labour Migration from Malawi, 1875-1966”, *Journal of Social Science*, Vol. 2: 31-46
- Constant, Amelie F. and Klaus F. Zimmerman, 2013 “The economics of circular migration”, in *International Handbook of Migration*, edited by Constant, Amelie and Klaus Zimmerman. Edward Elgar Publishers. Ch. 3: 55-74
- Cox Edwards, Alejandra and Manuelita Ureta, 2003 “International Migration, Remittances and Schooling: Evidence from El Salvador,” *Journal of Development Economics* in Vol. 72 (2): 429–461
- Crush, Jonathan. S., 1986 “The Extrusion of Foreign Labour from the South African Gold Mining Industry” *Geoforum* Vol. 17 (2): 161-172
- Crush, Jonathan. S., Alan Jeeves and David Yudelman, 1991 *South Africa’s Labor Empire*. Westview Press
- De Brauw, Alan and John Giles, 2006 “Migrant Opportunity and the Educational Attainment of Youth in Rural China”, *IZA Discussion Papers* No. 2326

- De Brauw, Alan, Valerie Mueller, and Hak Lim Lee, 2013 “The role of rural-urban migration in the structural transformation of sub-Saharan Africa” *World Development*, in press
- Foster, Andrew D. and Mark R. Rosenzweig, 1996 “Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution”, *American Economic Review*, Vol. 86 (4): 931-53
- Gibson, John, David McKenzie and Steven Stillman, 2011 “The impacts of international migration on remaining household members: Omnibus results from a migration lottery program”, *Review of Economics and Statistics*, Vol. 93 (4): 1297–1318
- Gibson, John, David McKenzie and Steven Stillman, 2013 “Accounting for selectivity and duration-dependent heterogeneity when estimating the impact of emigration on incomes and poverty in sending areas”, *Economic Development and Cultural Change*, Vol. 61 (2): 247-280
- Gibson, John, David McKenzie and Halahingano Rohorua, 2013 “Development impacts of seasonal and temporary migration: A review of evidence from the Pacific and Southeast Asia”, *CREAM Discussion Paper Series No. 08/13*
- Green, Erik, 2007 “Modern Agricultural History in Malawi: Perspectives on Policy Choices”, *African Studies Review* Vol. 50 (3): 143–67
- Green, Erik, 2013 “Land Concentration, Institutional Control and African Agency—Growth and Stagnation of European Tobacco Farming in Shire Highlands, c 1900 –1940”, in E. Hillbom and P.Svensson (eds.) *Agricultural Transformations and Global History*, pp.229–262. London: Routledge
- Government of Nyasaland, 1961 “Aide Memoire on WENELA activities, Provincial Reports, PCN1-21-16”, Zomba National Archives
- Hanson, Gordon and Christopher Woodruff, 2003 “Emigration and educational attainment in Mexico”, Mimeo
- Haviland, William E., 1953 “The Rise of the African Tobacco Industry in Nyasaland and Its Production Problems” *The South African Journal of Economics*, Vol. 22 (2): 141–52
- Heyneman, Stephen P., 1980 “The evaluation of human capital in Malawi”, *World Bank Staff Working Papers* No. 420. The World Bank, Washington D.C.
- Hoddinott, John, John Maluccio, Jere R. Behrman, Reynaldo Martorell, Paul Melgar, Anges R. Quisumbing, Manuel Ramirez-Zea, Aryeh D. Stein, Kathryn M. Yount, 2011 “The consequences of early childhood growth failure over the life course”, *IFPRI Discussion Paper* No. 01073
- Jeeves, Alan, 1987 “William Gemill and South African Expansion 1920-1950”, Paper prepared for University of the Witwatersrand History Workshop: The making of Class, 9-14 February, 1987
- Kosack, Edward. 2015 “The Bracero Program and Effects on Human Capital Investments in Mexico, 1942-1964”, Working Paper, Department of Economics University of Colorado Boulder
- Kydd, Jonathan and Robert Christiansen, 1982 “Structural change in Malawi since independence: Consequences of a development strategy based on large-scale agriculture” *World Development* Vol. 10 (5): 355-375
- Lucas, Robert E. B., 1985 “Mines and migrants in South Africa”, *American Economic Review*, Vol. 75 (5): 1094-1108
- Lucas, Robert E. B. and Stark, Oded, 1985 “Motivations to Remit: Evidence from Botswana,” *Journal of Political Economy*, vol. 93(5), pages 901-18, October.
- Lucas, Robert E. B., 1987 “Emigration to South Africa’s Mines” *American Economic Review* Vol. 77 (3): 313-330
- Lucas, Robert E. B., 2005 *International Migration and Economic Development: Lessons from Low-Income Countries*. EGCI: Edward Elgar
- Malawi Ministry of Education, 1977 “Major trends in education in Malawi 1974-1976”, UNESCO International Conference on Education
- Malawi National Statistical Office, 1972 *Malawi Statistical Yearbook*. Government Printer, Zomba
- Malawi National Statistical Office, 1995 *Malawi Statistical Yearbook*. Government Printer, Zomba
- Malawi National Statistical Office, 1979 *Malawi 1977 Population Census Final Report Volumes I and II*, Zomba

- Malawi National Statistical Office, 1967 Malawi 1966 Population Census Final Report, Government Printer, Zomba
- Mariotti, Martine 2015 “Father’s employment and sons’ stature: The long run effects of a positive regional employment shock in South Africa’s mining industry” *Economic Development and Cultural Change*, DOI: 10.1086/679755.
- McCracken, John, 2012 A History of Malawi, 1959-1966, Boydell and Brewer
- McKenzie, David and Hillel Rapoport, 2011 “Can migration reduce educational attainment? Evidence from Mexico,” *Journal of Population Economics*, Vol. 24 (4): 1331-1358
- Mountford, Andrew, 1997 “Can a brain drain be good for growth in the source economy?”, *Journal of Development Economics*, Vol. 53: 287-303
- Mponda, Felix, 2013 “Malawi offers South Korea 100,000 workers”, Business Day, June 11.
<http://www.bdlive.co.za/africa/africanews/2013/06/11/malawi-offers-south-korea-100000-workers>
- Nyasaland Governor, 1956 “Memorandum on Labor Migrancy in Malawi” 1956 Nyasaland Provincial Office Memo, December 7 1961
- Nyasaland Ministry of Education, 1960 Annual Report. Government Printer, Zomba
- Nyasaland Protectorate, 1931 Report on the Census of 1931. Government Printer, Zomba
- Nyasaland Protectorate, 1946 Report on the Census of 1945. Government Printer, Zomba
- Paton, Bill, 1995 Labour Export Policy in the Development of Southern Africa. Macmillan.
- Pritchett, Lant, 2006 Let their people come: Breaking the gridlock on international labor mobility. Center for Global Development: Washington D.C.
- Prothero, R. Mansell, 1974 “Foreign migrant labour for South Africa”, *International Migration Review*, Vol. 8 (3): 383-394
- Pryor, Frederic. L. and Chinyamata Chipeta, 1990 “Economic Development through Estate Agriculture: The Case of Malawi”, *Canadian Journal of African Studies*, Vol. 24 (1): 50-74
- Theoharides, Caroline, 2014 “Manila to Malaysia, Quezon to Qatar: International Migration and its Effects on Origin-Country Human Capital”, Working Paper December 29, 2014
- Treaty Series No.10/1967, 1967 “Agreement between the Governments of the Republic of South Africa and Malawi relating to the Employment and Documentation of Malawi Nationals in South Africa”
 Place and date of signing: Pretoria and Blantyre, 1 August 1967, Date of entry in force: 1 August 1967
- Weyl, Ulrich, 1981 “Labour migration in Malawi as a barrier to development”, PhD thesis, Universität Osnabrück
- Woodruff, Christopher, 2007 “How does migration affect local development? What Mexico’s experience tells us”, in Susan Pozo (ed.) *Immigrants and their International Money Flows*, Upjohn Institute.
- World Bank and UNICEF, 2009 Abolishing School Fees in Africa: Lessons from Ethiopia, Ghana, Kenya Malawi and Mozambique IBRD, Washington D.C.
- Wilson, Francis, 1972 Labour in the South African Gold Mines 1911-1969, Cambridge University Press
- Yang, Dean, 2008 “International Migration, Remittances and Household Investment: Evidence from Philippine Migrants’ Exchange Rate Shocks,” *Economic Journal*, Vol. 118 (528): 591-630
- Yang, Dean, 2011 “Migrant Remittances,” *Journal of Economic Perspectives*, American Economic Association, Vol. 25 (3):129-52

Figure 1 The spatial distribution of Wenela mine recruiting stations in 1937

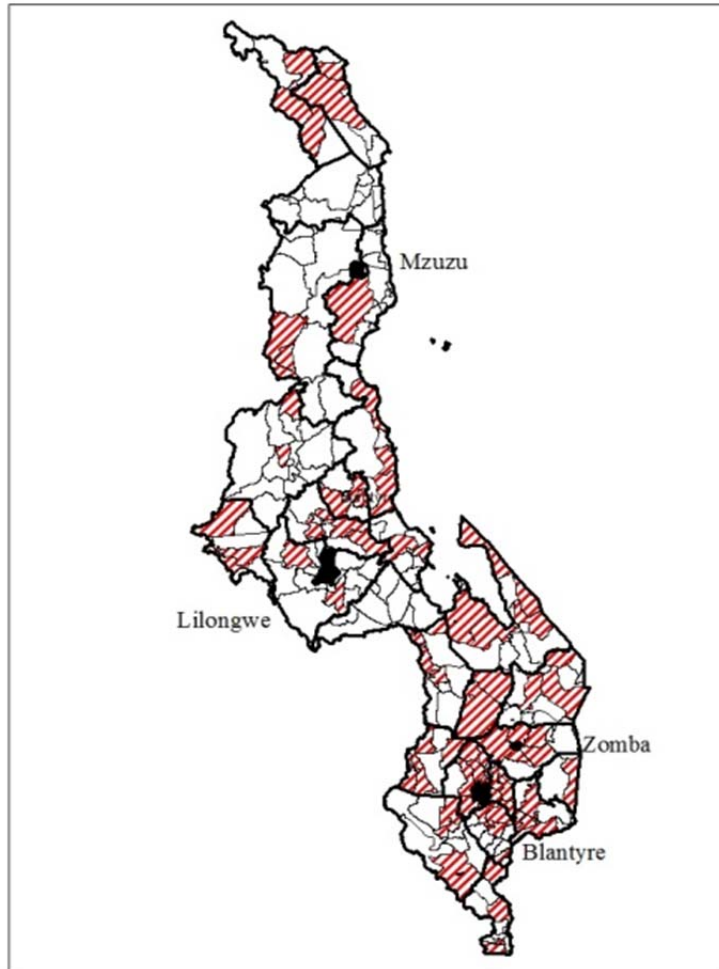


Figure 1 shows district boundaries (thick black lines), sub-district/traditional authority boundaries (thinner black lines) and the distribution of Wenela recruiting stations established by 1937 (hatched areas) across Malawi. Malawi's four cities are the shaded areas.

Figure 2: Annual employment of Malawian miners on South African mines, 1950-1994

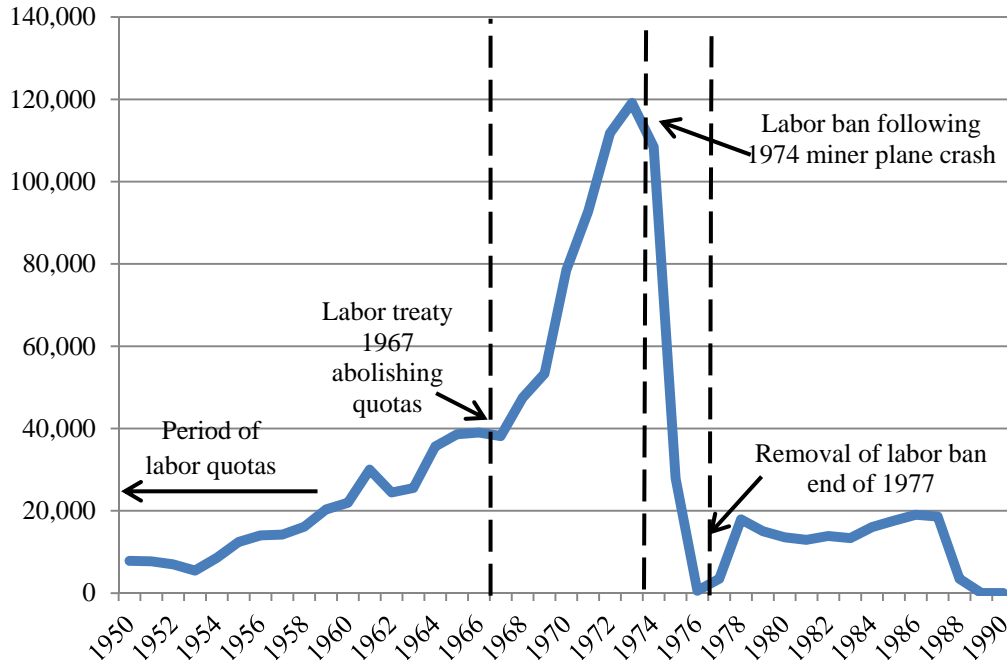


Figure 2 shows number of workers contracted by Wenela to work on South African mines in each year. The three dotted lines represent (from left to right) the abolition of labor quotas in August 1967, the moratorium on migration after the April 1974 Malawian plane crash and the legal resumption of mine migration in 1978.

Figure 3: Treatment assignment: Primary school eligibility in labor shock years 1967 -1977

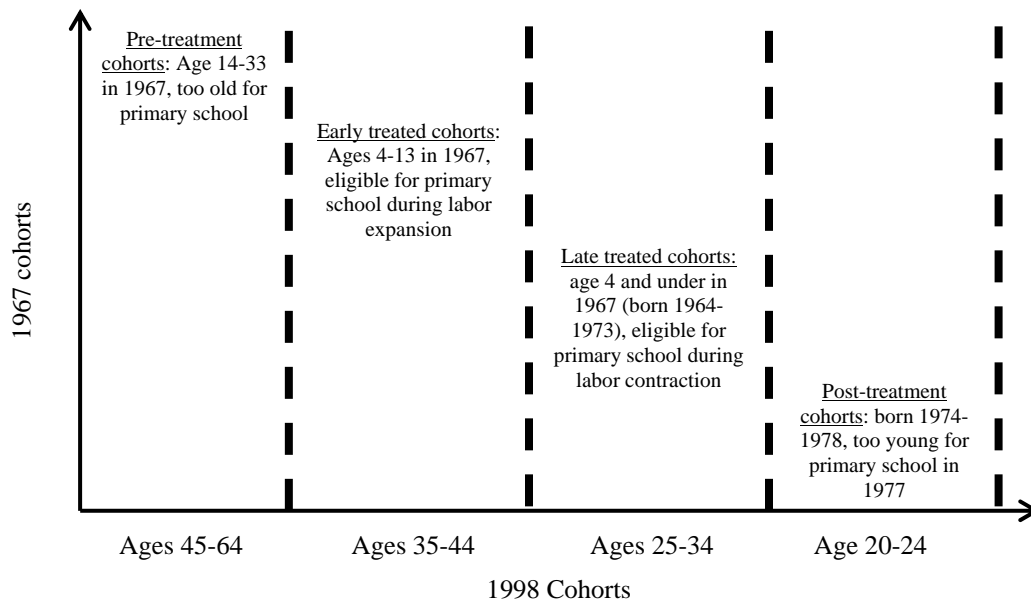


Figure 3 shows which cohorts are eligible for primary school during the labor shock years, 1967 to 1977. The horizontal axis shows age of cohorts in 1998, the vertical axis shows the range of ages for each cohort in 1967.

Figure 4: Estimated differences in education by cohort and Wenela status of district

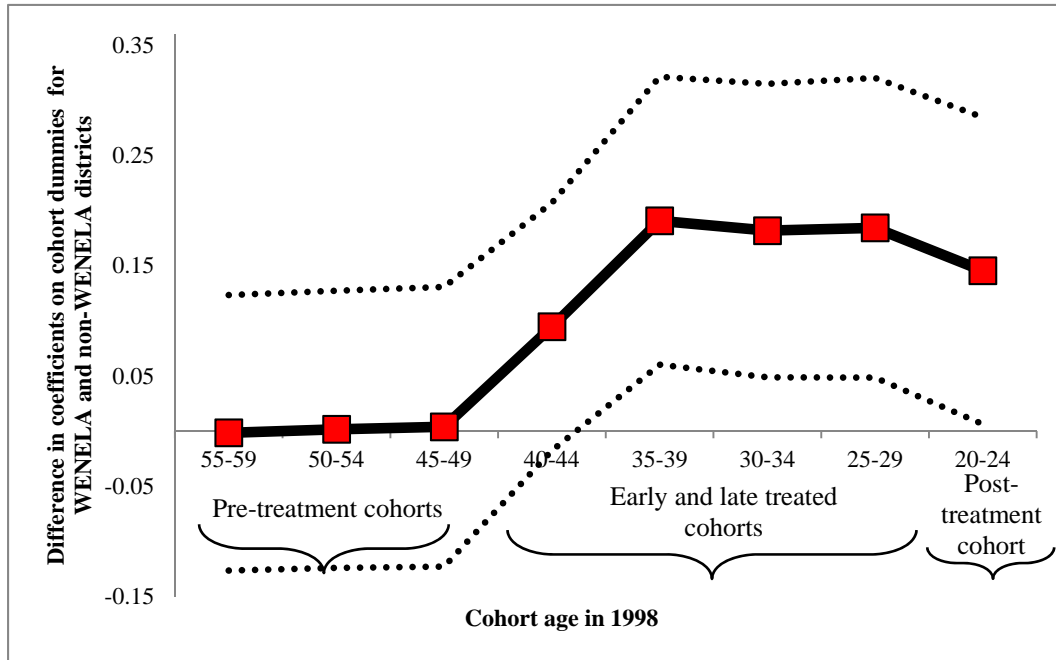
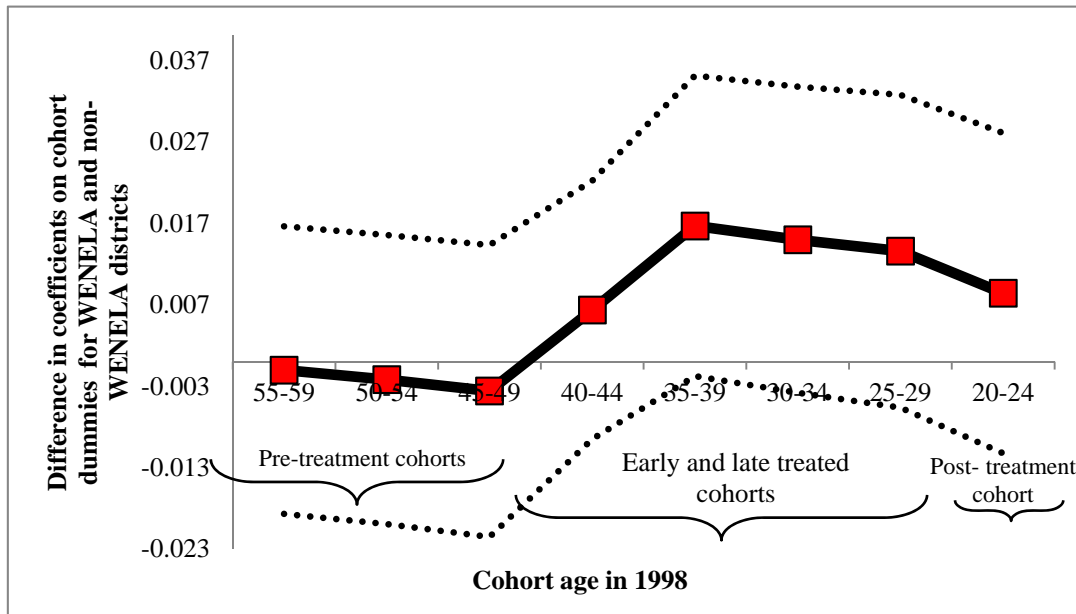


Figure 5: Estimated differences in primary schooling by cohort and Wenela status of district



Figures 4 and 5: show estimated interaction term coefficients (solid line) and 95% confidence intervals (dotted lines) from a regression of education (figure 4) and primary schooling (figure 5) on nine age group dummies and their interaction with the number of WENELA stations in the district. Controls include: female, log population density in 1931, share literate in 1945, district and region fixed effects, and trend terms interacted with region fixed effects, baseline population density and baseline literacy rates. The x-axis shows five-year age cohorts in 1998.

Table 1: Share of children aged 10-14 working and share in school, 1966-1968

	Full sample	Wenela Recruiting Districts	Non-Wenela Districts	<i>p</i> value of difference
<i>Panel A: All districts</i>				
Share in school	0.34	0.39	0.24	***
Share working	0.19	0.21	0.17	**
N	46	30	16	
<i>Panel B: Districts with agricultural estates</i>				
Share in school	0.27	0.33	0.22	*
Share working	0.18	0.19	0.18	
N	16	8	8	
<i>Panel C: Districts without agricultural estates</i>				
Share in school	0.37	0.41	0.25	***
Share working	0.20	0.21	0.16	**
N	30	22	8	

Statistical significance at the 1, 5, and 10 percent levels is indicated by ***, **, and *, respectively, and evaluated relative to the small sample *t* distribution to account for the small number of clusters. Enrolment data are from the 1966 Census. Work data are from the 1968 National Sample Survey of Agriculture. Excluded categories are home production and no activities. An observation is a district-cohort-gender group for the 10-14 age cohort.

Table 2: Historical and geographic predictors of Wenela recruiting station placement at district-level

	<i>Number of recruiting stations in the district</i>			
	(1)	(2)	(3)	(4)
Log population density, Census 1931	-0.979 (0.595)	-1.206* (0.585)	-0.758 (0.639)	-0.176 (1.035)
Estate district		-1.489 (1.126)	-0.891 (1.230)	-0.603 (1.535)
Altitude (meters)^*100			0.288** (0.117)	0.125 (0.204)
Literacy rate, Census 1945			2.678 (17.140)	10.010 (17.080)
Central region				1.019 (2.128)
Southern region				-1.411 (3.961)
Observations	24	24	24	24
R-squared	0.07	0.13	0.24	0.34
Mean of outcome	2.79	2.79	2.79	2.79

Robust standard errors in parentheses in all regressions. Statistical significance at the 1, 5, and 10 percent levels is indicated by ***, **, and *, respectively, and evaluated relative to the small sample t distribution to account for the small number of clusters. Outcome is the number of recruiting stations in the district in 1937. All variables are measured at district level.

Altitude is average altitude for each district and is a proxy for malaria risk.

Table 3: Wenela recruiting stations predict migration and money flows, 1966-1977

	<i>Δ in number of migrants from the district between 1966 and 1977</i>				<i>Deferred miner pay (Kwacha) per person, received between 1966 and 1977</i>			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of Wenela stations	1,138** (457)	1,356*** (362)	1,357*** (354)	1,532*** (427)	13.5 (10.24)	19.78* (10.71)	19.23* (10.61)	23.75** (9.07)
Central region		4,364 (3,359)	3,895 (3,816)	3,117 (4,390)		17.34 (47.56)	22.76 (54.92)	-0.848 (56.16)
Southern region		1,431 (4,901)	479.8 (5,629)	-821.3 (6,274)		104.1 (66.73)	111.9 (78.81)	72.79 (75.18)
Mean Literacy rate in 1945		-7,114 (5,886)	-6,923 (6,425)	-6,577 (6,524)		631.1 (465)	587 (498)	732.1 (543)
Log population density in 1931			827 (1,977)	1,293 (2,116)			-5.733 (18.25)	8.52 (15.43)
Estates			374.1 (3697)	2,448 (4,762)			-28.98 (34.84)	21.85 (25.17)
Estate district*Number of Wenela stations				-702.3 (1,086)				-17.41 (10.53)
N	24	24	24	24	22	22	22	22
R2	0.15	0.41	0.41	0.42	0.25	0.49	0.51	0.58
Mean of outcome variable	3,445	3,445	3,445	3,445	40	40	40	40

Robust standard errors in parentheses in all regressions. Statistical significance at the 1, 5, and 10 percent levels is indicated by ***, **, and *, respectively, and evaluated relative to the small sample *t* distribution to account for the small number of districts. Unit of observation is the district. Outcome are the change in the raw number (stock) of male migrants between 1966 and 1977, measured at district-level using Census data in 1966 and 1977 (columns 1 to 4) and total deferred pay per person (population measured in 1966) received by each district between 1966 and 1977 (columns 5 to 8). Number of Wenela stations is a count variable of all stations in the district in 1937 (mean is 2.79), estate is a dummy for whether the district contains a tea, tobacco, sugar or cotton plantation. Deferred pay data exist for a subset of districts.

Table 4: Summary statistics for Census, Geographic and Administrative Data

	Full sample		Wenela Recruiting Districts		Non-Wenela Districts		<i>p</i> value of difference
	Mean	s.d.	Mean	s.d.	Mean	s.d.	
<i>Panel A: Outcomes and control variables for District-5 Year Age Group-Gender cells</i>							
<i>Variables measuring exposure to mining employment shocks</i>							
Number of Wenela stations	2.79	2.77	4.47	2.19	0.00	0.00	0.00
Any Wenela station	0.63	0.48	1.00	0.00	0.00	0.00	0.00
Early Treatment Cohorts: Primary school eligible in 1967-1973	0.30	0.46	0.30	0.46	0.30	0.46	0.50
Late Treatment Cohorts: Primary school eligible in 1974-1977	0.20	0.40	0.20	0.40	0.20	0.40	0.50
Post-Treatment cohorts: Primary school eligible after 1977	0.10	0.30	0.10	0.30	0.10	0.30	0.50
Pre-Treatment cohorts: Primary school eligible before 1967	0.40	0.49	0.40	0.49	0.40	0.49	0.50
<i>Education outcomes</i>							
Total years of education for adult sample	2.56	2.20	2.85	2.36	2.07	1.79	0.00
Share with any primary school for adult sample	0.41	0.29	0.45	0.30	0.35	0.25	0.00
Number of observations	480		300		180		
<i>Panel B: Geographic and historical district variables</i>							
Number of Wenela stations	2.79	2.83	4.47	2.26	0.00	0.00	0.00
Any Wenela station	0.63	0.49	1.00	0.00	0.00	0.00	0.00
Fraction of men ever been abroad by 1977~	0.19	0.06	0.20	0.05	0.18	0.06	0.14
Δ number of migrants, 1966-1977	3,445	8,282	5,060	9,672	752	4,519	0.08
Migrant growth rate, 1966-1977	0.30	0.68	0.46	0.78	0.03	0.37	0.04
Total deferred miner pay (Kwacha) per person, 1966-1977	40.12	77.03	56.74	96.43	16.10	23.03	0.08
Altitude: high malaria area=1	0.28	0.35	0.20	0.32	0.43	0.36	0.06
Population density in 1931	24.67	22.13	15.71	9.59	39.60	29.09	0.01
District contains an estate	0.21	0.41	0.27	0.46	0.11	0.33	0.17
English and vernacular literacy, share of youth in 1945	0.08	0.04	0.09	0.04	0.06	0.02	0.00
Fraction of men doing wage work (farm, cash or other) 1966	0.63	0.10	0.62	0.12	0.66	0.08	0.14
Fraction of men not earning any wage 1966	0.37	0.10	0.38	0.12	0.34	0.08	0.14
Number of districts	24		15		9		

Panel A contains summary statistics using 1998 micro data and 1977 aggregate data, reported at the district-5 year age group-sex level. Panel B contains summary statistics from geographic data, aggregate Census data in 1931, 1945, 1966, 1977 and administrative data on migrant and deferred pay flows. *p* values are reported for the test of the difference in means across recruiting and non-recruiting station areas using robust standard errors and evaluated using the small sample *t* distribution to account for the small number of clusters. Estate is a dummy variable as described in the text.

Table 5: Long run effects of labor migration shocks on education: Difference-in-differences results

	<i>Total years of schooling attained</i>		<i>Share with any primary school</i>	
	(1)	(2)	(3)	(4)
Early Treatment Cohorts*Num. Wenela stations (β_1)	0.137*** (0.045)	0.124** (0.046)	0.014** (0.005)	0.011** (0.004)
Late Treatment Cohorts*Num. Wenela stations (β_2)	0.202*** (0.066)	0.179** (0.063)	0.020** (0.007)	0.016*** (0.005)
Post-Treatment Cohorts*Num. Wenela stations (β_3)	0.170** (0.061)	0.141** (0.066)	0.016*** (0.005)	0.010 (0.006)
Trend interactions	N	Y	N	Y
N	480	480	480	480
R2	0.83	0.86	0.82	0.84
Mean of outcome variable	2.56	2.56	0.41	0.41
p value of F test $H_0: \beta_1=\beta_2$	0.01	0.01	0.01	0.03
p value of F test $H_0: \beta_1=\beta_3$	0.15	0.57	0.33	0.73
p value of F test $H_0: \beta_2=\beta_3$	0.02	0.07	0.01	0.06

Robust standard errors clustered at the district level. Statistical significance at the 1, 5, and 10 percent levels is indicated by ***, **, and *, respectively, and evaluated relative to the small sample t distribution to account for the small number of clusters. Unit of observation is the district-5 year age group-gender cell. The outcomes are average years of schooling attained by the district-cohort-gender group and the share of each district-gender-cohort group with any primary schooling. Vector of controls included in every regression: female dummy, cohort dummies, region fixed effects, the log of district-level population density in 1931 and the share of literate youths in 1945, and district fixed effects. In the second specification for each outcome, a trend term is interacted with each of the region fixed effects, baseline population density and baseline literacy rates variables. Number of Wenela stations in the district is a count variable. Sample includes adults ages 20 to 44 in 1977 and 1998 census.

Table 6: Scaling human capital effects of migration shocks using migrant outflows and money inflows, IV results

	<i>Total years of schooling attained</i>		<i>Share with any primary school</i>	
	(1)	(2)	(3)	(4)
Total Migrants (1967-1977)*Early Treatment Cohorts (λ_1)	0.341** (0.125)		0.032*** (0.011)	
Total Migrants (1967-1977)*Late Treatment Cohorts (λ_2)	0.482*** (0.162)		0.043*** (0.014)	
Total Migrants (1967-1977)*Post-Treatment Cohorts (λ_3)	0.347* (0.168)		0.024 (0.016)	
Deferred Pay per capita (1967-1977)*Early Treatment Cohorts (λ_1)		0.316*** (0.103)		0.029** (0.013)
Deferred Pay per capita (1967-1977)*Late Treatment Cohorts (λ_2)		0.428*** (0.134)		0.038*** (0.015)
Deferred Pay per capita (1967-1977)*Post-Treatment Cohorts (λ_3)		0.252*** (0.080)		0.016 (0.011)
Additional controls, District FE	Y	Y	Y	Y
Trend interactions	Y	Y	Y	Y
N	480	440	480	440
Mean of outcome variable	2.56	2.56	0.41	0.41
p value of Chi2 test $H_0: \lambda_1=\lambda_2$	0.00	0.00	0.02	0.02
p value of Chi2 test $H_0: \lambda_1=\lambda_3$	0.96	0.66	0.63	0.52
p value of Chi2 test $H_0: \lambda_2=\lambda_3$	0.20	0.27	0.17	0.29

Robust standard errors clustered at the district level. Statistical significance at the 1, 5, and 10 percent levels is indicated by ***, **, and *, respectively, and evaluated relative to the small sample t distribution to account for the small number of clusters. Unit of observation is the district-5 year age group-gender cell. The outcomes are average years of schooling attained by the district-cohort-gender group and the share of each district-gender-cohort group with any primary schooling. Vector of controls includes age eligibility dummies, female, age group dummies, a Census year indicator, two region fixed effects, the log of district-level population density in 1931 and share of literate youth in 1945 and the interaction of density and literacy with trend.

Instruments include: Number of Wenela stations in the district interacted with eligibility in 1967-1973, in 1974-1977, and post-1977 dummies.

Coefficients evaluated at the mean number of migrants (3,445) and mean Kwacha per person (40). Sample includes adults ages 20 to 44 in 1977 and 1998 census. See Appendix Table 1 for first stage regressions.

Table 7: Long run effects of labor migration shocks on education: Heterogeneous effects in estate and non-estate districts

	<i>Total years of education</i>		<i>Share with any primary school</i>	
	Impacts in districts with		Impacts in districts with	
	Estates	No Estates	Estates	No Estates
	(1)	(2)	(3)	(4)
Early Treatment Cohorts*Num. Wenela stations	0.097** (0.039)	0.149*** (0.044)	0.014* (0.006)	0.014*** (0.004)
Late Treatment Cohorts*Num. Wenela stations	0.147** (0.059)	0.215*** (0.054)	0.019* (0.009)	0.021*** (0.005)
Post-Treatment Cohorts*Num. Wenela stations	0.076 (0.065)	0.200*** (0.038)	0.010 (0.008)	0.018*** (0.003)
Additional controls, district FE and trend interactions	Y	Y	Y	Y
N	160	320	160	320
R2	0.86	0.86	0.85	0.84
Mean of outcome variable	2.30	2.69	0.39	0.42

Robust standard errors clustered at the district level. Statistical significance at the 1, 5, and 10 percent levels is indicated by ***, **, and *, respectively, and evaluated relative to the small sample t distribution to account for the small number of clusters. Unit of observation is the district-5 year age group-gender cell. Estate denotes those districts which have substantial presence of tobacco and sugar estates, as described in the text. All regressions control for female, age group dummies, the log of district-level population density in 1931, the share of literate youths in 1945, historical literacy and population density interacted with a national trend, a full set of district and region fixed effects and region-specific trends. Number of Wenela stations is a count variable.

Table 8 : Effects of labor migration shocks on change in shares of children working, 1968 to 1977

	Full sample	Estate districts	Non-estate districts
	(1)	(2)	(3)
Post Dummy (=1 in 1977)*Num. Wenela stations	-0.010** (0.004)	0.195*** (0.024)	-0.013*** (0.004)
Additional controls	Y	Y	Y
District FE	Y	Y	Y
Region trends	Y	Y	Y
N	92	32	60
R2	0.71	0.86	0.76
Mean of outcome variable	0.18	0.20	0.17

Robust standard errors clustered at the district level. Statistical significance at the 1, 5, and 10 percent levels is indicated by ***, **, and *, respectively, and evaluated relative to the small sample t distribution to account for the small number of clusters. Unit of observation is the district-5 year age group-gender cell for the age group 10-14 years. Outcome is the share of children working in 1968 or in 1977. Vector of controls includes a dummy for the year of observation being 1977, female, two region fixed effects, the log of district-level population density in 1931, the share of literate youths in 1945 and historical literacy and population density interacted with post dummies. Number of Wenela stations in the district is a count variable.

Appendix Table 1: First stage estimates for IV scaling

	<i>Total Migrants (1967- 1977)*Early Treatment Cohorts</i>	<i>Total Migrants (1967- 1977)*Late Treatment Cohorts</i>	<i>Total Migrants (1967- 1977)*Post- Treatment Cohorts</i>	<i>Deferred Pay(1967- 1977)*Early Treatment Cohorts</i>	<i>Deferred Pay(1967- 1977)*Early Treatment Cohorts</i>	<i>Deferred Pay(1967- 1977)*Late Treatment Cohorts</i>
	(1)	(2)	(3)	(4)	(5)	(6)
Early Treatment Cohorts*Num. Wenela stations	1.14740e+03** (454.094)	47.663 (82.981)	38.130 (66.385)	13.779 (10.394)	1.422 (0.849)	1.137 (0.680)
Late Treatment Cohorts*Num. Wenela stations	16.822 (29.287)	1.22198e+03*** (380.489)	67.289 (117.149)	0.502 (0.300)	16.004 (10.239)	2.007 (1.199)
Post-Treatment Cohorts*Num Wenela stations	21.869 (38.074)	109.344 (190.368)	1.22535e+03*** (377.883)	0.652 (0.390)	3.262 (1.949)	16.104 (10.236)
Additional controls, district fixed effects	Y	Y	Y	Y	Y	Y
Trend interaction terms	Y	Y	Y	Y	Y	Y
N	480	480	480	440	440	440
R2	0.47	0.45	0.40	0.56	0.55	0.51
F statistic on instruments (interaction term)	7.73	7.73	7.73	2.77	2.77	2.77

Statistical significance at the 1, 5, and 10 percent levels is indicated by ***, **, and *, respectively, and evaluated relative to the small sample t distribution to account for the small number of clusters. Robust standard errors clustered at the district level. Unit of observation is the district-5 year age group-gender cell. Outcomes are eligibility dummies interacted with the change in the number (stock) of migrants between 1966 and 1977 (columns 1 to 3) or with the total deferred pay flows per capita in the district between 1966 and 1977 (columns 4 to 6). All regressions control for female, age dummies, the log of district-level population density in 1931, the share of literate youths in 1945 and a full set of district fixed effects. Trends are interacted with baseline population density, baseline literacy, and region fixed effects.

FOR ONLINE PUBLICATION ONLY

Appendix 1: Data

This appendix describes the main data sources used in the paper and the construction of main outcome and explanatory variables

1. *Education and demographic variables from 1998 Census*

- We use 100% microdata from the 1998 Census. These data are available from the Malawi National Statistics Office and from IPUMSI (<https://international.ipums.org/international/>).
- Variables include: total years of schooling attained for everyone in the data, current geographic location (region and district of the individual), age and gender. We create additional education variables: whether someone has attained any primary schooling, whether someone has completed primary schooling, and whether an individual reports being bilingual or not

2. *Education and demographic variables from the 1977 Census*

- We digitize aggregate data tables constructed from the 100% microdata of the 1977 Census, reported in *Malawi 1977 Population Census Final Report Volumes I and II*, Malawi National Statistics Office, Zomba
- Data are available at national, region and district level, and sometimes at district, sex and five-year cohort level.
- Variables we use include: total years of schooling attained by each gender-five year age group at district level, the share of each district-gender-five year age group cell that has ever been to primary school, and the cell counts for each district-gender-five year age cell.
We also use data on the number of men reporting a return from working abroad by district and five year age group, since the prior 1966 Census, and the number of boys and girls aged 10 to 19 who are employed outside the home, employed in the home, and enrolled in school

3. *Historic variables from older Census data*

- Aggregate tables presented at the district level are available from published reports for the 1931 (*Report on the Census of 1931*, Nyasaland Protectorate), 1945 (*Report on the Census of 1945*, Nyasaland Protectorate) and 1966 (*Malawi 1966 Population Census Final Report*, Malawi National Statistics Office, Zomba) Malawian Census. We digitized various tables from these reports and matched them to current definitions of district boundaries

- Variables include: the log of population density in 1931 and 1945, the share of youth who are literate (English and the vernacular) in 1945, the fraction of men employed in different sectors (farming/non-farming, working for wages/no wages, unemployed) in 1966, and the number of adult men who work abroad in 1966, reported at the district level

4. *Geographic variables*

- Altitude: we compute altitude for each point on the Malawian grid map using data from the national map seamless server (<http://seamless.usgs.gov/index.php>) and the Viewshed tool in ArcGIS. We aggregate these measures to district level.
- We define areas of high, medium or low malaria susceptibility based on standard measures of altitude: high malaria areas (altitude below 650m), medium malaria areas (altitudes between 650m and 1100m) and low malaria areas (altitudes over 1100m).
- We create a district boundary crosswalk that links districts over time (across Census waves) and across name changes. We assign variables measured in earlier years to later Census district boundaries in this way:
 - For districts that were eventually combined in later years, we add district level values together
 - For districts split apart in later years, we apportion district totals to split districts using the fraction of physical area that each split district accounts for within the total district.
- We identify which districts contain a large tea or tobacco plantation using information in Christiansen (1984). The FAO's crop suitability index measuring whether a district is highly suitable for tobacco or tea production significantly predicts this estate district indicator. Because of the coarseness of this measure, there is uncaptured variation within estate districts in the prevalence of estate lands out of total district agricultural lands.

5. *Administrative data*

- Figure 1 is constructed using the location of Wenela/TEBA recruiting stations in 1937. We collected and digitized this historical data using a variety of sources. The main source included "Correspondence from the Secretariat, Zomba, Nyasaland 1935 (Circular number 8 1935, S1/169/35). We verified these stations were still open in later years using information from later Provincial Administration Reports (Northern Province: 7th December 1961 Ref. No. O.3.37 and Commissioner for Labour Circular, 25th March 1957)

- Figure 2 is constructed using national labor migration totals from a variety of sources including: Chirwa (1991 for years 1950-1958); Lipton (1980: for years 1959-1994); Crush, Jeeves and Yudelman (1991: pp234-235) and various years of TEBA Annual Reports

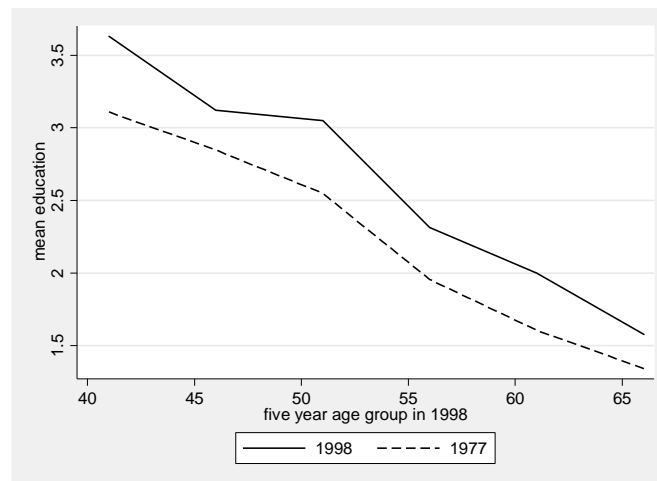
FOR ONLINE PUBLICATION ONLY

Appendix 2: Diagnosing selective attrition in the 1998 Census

Because life expectancy in Malawi was only 46 years in the late 1990s (http://www.theglobaleconomy.com/Malawi/Life_expectancy/), we are concerned about mortality selection at ages over 40 affecting the composition of our sample. Specifically, we worry that this mortality selection may be differential across districts with and without recruiting stations. In this appendix, we use the 1977 and 1998 Census to diagnose this selective attrition and motivate our use of the 1977 Census data to construct estimates of educational attainment among the older cohorts in our analysis.¹

Assuming that education is completed by age 20 in 1977, we can compare mean educational attainment for five year cohorts of those aged 20 to 44 in 1977 with the mean educational attainment of the five year age cohorts for those aged 41 to 65 in 1998.² At one extreme, if there is no mortality at all, mean education rates should match up for the same cohorts across Census waves. If there is any attrition (mortality) of those less educated between 1977 and 1998, then the mean education gap by cohort should be positive. Appendix 2 Figure 1 below shows just this.

Appendix 2 Figure 1: Mean education by cohort in 1998 using 1977 and 1998 Census data



Mean years of education (on the y axis) for each age group is higher in 1998 than for the corresponding age group in 1977, suggesting higher mortality rates among the less educated between 1977 and 1998.

¹ The question “What is the highest level of schooling you have attended?” is identical in the 1977 and 1998 Census’ and the same coding system for different levels of education is used in both waves.

² Note that we cannot do this reliably for younger cohorts in 1977 (i.e for those 36-40 in 1998), since many of those under 20 are still in school.

Of greater concern for our identification strategy is differential mortality selection of the less educated in *Wenela* districts. The table below presents coefficients from a regression of the gap in mean completed years of education at district level (1998 levels minus 1977 levels) on dummies for each five year age cohort and interactions of each cohort dummy with number of recruiting stations in the district (the constant coefficient and coefficient on the number of recruiting stations are suppressed). All *Wenela* interaction terms are positive, and the interaction terms are jointly significantly different than zero. This means that the less-educated cohorts in *Wenela* regions are less likely to survive to 1998 than the less-educated cohorts in non-*Wenela* regions.

Appendix 2 Table 1: Diagnosing selective attrition between 1977 and 1998

Age 40-44*Number of <i>Wenela</i> Stations	0.0413 (0.0255)
Age 45-49*Number of <i>Wenela</i> Stations	0.0288 (0.0212)
Age 50-54*Number of <i>Wenela</i> Stations	0.0508*** (0.0192)
Age 55-59*Number of <i>Wenela</i> Stations	0.0377** (0.0173)
Age 60-64*Number of <i>Wenela</i> Stations	0.0313* (0.0183)
Age 40-44	0.348*** (0.0938)
Age 45-49	0.193** (0.0818)
Age 50-54	0.405*** (0.0878)
Age 55-59	0.322*** (0.0742)
Age 60-64	0.423*** (0.0821)
N	135
R2	0.654

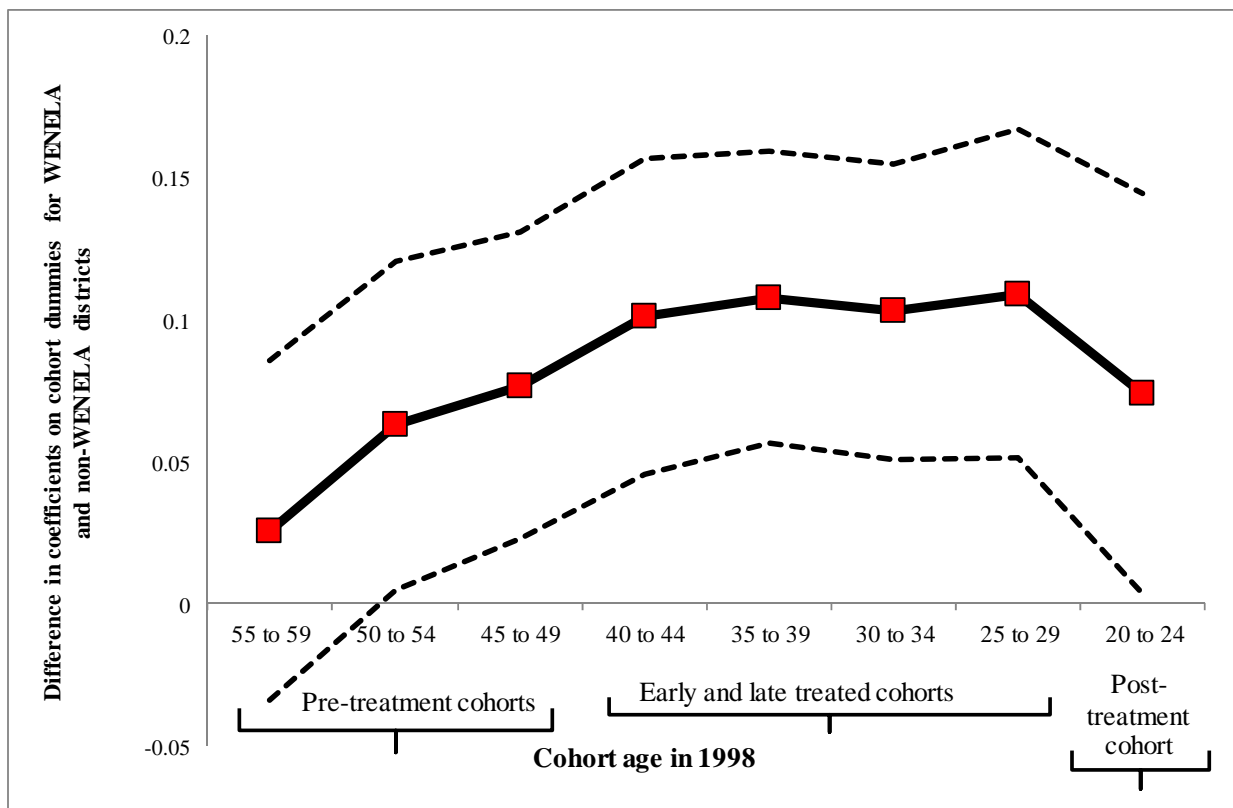
N=189. Observation is the district-five year cohort. Outcome is the difference in mean level of education (1998 and 1977) measured for the district-five year cohort. Robust standard errors. F-statistic for interaction terms is 3.83 (p value is 0.003)

Together, these results suggest that using only the 1998 Census for our analysis would result in downwards-biased estimates of education gaps across *Wenela* and non-*Wenela* districts in the pre-treatment cohorts. That is, we would end up controlling for larger positive education gaps between the older cohorts in *Wenela* relative to non-*Wenela* districts, and would estimate smaller effects on educational attainment among the treatment cohorts in our difference-in-differences strategy.

We can illustrate the impact of ignoring the differential selective mortality across the *Wenela* and non-*Wenela* districts on our estimates. In Appendix 2 Figure 2, we present point estimates from the difference-

in-differences specification that generates our main result in Figure 4 of the paper. Instead of constructing the synthetic cohort of older pre-treatment age groups (ages 45 and over) from the 1977 Census as we do in Figure 4, we only use the 1998 Census to produce the graph below. The results in Appendix 2 Table 2 indicate that among the older age groups, less educated individuals have been differentially selected out in *Wenela* areas. Because we lose these people by 1998, the figure indicates an increase in the education gap between *Wenela* and non-*Wenela* cohorts even among the pre-treatment cohorts. This is the effect of differential selective attrition in *Wenela* districts.

Appendix 2 Figure 2: Estimated differences in education by age group and *Wenela* status of district using 1998 cohorts only



We still see the largest impacts on education gaps between *Wenela* and non-*Wenela* areas concentrated in the Early and Late treatment cohorts, as we do in our main results. However, the size of the effect is smaller because of the bias coming from pre-treatment differences in the education gap. The point estimates that correspond to the specification used to produce our main results in Table 5 but using only the 1998 Census in Appendix 2 Table 2. The difference in differences terms are still significantly different than zero for years of education, share with any primary schooling and two additional outcomes,

the share literate in English and the share bilingual (English and Chichewe), although smaller, as we might expect from the discussion in this appendix.

Appendix 2 Table 2: Long run effects of labor migration shocks on education: Difference-in-differences results 1998 data only

	<i>Total years of schooling attained</i>		<i>Share with any primary schooling</i>		<i>Share English literate</i>		<i>Share bilingual</i>	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Early Treatment Cohorts*Num. Wenela stations (β_1)	0.064** (0.024)	0.060*** (0.020)	0.004* (0.002)	0.004** (0.002)	0.006** (0.003)	0.005** (0.002)	0.006** (0.003)	0.005** (0.002)
Late Treatment Cohorts*Num. Wenela stations (β_2)	0.066** (0.028)	0.059*** (0.018)	0.003 (0.002)	0.004 (0.003)	0.008** (0.004)	0.005*** (0.002)	0.008** (0.004)	0.005*** (0.002)
Post-Treatment Cohorts*Num. Wenela stations (β_3)	0.034 (0.028)	0.026 (0.022)	-0.002 (0.003)	0.000 (0.004)	0.004 (0.004)	0.001 (0.002)	0.004 (0.004)	0.001 (0.002)
Trend interactions	N	Y	N	Y	N	Y	N	Y
N	432	432	432	432	432	432	432	432
R2	0.96	0.97	0.95	0.95	0.95	0.96	0.95	0.96
Mean of outcome variable	3.66	3.66	0.60	0.60	0.36	0.36	0.36	0.36
p value of F test $H_0: \beta_1 = \beta_2$	0.88	0.96	0.49	0.87	0.46	0.76	0.46	0.77
p value of F test $H_0: \beta_1 = \beta_3$	0.17	0.15	0.04	0.17	0.32	0.11	0.32	0.11
p value of F test $H_0: \beta_2 = \beta_3$	0.02	0.04	0.01	0.04	0.02	0.02	0.02	0.02

Robust standard errors clustered at the district level. Statistical significance at the 1, 5, and 10 percent levels is indicated by ***, **, and *, respectively, and evaluated relative to the small sample t distribution to account for the small number of clusters. Unit of observation is the district-five year age cohort-gender cell. Vector of controls included in every regression: female dummy, cohort dummies, two region fixed effects, the log of district-level population density in 1931 and the share of literate youths in 1945, as well as district fixed effects. In the second specification for each outcome, trend terms are interacted with region fixed effects, baseline population density and baseline literacy rates. Number of Wenela stations in the district is a count variable. Sample includes adults ages 20 to 64 in the 1998 census.

FOR ONLINE PUBLICATION ONLY

Appendix 3: Bounding results for composition effects from internal migration

Internal migration poses one possible threat to the validity of our main results. Because neither the 1977 nor 1998 Census captures district of birth, we potentially mismeasure childhood exposure to Wenela recruiting stations among those people who move across districts after completing education, but before we see them in the relevant Census year. Internal migration flows are unlikely randomly allocated across districts. Without knowing more about differences in the magnitude and direction of migrant flows across districts, this possible misclassification of exposure to Wenela stations generates unpredictable biases in our estimates.

To see why, consider the following. Suppose all districts have the same average level of education before internal migration. If more educated adults move from non-Wenela to Wenela districts while less educated adults move in the opposite direction, this generates artificially positive differences in adult educational attainment across districts that we would ascribe to exposure to Wenela stations. As long as this sorting is constant over time, then the Wenela dummy in our regressions (as well as district fixed effects) controls for these observed differences in educational attainment driven by internal migration. If, however, internal migration flows differ by district as well as cohort then our results could be the result of complicated changes in the composition of population at the district level.

In the absence of individual level data on birth districts, we bound our effect sizes for possible composition changes induced by internal migration. We combine information on net migration rates from the 1977 Census with assumptions about possible levels of education of net migrants. First, we use 1977 Census data to construct the number of net migrants per person currently living in the district for each district in each five-year cohort and gender cell. We call this the net migration rate, or $NetMigRate_{asd}$. In our data, this number is always between -0.35 and 0.29.¹ We need to assume that this net migration rate is the same in 1977 and 1998, since the 1998 Census contains no information on district of birth. Second, we assume that all migrants – whether they show up as in- or outmigrants in a particular district – have

¹ Census 1977 counts the number of people in each district, cohort, and gender cell and enumerates how many of these individuals were born in each district. The net migration rate is computed as the difference between total in-migrants and total out-migrants divided by total current population in the district; it is the number of net migrants (in-migrants – out-migrants) per person living in the district. A 0.2 net migration rate means that for every person living in the district, there are 0.2 net in-migrants.

the same level of education and therefore we need only account for the potential education of net migrants, the difference between in- and outmigrants.²

We adjust our education variables (\bar{Y}_{asd}) measured at district, cohort, and gender level:

$$\begin{aligned}\bar{Y}_{asd}^{BOUND} &= \frac{N_{asd}\bar{Y}_{asd} - NetMigrants_{asd} * \bar{Y}_{as}^m}{N_{asd} - NetMigrants_{asd}} & (A.1) \\ &= \frac{N_{asd}\bar{Y}_{asd} - NetMigRate_{asd} * N_{asd} * \bar{Y}_{as}^m}{N_{asd} - NetMigRate_{asd} * N_{asd}} \\ &= \frac{\bar{Y}_{asd} - NetMigRate_{asd} * \bar{Y}_{as}^m}{1 - NetMigRate_{asd}}\end{aligned}$$

where $BOUND = \{upper, lower\}$, \bar{Y}_{asd}^{BOUND} represents the adjusted mean education outcome at district, cohort, and gender level, N_{asd} is total population in a district-cohort-gender cell, \bar{Y}_{as}^m is either the maximum or minimum value of the relevant education variable across all districts at cohort and gender level, and $NetMigrants_{asd}$ is the total number of net migrants in a district-cohort-gender cell.

$NetMigrants_{asd}$ is estimated by multiplying the total population in that district-cohort-gender cell with the net migration rate ($NetMigRate_{asd}$) for that cell. Each component of (A.1) comes from the relevant Census wave, except for $NetMigRate_{asd}$ which is computed using 1977 Census data and applied to both Census waves. We estimate the main regression specifications for our sample after creating these adjusted education variables, one set for each of the extreme values of \bar{Y}_{as}^m , or

$$\begin{aligned}\bar{Y}_{asd}^{lower} &= \frac{\bar{Y}_{asd} - NetMigRate_{asd} * \bar{Y}_{as}^{max}}{1 - NetMigRate_{asd}} \text{ and} \\ \bar{Y}_{asd}^{upper} &= \frac{\bar{Y}_{asd} - NetMigRate_{asd} * \bar{Y}_{as}^{min}}{1 - NetMigRate_{asd}}.\end{aligned}$$

There are two notable features of equation (A.1). First, the adjustments we make for internal migration imply that \bar{Y}_{asd}^{upper} and \bar{Y}_{asd}^{lower} provide upper and lower bounds on mean education and average share of adults with any primary school across the entire sample. Second, despite these names, these adjustments do not imply that the difference-in-differences regressions using these new variables will produce estimates that contain our main education results. This is because in a closed system (i.e. the whole of Malawi) some districts are receiving districts ($NetMigRate_{asd} > 0$) while others are sending districts

² For example: if there are 110 in-migrants and 100 out-migrants to a particular district, and in-migrants and out-migrants have the same levels of education, the only change in composition that occurs as a result of this net migration is due to the additional 10 people who migrated into the district.

($NetMigRate_{asd} < 0$). In order for $\bar{Y}_{asd}^{upper} > \bar{Y}_{asd}$ or $\bar{Y}_{asd}^{lower} < \bar{Y}_{asd}$, the following equations should hold (note that in our sample, $1 - NetMigRate_{asd} > 0$ in all cases):

$$NetMigRate_{asd} * (\bar{Y}_{asd} - \bar{Y}_{as}^{min}) > 0 \quad (A.2)$$

$$NetMigRate_{asd} * (\bar{Y}_{asd} - \bar{Y}_{as}^{max}) < 0 \quad (A.3)$$

Since $\bar{Y}_{asd} \geq \bar{Y}_{as}^{min}$ and $\bar{Y}_{asd} \leq \bar{Y}_{as}^{max}$ in all districts, these equations are only satisfied for receiving districts that have $NetMigRate_{asd} > 0$. To see this, assume that we impute the minimum level of education for net migrants, Then, $\bar{Y}_{asd}^{upper} > \bar{Y}_{asd}$ is satisfied only in receiving districts because our adjustments take out the low levels of education of net in-migrants to create a higher adjusted mean education variable. For sending districts, where $NetMigRate_{asd} < 0$, the inequality in (A.2) is reversed and $\bar{Y}_{asd}^{upper} < \bar{Y}_{asd}$. Similarly, when we impute the maximum level of education for net migrants, equation (A.3) will only be satisfied in receiving districts; subtracting high levels of net in-migrant education generates $\bar{Y}_{asd}^{lower} < \bar{Y}_{asd}$. In sending districts, (A.3) is reversed, so $\bar{Y}_{asd}^{lower} > \bar{Y}_{asd}$.

Because we have both sending *and* receiving districts in our sample, and because rates of internal migration in 1977 are different across Wenela and non-Wenela districts (rates of in-migration are higher in Wenela districts, results not shown), our adjustments have different effects on the bounds values in specific Wenela and non-Wenela districts. More complicated patterns of net migration that vary across exposed and non-exposed cohorts and across Wenela and non-Wenela areas imply that adjustments for internal migration may generate in difference-in-differences estimates that do not bound our main result.³ Nevertheless, it is still a useful exercise to check whether internal migration modelled in this way appears to confound our results.

Appendix Table 2 displays results from difference-in-differences regressions estimated using the adjusted education variables, first including all controls and district fixed effects, and then adding in trend interactions with region fixed effects, baseline literacy and baseline population density. We compare the coefficients in this table with the main estimates in Table 5.

First, assuming net migrants have the maximum level of schooling in the district-cohort-gender cell for a given Census year, the presence of anyone new in a receiving district raises mean education and their absence from a birth district artificially deflates that district's average education. Adjusting for these

³ Crudely, if net migration rates are more likely positive in Wenela districts among exposed cohorts, we would be doing more "receiving district" adjustments in our core treatment groups and more "sending district" adjustments in our control groups.

educated net migrants, we still see large, positive impacts of exposure to treatment among exposed cohorts: those exposed during the labor expansion have 0.14 more years of education, while those exposed during the labor contraction have 0.2 more years of education. The education gap between Wenela and non-Wenela districts for the post-treatment group continues to be positive, at around 0.16 more years of education. Second, if we instead assume that net migrants are uneducated, removing them from our outcome measure in receiving districts and adding them back to sending districts reveals similar, large positive impacts of exposure to mine employment shocks. The difference-in-differences estimates in columns (3) and (4) imply that exposed cohorts in Wenela areas gained between 0.16 and 0.25 more years of education. Post-treatment groups continue to have about 0.22 more years of education in Wenela relative to non-Wenela areas, controlling for differences between these areas using the oldest pre-treatment cohorts. These bounds compare favorably to our main results in Table 5, 0.12 and 0.179 more years of education (Table 5, column 3).

Results are similar when we use the share with any primary school as outcome. In Table 5, directly exposed cohorts from districts with more Wenela stations are 1.1 to 2.6 percentage points more likely to have ever attended primary school. After adjusting for internal migration in Appendix Table 2 columns (5)-(8), these exposed cohorts from districts with Wenela stations are between 1.2 and 2.4 percentage points more likely to have ever been to primary school. All of our estimates are statistically different from zero at the 1, 5 or 10% level. Looking at the final three rows of the table, we see that we can strongly reject that the impacts of the migration shock on education are the same for the labor expansion and labor contraction cohorts, and we can reject that the impacts on primary school access for the labor contraction cohorts and the youngest post-treatment cohorts are the same. However, as in the case of our main results in Table 5, we cannot reject that the impacts for the post-treatment cohorts and the labor expansion cohorts are the same. We see the same inverted U-shaped pattern of coefficients in our bounded results as in the main results. The results of this bounding exercise suggest that selective internal migration and any resulting measurement error in $Wenela_d$ cannot account for our main effects.

Appendix 3 Table 1: Long run effects of labor migration shocks on education: Bounds for internal migration

Assumptions about migrant education:	<i>Total years of education</i>				<i>Share with any primary school</i>			
	Max. schooling		Min. schooling		Highest share with primary school		Lowest share with primary school	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Early Treatment Cohorts*Num. Wenela stations (β_1)	0.142*** (0.048)	0.136*** (0.044)	0.163** (0.059)	0.164** (0.063)	0.014*** (0.005)	0.012*** (0.004)	0.015*** (0.005)	0.014*** (0.005)
Late Treatment Cohorts*Num. Wenela stations (β_2)	0.202*** (0.067)	0.191*** (0.060)	0.251*** (0.084)	0.252** (0.090)	0.020*** (0.007)	0.016*** (0.005)	0.024*** (0.007)	0.021** (0.007)
Post-Treatment Cohorts*Num. Wenela stations (β_3)	0.168** (0.062)	0.153** (0.062)	0.220*** (0.076)	0.222** (0.088)	0.015** (0.005)	0.011* (0.006)	0.019*** (0.006)	0.015* (0.007)
District FE	Y	Y	Y	Y	Y	Y	Y	Y
Region trends	N	Y	N	Y	N	Y	N	Y
N	480	480	480	480	480	480	480	480
R2	0.83	0.86	0.83	0.85	0.824	0.84	0.82	0.83
Mean of outcome variable	2.53	2.53	2.59	2.59	0.41	0.41	0.41	0.41
p value of F test $H_0: \beta_1=\beta_2$	0.01	0.00	0.00	0.01	0.01	0.02	0.00	0.02
p value of F test $H_0: \beta_1=\beta_3$	0.22	0.56	0.02	0.09	0.56	0.69	0.10	0.76
p value of F test $H_0: \beta_2=\beta_3$	0.01	0.09	0.04	0.12	0.01	0.07	0.01	0.04

Statistical significance at the 1, 5, and 10 percent levels is indicated by ***, **, and *, respectively, and evaluated relative to the small sample t distribution to account for the small number of clusters. Robust standard errors clustered at the district level. Unit of observation is the district-5 year age group-gender cell. Vector of controls includes female, age group dummies, two region fixed effects, the log of district-level population density in 1931 and the share literate in 1931, and historical density, historical literacy and region fixed effects interacted with a trend term. Number of Wenela stations in the district is a count variable. Outcomes are our estimates of the bounds on education and share in primary school, after accounting for maximum and minimum possible values of each variable for the number of net migrants in each age-gender cell. Details of variable construction are explained in the text. Sample includes adults ages 20 to 44 in 1977 and 1998 census.

FOR ONLINE PUBLICATION ONLY

Appendix 4: Constructing exact p values for Table 5 results using randomization inference

Our empirical strategy exploits pre-existing spatial variation in migration costs across 24 districts within Malawi. The relatively small number of districts leads to a concern that standard inference procedures will over-reject the null hypothesis of zero impact of district-level exposure to the labor migration shock. We deal with this concern by presenting robust standard errors clustered at the district-level in Table 5 and obtain p values to indicate statistical significance using the small sample t distribution adjusted for the number of covariates that are constant within the cluster.

An alternative approach is to construct exact p values using randomization inference (Fisher, 1935; Rosenbaum 2002; see Cohen and Dupas 2010 for an example of how this is done in the context of a randomized controlled trial with 16 clusters and a single treatment). The idea behind this approach is as follows:

- We randomly assign the actual distribution of Wenela stations to districts and estimate the difference-in-difference models of Table 5 for this “false” assignment. The false allocation mimics the true distribution of stations
- Since there are over 1.3 million ways to allocate Wenela stations (the range of this variable is 0 to 10) to the 24 districts, we generate 1,000 different random assignments of stations to districts and estimate the difference-in-differences regression for each allocation.
- We compute the empirical distribution of t statistics for each of the three main parameters (β_1 , β_2 and β_3) generated by these 1,000 false assignments
- We compare the actual t statistics from Table 5 to the empirical distribution of test statistics for each parameter and compute the probability of observing a t statistic in the tails of this distribution. The resulting p values, denoted randomization inference p values are presented in the table below.

In all cases, we can reject the null of zero impact at the 10% level. In most cases, for the estimates of β_1 and β_2 , we can also reject the null of zero impact at the 5% level.

References

- Cohen, Jessica and Pascaline Dupas, 2010 “Free distribution or cost-sharing? Evidence from a randomized malaria prevention experiment”, *Quarterly Journal of Economics*, Vol. 125 (1): 1-45
- Fisher, Ronald A. 1935 The Design of Experiments London: Oliver and Boyd.
- Rosenbaum, Paul R. 2002 Observational Studies New York: Springer-Verlag.

Appendix 4 Table 1: Randomization inference p values for difference-in-differences estimates of Table 5

p value for:	<i>Outcome is: Years of schooling</i>		<i>Outcome is: Share with any primary school</i>	
$H_0: \beta_1=0$	0.001	0.001	0.001	0.016
$H_0: \beta_2=0$	0.009	0.012	0.013	0.011
$H_0: \beta_3=0$	0.013	0.031	0.009	0.084
Other controls?	Y	Y	Y	Y
District FE?	Y	Y	Y	Y
Trend interactions	N	Y	N	Y

FOR ONLINE PUBLICATION ONLY

Appendix 5: Correlations between human capital formation and household well-being in Malawi

Given concerns about the quality of education provided in poor countries, it is worthwhile asking whether there is any evidence that more years of schooling and more skills learned in school are valuable in Malawi. We use data from the 2004 Malawi Integrated Household Survey to estimate the relationship between total assets owned by the household (as a summary measure of household well-being) and human capital attainment of the head of the household. There are too few wage earners in the data to examine the relationship between individual education and wages.

We restrict the sample to all rural households in the survey where the head is between the ages of 26 and 60 in 2004, leaving us with 85% of the initial survey. We calculate total assets owned by the household (the range is 0 to 31; the mean is 6.9) and regress this index on four different measures of human capital of the household head: total years of schooling attained, whether the head has been to primary school, and self-reported literacy in English and Chichewa (the person reports that they can read a one page letter in the relevant language). Using the education of the head of household (rather than the maximum level of education in the household) is a meaningful way to characterize how much human capital the household has access to, since there are few three-generation households in the sample. Literacy reflects skills acquired in the lower levels of primary school: 88% of household heads with four years of completed education report being literate in Chichewa, and literacy rates increase from 47% to 72% between two and three years of total schooling.

Appendix 4 Table 1 presents the results of regressing our asset index on the four human capital measures for three specifications: one without controls, a second adding controls for age and gender of the household head, household size, historical district characteristics (literacy, population density) and region fixed effects, and a third that includes district fixed effects. Results indicate a robust, large and positive relationship between each human capital measure and the household asset index. In the first three columns of the table, we see that an additional year of schooling is correlated with about one third more total assets (columns 1-3), which is a 5-6% return per year of schooling. For a household head with any primary school, assets are 7% higher (columns 4 and 5). Returns to literacy are particularly large: having a literate household head raises total assets in the household by over two; a 30% gain. While usual concerns about selection and measurement error caution us against interpreting these point estimates as exactly causal, the strong positive relationship between human capital attained in childhood and measures

of household well-being in adulthood provide some evidence that education, and skills learned at school, are indeed valuable in Malawi.¹

¹ Our results are consistent with findings in Chirwa and Matita (2009), who use Mincerian regressions to show that among wage earners working in urban areas of Malawi, the return to completed primary education is 5 percent.

Appendix 5 Table 1: Correlation between household asset index and human capital of household head

	<i>Outcome: Total assets owned by household [mean=6.9]</i>											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
<i>Human capital of household head</i>												
Years of education	0.380***	0.362***	0.361***									
	(0.023)	(0.023)	(0.019)									
Any primary school				0.253	0.471***	0.473***						
				(0.165)	(0.148)	(0.146)						
Literate in English							2.531***	2.217***	2.185***			
							(0.129)	(0.130)	(0.130)			
Literate in Chichewa										2.435***	2.061***	2.005***
										(0.148)	(0.166)	(0.147)
N	6,771	6,771	6,771	6,771	6,771	6,771	6,794	6,794	6,794	6,794	6,794	6,794
Effect size: % of mean assets	6%	5%	5%	4%	7%	7%	37%	32%	32%	35%	30%	29%
Additional controls?	N	Y	Y	N	Y	Y	N	Y	Y	N	Y	Y
District FE?	N	N	Y	N	N	Y	N	N	Y	N	N	Y

Robust standard errors clustered at the district level. Statistical significance at the 1, 5, and 10 percent levels is indicated by ***, **, and *, respectively. Unit of observation is the household; household head's level of education is the main regressor in each specification. Sample includes all rural households where the household head is between the ages of 26 and 60 in 2004. Additional controls include age of household head, whether the head is female, household size, the log of district-level population density in 1931, the share of literate youths in the district in 1945, and region fixed effects. All regressions are weighted using household weights. Data are from the 2004/2005 Malawi Integrated Household Survey.