# What are the long run effects of labor migration on human capital? Evidence from Malawi

Taryn Dinkelman and Martine Mariotti[1]

First draft: October 31, 2013

This draft: March 2014

Abstract: Circular labor migration is a core feature of low-income labor markets. Yet, evidence on how this migration affects education investments in sending communities is limited due to lack of high quality data and challenging identification issues. This is especially true in Africa, where children can substitute for migrant adult labor. In this paper, we estimate the net effect of international migration on human capital accumulation of children by exploiting two large migrant labor shocks in sending communities in Malawi. An international mine labor treaty signed in 1967 initiated a 300% increase in the flow of Malawians to South Africa. Seven years later, a mining plane crash prompted the Malawian government to halt and reverse this expansion until 1977. Our strategy compares differences in long run human capital accumulation across high and low shock areas, among cohorts eligible and ineligible for primary school during the shock years. We use historical locations of mining recruiting stations to construct measures of district-level exposure to this rapid expansion and contraction of foreign employment and earnings. We match this spatial variation in migration costs to cohort-specific education outcomes from newly digitized 1977 and 1998 Census data. Both shocks to migration had large, positive impacts on education. Age eligible cohorts with the highest shock exposure attained 10 to 15% more schooling and the share with any primary schooling rose by 5 to 8%. Neither school supply-side interventions nor internal migration dynamics account for our results. These long run effects are only apparent in districts without agricultural estates, where child labor is less substitutable for missing male labor.

JEL CODES: O15, O55, O12, J24

Keywords: labor migration, human capital, child labor, Africa, agricultural production technologies

[1] Dinkelman: Dartmouth College, NBER and BREAD, Taryn.L.Dinkelman@Dartmouth.edu. Mariotti: Australian National University, martine.mariotti@anu.edu.au. Freed 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 the Rhodes House Library for their invaluable assistance in data collection.

## 1. Introduction

International labor migration is one of the most powerful ways in which poor households can improve their standard of living (Clemens 2011). Such migration, increasingly important in poor countries, often entails large remittance flows.[2] Despite these large capital and labor flows, there is no consensus in the economic literature on whether and how international labor migration affects human capital accumulation over the longer run. Remittance research focuses either on understanding motives for remitting, or on detecting income effects of remittances on schooling and other expenditures among credit-constrained households (Yang 2008, 2011). Debates about human capital externalities of labor migration focus on whether migration encourages brain drain (Collier et al 2004; Stark 2004; De Braauw and Giles 2008; Docquier and Rapoport 2012), or its alternative, brain gain, among left-behind adult members of the household who are incentivized to get more schooling (Gibson, McKenzie and Stillman 2011; McKenzie and Rapoport 2011). These different strands of research often overlook the direct impact of losing unskilled adult labor on human capital attainment in contexts where children can substitute for adults in production. Such impacts on the next generation could be particularly important in affecting long run human capital stocks in a country, and consequently, economic growth.

In theory, adults migrating away from rural economies might improve or undermine educational attainment of the next generation. If migrants send money home and households get wealthier, credit constraints may be relaxed and the demand for schooling may increase, especially where school expenses are high. However, the loss of male labor in agricultural production may require families and farms to substitute towards child labor, thereby raising the opportunity cost of going to school. Whether these substitution effects dominate the effects of rising migrant incomes or not is an empirical question. And, the size and sign of the net effect on education is likely to depend on the technology of production of the local economy, in particular, how well child labor substitutes for adult labor.

This paper asks two related research questions: Does adult labor migration increase or lower total human capital attained by the next generation? And, do these impacts of labor migration depend on local technologies of production? We address these questions in the context of international labor migration of unskilled male workers from Malawi to South Africa between 1950 and 1990. We use an unexpected and large expansion and subsequent sudden contraction of mine employment to test whether or not the substitution effect outweighs the effects of income on the demand for education in a rural, agricultural economy, and to find out whether the impacts of this labor migration had long-lasting effects on human capital stocks.

[2] Global remittance flows currently exceed flows of official development assistance to poor countries (Yang 2011).

Beginning in the early colonial period, Malawi has had a long history of labor migration to the South African gold mines. This migration entailed sending men abroad for long periods of time to earn substantially higher incomes than they could at home. In 1967, a newly signed labor treaty between Malawi and South Africa opened the way for a 300% increase in the flow of Malawians to South African mines. In the following seven years, over 20% of adult men became contract mineworkers. Access to these foreign jobs and income earning opportunities came to an abrupt halt after a mining plane crashed in April 1974 and killed a group of Malawian miners. This event precipitated a four-year ban on all labor migration to the mines, leading to the immediate return of over a hundred thousand men. We use both sides of this natural experiment to test whether income or substitution effects dominate in the demand for education and to understand more generally the role that local agricultural production technologies play in determining the net effect of this migration.

To identify these human capital effects, we use a difference-in-differences research design that combines spatial and cohort-level variation in exposure to these employment shocks. We estimate long run effects on education of missing male migrants and their remittances during the massive employment expansion and of the returning male migrants and their deferred payments during the labor ban by combining district level variation in the costs of accessing a mining job with cohort variation in exposure to these foreign mining employment shocks during primary school age eligibility years. Our measures of education (total years of schooling, share of cohort with any primary school) come from newly digitized 1977 Census matched and 1998 Census data matched at cohort, sex and district level. We define exposed cohorts as adults who were age eligible for primary school during each of the labor expansion and contraction periods; our comparison cohorts are eligible before or after these periods. All of these cohorts had to pay to go to school, unlike the current policy of free primary school. To create spatial variation in the costs of labor migration that help us to measure high and low exposure to these foreign employment shocks, we use newly collected administrative data on historical locations of mine recruiting stations within districts across the country.

As a specific test of whether the local technology of production matters for how international labor migration affects the demand for education, we estimate difference-in-differences regressions for two different areas: estate and non-estate districts. As we discuss in the background section, child labor is a good substitute for adult male labor on estates, where tea and tobacco plantations produce cash crops for export (Eldring 2003, Chirwa 2005 and Otañez 2006). In non-estate districts without large cash crop plantations, farms and smallholdings are less suited to using child labor in place of men. We test whether the impacts of international migration differ across areas with these pre-existing local production technologies.

The historical experience of labor migration from Malawi provides an interesting and advantageous setting for our research questions. First, there is only a single selection problem to consider – only adult men (and not the entire household) were permitted to move. Second, unlike many other studies of migration, migrants were required to return home after two years and required to have two thirds of their earnings diverted into deferred pay. This means we can be sure that flows of capital back to the rural areas were substantial and largely exogenously determined. Third, rather than compare outcomes across households with and without migrants where we might worry that household level unobservables influence both the decision to migrate and the decision to invest in education, we compare outcomes across districts with historically different access to mine jobs. Although international migration to South Africa was a large phenomenon in most districts of Malawi (meaning that potentially all districts exposed to the foreign employment shocks), proximity to a mine recruiting station substantially reduced the costs of migrating, which we explain in Section 3. Since these recruiting stations were established at least thirty years before the 1967 labor treaty and 1974 labor ban, we use the location of these stations as a pre-determined measure of differences in the costs of labor migration. As demand from the South African mines surges after 1967 and then is choked off in 1974, districts with the lowest historical costs of sending men to the mines experience these employment shocks more intensely. Our difference-in-differences strategy therefore identifies the long run impact on education of increases in international migration induced by lower migration costs.

Although districts with and without access to these recruiting stations differ on some baseline variables, we are able to show that these districts share the same trends in education outside of the labor expansion and contraction periods. We can control for a range of observable and unobservable differences across high and low migration cost districts by adding controls for historical and geographic variables that could have affected station placement in the past, as well as district fixed effects and region-specific trends. Under the assumption that there are no other contemporaneous shocks to district- and cohort-specific education concentrated in recruiting station districts, our difference-in-differences estimator gives us the causal impact of labor migration and related remittance earnings on long run human capital attainment.

Our results indicate that in only ten years, exposure to mining employment shocks and concomitant migrant remittances enabled districts with mine recruiting stations to overtake districts without recruiting stations in their total amount of human capital, with long-lasting effects. The main results suggest that greater exposure to international labor migration contributed to a 10 to 15% gain in total years of education and a 5 to 8% increase in any primary schooling attendance. The primary school enrollment effects account for between one third and one half of the total increase in enrollment rates between 1967 and 1978.

These large positive net effects are evident for those eligible for primary school during the labor expansion years, as well as for those eligible during the labor ban years. Although mine work dries up during the labor ban, we use the historical evidence to show that that returning miners would have had up to three years' worth of earnings to draw down upon return, meaning that the income effects are likely larger for the later treated cohorts, while substitution effects are reversed in sign.

Our estimates cannot be explained by school supply-side programs during this period; the number of primary schools in Malawi did not grow between 1960 and 1990. We also show that our results are robust to adjustments for patterns of internal migration between district of birth and district of current residence, a key potential threat to validity.

The paper's main results suggest that the effects of income on human capital accumulation of children outweigh the substitution effect of adult labor migration. However, we also find some evidence that substitution technologies in agricultural production contribute to these long run education impacts. The positive impacts of labor migration on education are only present among districts that do not have large agricultural estates. That is, we estimate positive impacts for total years of schooling and share of population with any primary school, but only in areas where child labor was less substitutable for male labor. This indicates a role for local production technologies to determine the impacts of labor migration on sending communities.

This paper is part of a broader research agenda trying to pin down how international, circular labor migration has long-run effects on rural agricultural labor markets. Our paper relates to several literatures. At the broadest level, our finding that migration promotes educational attainment in the next generation is hopeful for the very poorest of countries, where the migration of one adult towards places of improved employment opportunities is often the best escape from poverty for an entire household. The empirical literature on the effects of international migration on education outcomes has been mixed (Yang 2008, Gibson et al 2011, Edwards and Ureta 2003, McKenzie and Rapoport 2011)[3], often focused on middle income countries, and almost non-existent in the case of Africa. Our paper deepens the evidence for Africa, and broadens the discussion by pointing out that local production technologies matter for the human capital effects of migration. These positive effects are also not necessarily limited to international

[3] Cox-Edwards and Ureta (2003) and Yang (2008) both find positive impacts of migrant remittances on education spending and education outcomes. McKenzie and Rapoport (2011) estimate the impact of Mexican migration to the US on educational attainment of teen children left behind using historical migration as an instrument for current migration. They find that migration of a parent significantly reduces the total amount of schooling attained and provide evidence that boys drop out of school in order to migrate (along the lines of a "brain drain" effect) while girls who drop out do more housework. The mechanism they have in mind is that migration possibilities for teens reduce the expected future return to school because such returns are lower in the US than in Mexico. There is no explicit discussion of child labor substituting for adult labor in the Mexican context.

labor migration but may apply to any type of internal, circular migration, a widespread phenomenon in developing countries.

Second, our results contribute to the child labor literature that focuses on crowding out of schooling (Edmonds 2009). This literature generally finds that in the wake of transitory income shocks, child labor declines and educational outcomes improve (Kruger 2007, Edmonds and Schady 2012). However, labor migration is almost never the source of these income shocks, so there is rarely a chance for the substitution effect of child for adult labor to operate. We provide new evidence that even when men do move away, replacement income in the form of remittances can still improve education outcomes.

Third, we open a new area of research in the economic history of *apartheid*. The vast majority of this literature focuses on the South African experience of *apartheid* and does not consider the impacts of the structures of *apartheid* on South Africa's many northern neighbors.[4] Given the extent of involvement of the Chamber of Mines in recruiting large numbers of men from Malawi, Zambia, Zimbabwe, Tanzania, Angola, Lesotho, Swaziland and Botswana over a number of years, it is perhaps surprising that there is such a gap in the literature. We assemble newly collected disaggregated Census data from the height of labor migration out of Malawi, and provide new evidence of the long-run impacts of mine migration to South Africa that may be relevant for these other countries, all of which have experienced large fluctuations in employment flows in different periods.

Finally, although we discuss Malawi's historical experience in this paper, our results have relevance for current migration policy. There are many countries with substantial outflows of foreign labor migrants and inflows of remittances. To understand the potential long run impacts of this migration on education, it is important to know more about the characteristics of local labor markets of sending communities and specifically whether children are good or poor substitutes for their male (or female) migrant family members. Furthermore, abrupt cessations of foreign guest worker programs may have unintended negative consequences for human capital outcomes in contexts where the income effects of adult labor migration outweigh the substitution effects on child labor.

Our paper begins by setting out some facts about education in Malawi and child labor during the pre- and post-colonial period. To motivate our empirical strategy, Section 3 describes in detail the background to mine migration from Malawi and local alternatives to migration. Section 4 sets out our empirical strategy and addresses potential threats to validity. Sections 5 and 6 present data and results and Section 7 concludes.

---

[4] Lucas (1985) and Lucas (1987) are exceptions, but these papers focus on national aggregate effects of the system of mine migration.

## 2. Human capital accumulation in Malawi: Background

Levels of human capital accumulation in Malawi are generally low, although they have increased over time. Prior to independence in 1964, missionaries were responsible for education and emphasized vocational training rather than literacy. In 1945, less than 6% of the population was literate. At independence in 1964, enrolment of school-aged children was under 35%. Between 1959 and 1978 however, total student enrolment 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.[5] By the early 1990s, primary school enrolment had grown to 50% of the relevant age range (UNICEF: http://www.childinfo.org/files/ESAR_Malawi.pdf). National trends will be important to take account of in our analysis.

Throughout our period of interest, cost was an important constraint to going to school. Unlike many African countries that instituted free primary education at independence, school fees for primary school education were in place in Malawi until 1994. Average annual school fees were the same across the country, around 2.75 Kwacha for lower primary and around 5.75 Kwacha for upper primary school.[6] Parents were also responsible for other expenditures such as textbooks, exercise books, writing materials and school uniforms (Heyneman 1980). The 50% increase in student enrolment following the enactment of universal free primary education in 1994 shows that school fees were a substantial impediment to enrolment during the 1950-1990 period (World Bank and UNICEF 2009).

Apart from the cost of education, the existence of outside options and obligations also played a role in the demand for education between 1950 and 1990. Children (ages 10 to 19) could work on family farms, in the household, or for a wage for other farmers or landlords, including owners of estates producing tea and tobacco for export. Rates of child labor in 1977 (the only year for which we can compute these rates) were indeed high. Around four in ten children aged 10 to 19 were working for pay for someone else, working without pay on family farms, or working in home production. There is also substantial variation in the share of 10 to 19 year olds employed across districts: some districts have only 20% of children working, while others report employment rates of 62%.

[5] There were just over 2,000 primary schools for over 600,000 enrollees in 1974, and only 18 high schools in the country (Malawi Ministry of Education 1977). Data on number of primary schools in Malawi in 1960 and in 1992 are taken from Malawi's Ministry of Education Annual Report (1960, p 37 Table 1) and Malawi's Statistical Yearbook (1995, p 57).

[6] Lower primary consisted of the first four years of primary school, upper primary of the next 4 years, 8 years in total (Heyneman 1980).

One of the main factors distinguishing between these high and low child labor districts is the presence of agricultural estates. In the next section, we discuss in more detail the evolution of these estates between 1950 and 1990. By the late 1970s, the share of 10 to 19 year olds working in estate districts is almost 50% while in non-estate districts, child labor is a lower 40% (results not shown). There are at least two reasons for these differences. The first is that the returns to child labor are higher on agricultural estates where children could expect to earn a higher wage relative to what children engaged in household or agricultural chores on the family plot could expect to earn, which was often nothing (Chirwa 1993). The second is that families who worked on these estates as tenants may have required their children to help, without pay, to increase the amount of output the family could produce, and in some cases, meet landlord specified quotas. We will use spatial variation in the presence of estates across the country to investigate whether and how the local technologies of agricultural production (estate versus non-estate production) contribute to the main impact of mining labor demand on educational outcomes.

## 3. Labor Migration from Malawi: Background and Context

Malawi is today one of the poorest and most densely populated countries in Africa and the world. Over half of the population lives in poverty, commensurate with its 170$^{th}$ (out of 186) ranking on the Human Development Index (United Nations Human Development Report, 2013). The agricultural sector currently employs around 50% of the labor force and 85% of the population lives in rural areas. Malawi's central location in southern Africa, high population densities, lack of natural resources, and few non-agricultural economic opportunities have all contributed to the importance of labor exports in this country's economic profile.

The bureaucratic infrastructure for labor exports was already established early in the colonial period and our analysis takes advantage of massive and largely unexpected fluctuations in these labor export flows and concomitant flows of remittances in the post-colonial period. To motivate our empirical strategy, the next sections describe the early establishment of the labor recruitment system in Malawi, the source of shocks to labor exports between 1950 and 1990, and the relative importance of earnings from mine work versus domestic earnings opportunities during the period.

*International labor migration from Malawi*

Throughout the twentieth century, Malawians took advantage of employment opportunities on Rhodesian farms and South African mines to boost local incomes. Over the first half of the twentieth century, labor recruitment to both destination countries became institutionalized. The South African mining industry established a physical presence in Malawi, opening and operating a series of recruiting stations by 1937.

These stations were run by the Witwatersrand Native Labour Agency (WNLA), the mining industry's centralized labor recruitment organization that coordinated all recruitment activities north of the Tropic of Capricorn (Crush et al. 1991, pg40; Jeeves 1987).[7] We collect and digitize information about the historical locations of these WNLA recruiting stations to use in our analysis.

The centralized and highly bureaucratic procedure of recruitment of foreign labor allowed the mines to maintain a uniform wage across the industry thereby eliminating labor competition within the industry.[8] Colonial Malawi (then Nyasaland) benefited from this system because WNLA agreements required mines to defer a large fraction of workers' pay and send it directly to Malawi, from where workers could withdraw their wage upon return. In this way, the government could be certain that incomes earned abroad were spent in Malawi.

Between 1950 and 1990, migrant labor from Malawi to South Africa's 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 1).[9] To put these numbers in perspective, at the peak of labor migration, around 20% of the adult male population was missing from the country and by 1977, 35% of all adult males had ever worked abroad. Most of these mineworkers were engaged on two-year contracts, after which they had to return home. Workers could reengage for subsequent two-year contracts after spending some time at home (Wilson 1972, pg 68, Prothero 1974 and Lucas 1985). Labor migration was therefore highly prevalent (we show below that almost all districts experienced some of this migration), circular, and reasonably long-term.

Until 1967, labor recruitment agreements between Nyasaland and WNLA restricted the number of workers WNLA could recruit to a few thousand workers per year. These restrictions were lobbied for by European plantation owners in Nyasaland as a way to protect access to a cheap source of labor (Paton 1995, pg 46) as well as by the Southern Rhodesian government which was concerned that Southern Rhodesian enterprises could not compete with the wages paid on South African mines (Jeeves 1987).[10]

---

[7] WNLA, or Wenela, merged with the Native Recruitment Corporation, NRC, in 1977 to form The Employment Bureau of Africa or TEBA. Many of the decisions about where and when to set up recruiting practices in southern Africa were spearheaded by the mining industry's "labor czar", William Gemmill (Jeeves 1987).

[8] Wilson (1972) and Lucas (1985) provide accounts of how the WNLA and the NRC operated as labor monopsonists in the colonies, keeping wages in the industry low by hiring workers from Mozambique, Northern and Southern Rhodesia, Nyasaland, Lesotho, Swaziland, Tanzania, Angola, Botswana as well as South Africa.

[9] Prior to the 1950s, migrant labor from Malawi to Rhodesia was also prevalent. However, the number of men leaving for Rhodesia never reached the numbers migrating to South Africa; the system of recruitment and remittance flows was never as centralized or organized as it was for WNLA; and by the early 1960s, most of the employment in Rhodesian agriculture had dried up.

[10] The Southern Rhodesian government forced WNLA to restrict the amount of workers WNLA would recruit from Nyasaland (Jeeves 1987). This restriction remained in place until Southern Rhodesia reduced its reliance on the Nyasaland labor supply (Clarke 1977).

Between 1946 and 1959, the WNLA quota increased from 8,500 men to 20,000 men, which represented a relatively small (around 2%) share of the population of working age males (Coleman 1973, Chirwa, 1992). In 1967, newly independent Malawi's President Banda signed a new agreement with WNLA removing all quotas on recruitment of Malawian workers (Treaty Series No. 10/1967). Figure 1 shows the sudden increase in the number of workers recruited by WNLA in the six years following the new agreement.

The 1967 labor agreement also permitted WNLA to withhold two thirds of miner wages until the miners' return to Malawi. This fact is important will be important for interpreting our results. Unlike many other migration contexts, we can be sure that at least two thirds of earnings were returned to sending communities. Moreover, these large amounts were not endogenously chosen by the miners. In 1967, the value of deferred pay and remittances paid by WNLA was K2.114m (Malawi Statistical Yearbook 1972, Table 7.9). This was Malawi's third highest source of foreign exchange revenue (Prothero, 1974). If we conservatively assume that voluntary remittances are zero, then the total amount of income that a family could have received from a miner returning from a two-year contract would have been 342.71 Kwacha. The size of the income shock within a district is then a function of total mine employment within the district multiplied by this 342.71 Kwacha.

The expansion in labor recruitment came to an abrupt halt in April 1974, when a WNLA plane transporting miners back to Malawi crashed, and 74 miners were killed.[11] In response, President Banda rescinded the labor agreement, banned all recruiting by WNLA and recalled all Malawian migrant workers home (Lucas 1985; Chirwa 1996).[12] Although Banda had wanted to expand the domestic agricultural estate sector, and had reason to recall the mine workers, the timing of this recall, initiated in response to the plane crash, was clearly unexpected.[13] In the four years between 1974 and 1977, mine employment fell dramatically (see Figure 1) from a high of over 120,000 men to zero. Although the flows of foreign capital in the form of mine wages dried up after the labor ban in 1974, the initiation of this ban entailed a lump sum payout of deferred pay earnings for all returning miners.

WNLA recruiters were allowed to restart operations in Malawi in 1977, but employment levels for Malawians never returned to 1970 levels as the South African mines had turned their strategy of

[11] *The Observer-Reporter*, April 5 1974 "77 Persons Killed in Plane Crash" http://news.google.com/newspapers?nid=2519&dat=19740405&id=q85dAAAAIBAJ&sjid=f14NAAAAIBAJ&pg=3041,1287591

[12] Banda claimed "I have killed *Wenela*" (Chirwa 1996)

[13] 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).

recruitment inwards, substituting local labor for what they saw as unreliable foreign supplies (Crush 1986; Crush et al 1991, pg 129).

The unanticipated rise in external labor demand from 1967 to 1973 and sharp fall from 1974 to 1977 constitute the two sides of our natural experiment. The periods before 1967 and after 1977 represent relatively stable periods of labor demand from the foreign mining sector. In the empirical methods section, we describe how the periods before and after the employment shocks constitute our comparison periods.

Our discussion of labor migration has so far ignored domestic economic opportunities for Malawians. To understand how mass migration (and then return) of adult male labor and the related flows of remittance income could have affected investments in education of children, we next describe the changing context of local economic opportunities in rural Malawi between 1950 and 1990.

*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, and 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), wages in the local economy were always far below what workers could earn on the mines.

Malawians remaining at home had essentially three, not necessarily mutually exclusive, options for work in the agricultural sector. Until the early 1960s, workers could participate in the *thangata* system. This entailed offering between one and three months' work per year to a European landlord while tending own crops on the landlord's land for the rest of the year (Christiansen and Kydd, 1982). After *thangata* practices were outlawed two years prior to independence, workers could instead work for wages on the large tea, tobacco and sugar estates, or as visiting tenants on these farms. Landlords allocated land to visiting tenants, and bought back the output at prices they determined. The expansion of the cash crop industry on these estates underpinned Malawi's 6% annual growth rate from the late 1960s to the late 1970s. The remainder of income earners were absorbed into the peasant smallholding sector (Christiansen and Kydd, 1982). Peasants grew cash crops for export, or food crops to sell to estate laborers and visiting tenants who did not have time or land to cultivate their own food.

Estate wage-workers typically earned more than visiting tenants and smallholders. Average annual earnings on local estates were 94.4 Kwacha in 1968 rising to 112 Kwacha in 1973 and to 126.80 Kwacha in 1974 (Pryor and Chipeta 1990).[14] In contrast, average miner earnings were at least twice as much as the average agricultural wage in the late 1960s. Miners could earn183 Kwacha per year in 1966 (Wilson 1972, pg 46) and this grew to 375 Kwacha in 1974 (Crush et al 1991, pg. 19). Mine earnings were therefore always significantly higher than they would have been relative to the next best alternative: working on an estate farm in Malawi. Despite the growth of the estate sector immediately after independence and the opportunities for wage employment and tenancy status, the low wages, the challenges of being a visiting tenant, and the lack of growth in the small holding sector led many men across all districts to continue to seek employment with WNLA.

*WNLA recruiting stations: Facilitating labor migration by reducing migration costs*

Although mine wages may well have been attractive to Malawian farmers, migration was still costly. By 1937 WNLA was well established throughout Nyasaland with recruiting stations set up in many districts across the country. We have collected and digitized administrative data on the location of these WNLA stations as of 1937 and show their prevalence across the country in Figure 2. The red hatched areas in the figure represent the communities (Traditional Authorities, or TAs) that had a WNLA station; the white areas show TAs without a WNLA station; and the thick black borders represent district boundaries. Each of the three geographical regions had some access to a WNLA station. However, the logistics of getting access to a mining job were neither simple nor costless.

The procedure for getting from Malawi to South Africa involved several steps.[15] A man 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 WNLA recruiting 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 going back to the WNLA station to await transportation to a main WNLA depot for transfer to South Africa (Prothero 1974). The costs associated with signing up for mine work were therefore not negligible. They also varied depending on distance to the nearest WNLA recruiting station.

An important part of our identification strategy rests on comparing outcomes across different cohorts from locations facing substantially different costs of signing up for mine work. These costs of signing up are proxied for by the presence of a local WNLA station. Table 1 provides evidence that the presence of a

[14] In 1974 K1 (Malawian Kwacha) was worth USD1.22
[15] 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).

WNLA station in a district and the number of WNLA stations in a district both predict substantially higher labor outmigration. The first four columns show results from a regression of the log of the number of men who report having ever worked abroad in the 1977 Census on different measures of access to WNLA stations, region fixed effects, and historical population density. Both measures of exposure to recruitment predict significantly higher stocks of labor migrants by 1977. The second set of columns show the same results for a flow measure of migration: the log of the number of men issued foreign travel certificates between 1958 and 1967 at the district level. Again, proximity to a WNLA station positively and significantly predicts these outcomes. In the next section, we explain how we use this spatial variation in access to mine work to distinguish between districts with high versus low exposure to the employment shocks of 1967 to 1973 and 1974 to 1977.

**4. Empirical Strategy**

To understand how migration affects the educational attainment of those left behind, we test for the sign of the net effect of Malawian labor migration to South Africa between 1967 and 1977. We test between the following two hypotheses:

*H1: The effect of income dominates the substitution effect:* When men leave to work abroad and send remittances, school enrolment and attainment rise, as long as goods are normal. When men return (so the substitution effect reverses sign), bringing with them accumulated remittances, enrolment and attainment may continue at high rates until the income runs out. The net effect of exposure to international labor migration on long run education attainment among children left behind will be positive in both the 1967 to 1973 cohorts and the 1974 to 1977 cohorts, relative to older and younger cohorts. Effects for the 1974 to 1977 cohorts could be even larger, since they experience the same signed income and substitution effects related to these migrant flows.

*H2: The substitution effect dominates the effect of income*: When men leave to work abroad, schooling falls, as children replace fathers at work. When male labor returns, the demand for child labor declines, and enrolment and attainment rise at least among those young enough to return to school. The net effect is that education outcomes should decline for cohorts eligible for primary school in 1967-1973 and then improve for cohorts eligible for primary school in 1974-1977, relative to both older and younger cohorts.

To test between these two hypotheses, we estimate the gap in total schooling between adults who were age eligible for primary school between 1967 and 1977 and from areas that had a WNLA recruiting station (WNLA districts) and similarly aged adults from areas with no WNLA stations (non-WNLA districts). Since districts with and without WNLA recruiting stations might differ on a number of dimensions (e.g. WNLA districts could be wealthier, or have more schools, or more land for cultivation), we control for counterfactual difference in education outcomes across cohorts ineligible for primary school in these districts.

The main difference-in-differences regression model we estimate is:

$$\bar{Y}_{asd} = \beta_0 + \beta_1 Eligible67\text{-}73_{asd} * WNLA_d + \beta_2 Eligible74\text{-}77_{asd} * WNLA_d + \beta_3 Eligible67\text{-}73_{asd} +$$

$$\beta_4 Eligible74\text{-}77_{asd} + \beta_5 WNLA_d + G_d{'}\pi + X_{as}{'}\gamma + \lambda_d + \varepsilon_{asd} \qquad (1)$$

where $\bar{Y}_{asd}$ is total years of schooling attained or share of adults with any primary school by age-sex-district cell.[16] $G_d$ is log of historical population density (measured in the 1931 Census), $X_{as}$ is a set of demographic controls (age and sex controls) and $\lambda_d$ is a district fixed effect. In our most comprehensive specification, we also include region-specific trend terms in (1) to control for differences in trends in education across Northern, Central and Southern regions in Malawi.[17] Robust standard errors are clustered at the district level.

*WNLA* is either a count variable capturing the number of WNLA stations in the district or an indicator for whether the community contained a mine recruiting station established by 1937. $\beta_5$ therefore gives us an estimate of the schooling gap between WNLA areas and non-WNLA areas. We might expect this gap to be positive if recruiting stations are located in relatively wealthier areas (perhaps to target healthier workers), or negative if they are located in areas with high demand for child labor from other sectors.

*Eligible67-73* and *Eligible74-77* are binary variables denoting whether an adult was age eligible for primary school during the expansion of mine employment in 1967 to 1973 or during the labor ban years of 1974 to 1977. Since the decision to start and continue with the first four years of primary school is most relevant for the ages 6 to 12 years (the years in which one should be attending primary school), we define the age eligible cohorts in the district-level data as being between the ages of 5 and 14 in a given year. To construct appropriate comparison cohorts (i.e. those not eligible for primary schooling in either of the two periods) we consider cohorts age eligible for primary school in the post-1977 period (those

[16] In the absence of technical change in agricultural production during this time (Foster and Rosenzweig 1996), there would have been little reason to get much education beyond what was required for functional literacy (about 4 years of education). Positive returns to education in agriculture typically occur around this level of education (Appleton 2000). For this reason, we focus on measuring the impacts on total years of schooling attained and on the share of adults with any primary school enrolment, rather than secondary schooling or higher education; these are irrelevant outcomes in rural Malawi during this period.

[17] The Northern Region was the birthplace of formal schooling in Malawi (Heyneman 1980), originally established by missionaries. It is plausible that these different initial conditions with respect to education could influence subsequent trends in education across regions.

younger than 5 in 1977) and those eligible in the pre-1967 period (older than 14 in 1967). Both younger and older cohorts are eligible for primary school during more stable periods of mining employment (see Figure 1).[18]

For much of our analysis, we combine the older and younger cohorts into a single comparison cohort. $\beta_3$ and $\beta_4$ in equation (1) therefore describe differences in mean schooling attainment between age eligible and age ineligible cohorts in non-WNLA areas. These differences might be positive or negative, because although educational attainment is increasing over time, the non-eligible cohorts are comprised of both older and younger individuals.

The main parameters of interest in equation (2) are those associated with the difference-in-differences terms, *Eligible67-73*WNLA* and *Eligible73-77*WNLA*. These interaction terms allow us to estimate the differential difference in education gaps between age eligible and age ineligible cohorts in WNLA versus non-WNLA areas. As explained above, the signs of these interaction terms are ambiguous and driven by a combination of income and substitution effects. If the substitution effect outweighs the effect of income, we expect $\beta_1<0$ and $\beta_2>0$; while if the income effect dominates, $\beta_1>0$ and $\beta_2>0$.

Identification of these parameters relies on district-level variation in access to mine jobs proxied by the presence of an historical WNLA recruiting station combined with within-district level variation in age eligibility of different cohorts. It is important to note that because mine migration from Malawi is so prevalent, no districts are completely untouched by the labor expansion and contraction. However, as Table 1 shows, districts with at least one WNLA station have *more* men migrating abroad, so we are able to estimate the differential impact of additional migration (and return flows of men, and money) across high and low shock intensity districts. As with any difference-in-differences research design, to interpret the interaction terms as capturing causal effects of having more (or fewer) migrants, we must assume no differential trends in education across WNLA and non-WNLA districts. That is, the education gaps between ineligible cohorts in WNLA and non-WNLA areas should provide a valid counterfactual for the gap between eligible cohorts in WNLA and non-WNLA areas. This is where it is useful to have both younger and older comparison cohorts. We will be able to show parallel education trends among these comparison cohorts across areas with and without WNLA stations.

Although the locations of WNLA stations set up long before the period of analysis provide us with a good pre-existing measure of differential costs of getting access to mining jobs in South Africa, these stations were not randomly allocated to districts even in 1937. To get some insight into factors driving selective

[18] Appendix Table 1 shows in detail how we construct these cohorts from Census data and their respective ages in 1967, 1974, 1977 and in 1998.

placement, Table 2 presents correlations between the location of WNLA stations at the district and sub-district (Traditional Authority) level and a range of historical and geographic variables.[19] We use the number of WNLA stations in the district as the outcome in the first four columns and an indicator for WNLA station in the sub-district in the next four columns. Interestingly, log population density measured in 1931 is negatively correlated with WNLA stations at the district and sub-district level (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). Certainly, we see some evidence of this in the negative coefficient on the estate dummy: in districts that have large tea or tobacco plantations later in the period, the chances of having a WNLA station in 1937 are lower (although never significantly so). Districts at higher altitude (i.e. more likely to be free of malaria) and districts in the Central region are more likely to have a WNLA station, although these relationships are not always significant.

With this selective placement in mind, equation (1) has several controls that allow us to limit concerns about short, sharp, confounding shocks to the local economy of WNLA or non-WNLA areas during the 1967-1977 period that could complicate our results. By including district level controls for historical population density in 1931, region fixed effects, and eventually district fixed effects and region-specific trends, we compare schooling gaps between cohorts within the *same* district. Since geography, agricultural conditions, weather and population density vary much more down the length of the country and across regions and districts, we eliminate many potentially confounding factors with region and district level fixed effects. For robustness, we also estimate (1) using sub-district level variation in the 1998 Census, the lowest geographic unit for which we have information about WNLA stations. Although this means we have to give up our older control group to do this (the 1977 Census does not contain sub-district level data on education), we can use a richer set of controls and exploit finer variation in the WNLA measure than at the district level.

Regardless of the level of analysis (district or sub-district), our identification strategy relies on assumptions of no differential changes in school supply across WNLA districts, and no differential local shocks to agricultural production that might change household incomes and hence the demand for primary schooling among age eligible cohorts. Since there was no massive school expansion program established

[19] Note that at the time that WNLA was deciding where to locate these WNLA stations, there was no hard data on which to make location placement decisions. Decisions were likely made after visual inspection of the potential of these areas for recruiting. Jeeves (1987) describes this process: "As early as 1928, the Chamber (of Mines) began to lay its plans for expansion into the north. In that year, WNLA sent one of its senior employees on tour into the trans-Zambesi. Travelling by auto, rail and ferry, P. Neergaard saw huge areas of Nyasaland, Northern Rhodesia, Tanganyika and northern Mozambique. He returned with glowing reports on the labour potential of these areas and the ease with which WNLA could establish itself in them."

in the post-independence period (see Section 2), it is unlikely that changes in schooling access could confound the interpretation of our results. Dealing with changes in agricultural production is a little trickier. As noted in Section 2, Malawi's agricultural estate sector was growing during the years that mining employment was expanding and then contracting and this growth was unevenly distributed across the country.

How would this differential local economic growth affect interpretation of results in (1)? Since WNLA stations tend to be located in areas with fewer estates (see Table 2), it is likely that the positive effects on Malawi's agricultural-based economic boom from the mid-1960s to the late 1970s had larger impacts on local work opportunities for men in non-WNLA areas. Such differential employment growth concentrated in non-WNLA areas between 1967 and 1977 should lead us to underestimate the education effects of international migration. Since men do not leave Malawi to take up these estate jobs in non-WNLA areas, we would expect an income effect from more estate work, and no substitution effect pulling children out of school.[20] Controlling for region and district fixed effects, and region-specific trends help us address some concerns about the growth of the domestic estate sector, since tobacco plantations that were the major source of this growth tended to be located in the Central Region, and in certain districts.

In addition to controlling for potential threats to validity from the effects of local estate sector growth, we pursue another strategy that uses the spatial variation in estate prevalence to investigate a particular mechanism for our education results. That is, we can use what we know about the growth on tea and tobacco plantations to understand more about how local technologies of production contribute to the long run effects of labor migration on human capital investments. Specifically, we estimate separate education regressions for two samples of districts: districts that have large tea and tobacco estates, and districts without.

As discussed in Section 2, we know that child labor was used on agricultural plantations in Malawi. Certainly, among tenant farmers trying to meet annual output quotas set by landlords, missing male labor might have forced the household into marshalling all available labor resources to meet these quotas. If the removal of male labor through international migration does create pressure to substitute towards child labor at all, we should see smaller net effects of access to South African mine work on the long-run educational attainment of adults from estate districts, compared with adults in non-estate districts.

[20] Or, even if men are vacating own farm jobs for work on agricultural estates, thereby inducing children to replace them on their own fields, we would expect the substitution effect to be much smaller when these agricultural jobs are still in home districts.

A final and important threat to validity of our results arises because of the measurement error in $WNLA_d$ induced by internal migration within Malawi between childhood and adulthood. Our data do not contain information on birth district at the individual level, which would allow us to define $WNLA_d$ more accurately. We cannot know whether a person's current district of residence is the district in which she went to school. We do have some information on the district, age and sex-specific prevalence of internal migration between 1967 and 1977, and we use this information to create bounds for our education estimates in (1). We discuss details of how we adjust for internal migration in the results section.

**5. Data and descriptive statistics**

Our data are from the complete 1977 and 1998 population Census' of Malawi. The 1998 Census is comprised of individual level data for all three regions (North, Central and Southern Malawi), 27 districts and 227 sub-districts (Traditional Authorities, TA) of Malawi. A TA contains on average 44,000 people, while districts are larger, consisting of many villages and with an average population size of 370,000 people. The 1977 Census data are aggregated to the district, five-year age group, sex level. We digitized all of the Census 1977 aggregate reports and matched these data to appropriately aggregated 1998 Census data.

To estimate equation (1), we restrict the sample to adults ages 20 to 49 (in five-year age groups) in each Census wave and exclude adults living in any of the four towns of Malawi (Blantyre, Lilongwe, Zomba, and Mzuzu) since these locations would likely have offered additional economic opportunities outside of mine work and agriculture.[21] In a robustness check, we use only the 1998 Census data aggregated by age and sex to the sub-district level, and further restrict the sample to only those TAs with at least 100 individuals of each age and either sex (this leaves us with 210 TAs).

Our definition of whether a cohort was exposed to the labor expansion or contraction depends on whether the cohort was age eligible for the first four years of primary school during the 1967 to 1977 period. Appendix Table 1 describes how we map age groups and ages from the 1977 and 1998 Census waves into cohorts exposed to employment shocks between 1967 and 1977, or to either of the two unexposed cohorts. The younger comparison cohort comprises 20 to 26 year olds from the 1998 Census; these adults were eligible for primary school after 1977. The 27 to 36 year olds in 1998 would have been between ages 6 and 12 at some point during the 1974-1977 labor ban, while the 37 to 44 year olds would have been 6 to 12 in the 1967 to 1973 period. Note that in each of the eligible age groups, some individuals would have been exposed to the shock in every year between age 6 and 12, while others would have only been exposed in the first or last part of their 6 to 12 age range.

[21] This exclusion affects only XX% of the adult sample.

We do not use the 1998 Census to construct an older comparison cohort, eligible for primary schooling before 1967. This is because life expectancy in Malawi was only 46 years in the late 1990s (http://www.theglobaleconomy.com/Malawi/Life_expectancy/), and we are concerned about mortality selection at these upper ages. Instead, we construct synthetic older cohorts using adults aged 20 to 44 in the 1977 Census. Within this sample, adults aged 20 to 24 in 1977 would have been ages 10 to 14 and hence eligible for primary school during the 1967 to 1973 labor expansion; those older than 24 in 1977 were all 15 and older by 1967 and so unlikely to have had their educational attainment affected by the mining employment shocks.[22]

For all analyses, we use self-reported age in years to determine whether an individual was part of the exposed or non-exposed cohorts. Appendix Figure 1 presents histograms of reported ages by WNLA station status of the area from the 1998 Census. While there are some spikes on the decadal ages (20, 30 and 40), there are no other obvious age heaps on ages ending in 5. Also, since we are not creating an eligibility variable using a specific year of exposure, any age misclassification at the beginning or end sections of the eligible cohorts (defined in five year age groups) is likely to bias us against finding any significant differences between eligible and ineligible groups, within or across WNLA and non-WNLA areas. It simply means we have a fuzzier definition of treatment than would be ideal; it is likely this makes it harder to find any significant effects.

To construct proxy measures of the costs of getting access to mine work in South Africa, we collect data from administrative records on the location of WNLA recruiting stations prior to 1937. In our district level analysis, we use the number of WNLA stations in the district (results are similar if we use an indicator for at least one WNLA station in the district or the share of sub-districts in the district with any WNLA station). For the sub-district level analysis, we use an indicator for whether the TA has any WNLA station or not (shown in Figure 2). Importantly, there are recruiting stations across the length of the country, so we can make comparisons across exposed and non-exposed cohorts within regions and within districts. This allows us to create better counterfactuals by controlling for differences in outcomes that might arise because economic alternatives to mining differ across large districts and larger regions

Table 3 presents summary statistics for outcome and control variables at the five-year age group-sex-district level in Panel A and for some baseline historical and geographic variables at the district level in Panel B. Much of the data for Panel B comes from newly digitized historical Census data (from Census

[22] Because the 1977 educational attainment data are only reported in five-year age categories by sex and district, we create a slightly fuzzier measure of *Eligible67-73* that is one for the age group 20 to 24 in 1977, and otherwise zero.

1931, 1945, 1966 and 1977). We present means for the full sample and for districts with and without a WNLA recruiting station along with the *p* value of the difference in means.

Panel A shows that about 60% of districts have at least one WNLA station and the average number of stations per district is two. The fraction of cohorts eligible for primary school in either period, and in each of the younger and older comparison cohorts is balanced across WNLA and non-WNLA areas.

The middle part of Panel A gives an indication of the prevalence of child labor across areas, and the dismally low levels of education among adults in our sample. Almost 40% of children aged 10 to 19 are enrolled in school in 1977, while 42% report that they are working. This employment is comprised of wage work, work as a farmer (*mlimi*) for no pay, and working at home. 13% of this child labor is in home production, and this fraction is balanced across WNLA and non-WNLA areas. In contrast, the share of 10 to 19 year olds working for a wage or as *mlimi* is 25% in WNLA areas, and a significantly higher 34% in non-WNLA areas. Recall that these are outcomes measured in 1977, at the end of the labor ban. We use these child labor variables in some of our analysis to illustrate mechanisms.

Average schooling among adults aged 20 to 44 (across eligible and ineligible cohorts) is 2.56 years. Much of this average is driven by the older cohorts, since 40% of the sample falls within the older comparison cohort group. Average education is 2.85 years in WNLA areas and only 2.08 years in non-WNLA areas; the share of adults who have ever been to primary school is 45% in WNLA areas and only 35% in non-WNLA areas. Both of these differences are strongly significantly different than zero.

In Panel B of Table 3, we see some differences between WNLA and non-WNLA areas. As we saw from Table 1, in WNLA areas, a higher fraction of men report ever working aboard by 1977 (37%); although note that the share of men who have ever worked abroad is still high (31%) within non-WNLA districts. Population density in 1931 is significantly lower in WNLA districts and literacy rates in 1945 are twice as high as in non-WNLA districts, although still below 10%. Interestingly, the fraction of districts with a tea or tobacco estate is higher among WNLA districts than among non-WNLA districts, although not significantly so.

The last part of this table gives us some idea of how local economic opportunities differed across areas. Using variables constructed from the 1966 Census, we see that men are more likely to be working for a wage in 1966 in non-WNLA areas, and are more likely to report not earning any wage in 1966 in the WNLA areas (Note that not earning a wage does not imply unemployment: men may be self-employed on their own peasant farms). Rates of wage work in farming are similar in both areas, with men more likely to work in non-farm sectors for a wage in the non-WNLA areas. In the face of such differences in economic opportunities at the district level, it will be particularly important to control for district fixed effects in our analysis.

## 6. Results

*i. Long run education effects of exposure to mine employment shocks: District-level analysis*

Table 4 presents the raw difference-in-differences estimates for long run effects of exposure to international labor migration on the next generations' human capital. We show means, standard deviations and cell counts at the district-five year age group-sex level for two measures of education: total years of schooling attained by age 20 and the share of adults with any primary school.[23] Statistics are provided for districts with more and less exposure to the mine employment shocks (WNLA and non-WNLA districts). Within each set of districts, we show means for cohorts eligible for primary during the labor expansion period (row 1), for cohorts eligible during the labor ban (row 2), and for the combined comparison group of cohorts who are too old or too young for primary school between 1967 and 1977 (row 3). Columns (3) and (6) of the table show differences in means across WNLA and non-WNLA districts and the lower part of the table shows difference-in-difference comparisons without additional controls.

First, looking at cohort-specific differences across WNLA and non-WNLA districts, we see that educational attainment is uniformly higher in districts that had easier access to mine work, and these differences are statistically significantly different from zero. Controlling for education gaps between WNLA and non-WNLA districts among the younger and older comparison cohorts (row 3), we see that cohorts age-eligible for primary school during the labor expansion period and from a WNLA district gained about 0.58 more years of education. An even larger education gain – 0.94 years – is evident in the difference-in-differences comparison for the cohorts who were age eligible during the labor ban. And, both sets of exposed cohorts also exhibit gains in primary school enrollment. WNLA cohorts eligible for school between 1967 and 1977 were between 7 and 11 percentage points more likely to have ever attended primary school relative to non-WNLA cohorts, controlling for differences in enrollment rates across WNLA and non-WNLA areas in the combined younger and older comparison group.

These education effects are economically large and statistically significant. Relative to mean education levels and rates of primary school attendance in the sample, exposed cohorts in WNLA areas attain 23 to 37% more education and are 17 to 27% more likely to have ever been to primary school. On average, these education gains bring the exposed cohorts closer to the level of education required for functional literacy (about 4 years). And, the fact that estimates are uniformly larger for cohorts eligible for schooling

[23] Results using the 1998 Census data at the sub-district level are available in Appendix Table 2.

during the labor ban makes sense, since income flows from migrant earnings in this later period were still large and for some workers, the lump sum payout upon recall to Malawi would have been even larger than remittance flows in the earlier labor expansion period.

Table 5 presents regression-adjusted difference-in-differences results for the education outcomes, controlling for gender, age group dummies, region fixed effects and the log of historical district population in columns (2)-(4) and (6)-(8), adding district fixed effects in columns (3), (4), (7) and (8), and finally region-specific trends in columns (4) and (8). Note that here, instead of using a WNLA dummy variable, we control for the number of WNLA stations in the district (as a measure of access intensity) and the interaction of this number of stations with the exposed cohort dummies. We present coefficient estimates evaluated at the mean level of WNLA stations per district (two stations).

Results are very similar to those in Table 4, although effect sizes are somewhat smaller, since we are controlling for a number of variables that are important for explaining variation in educational attainment; in particular, district fixed effects and region-specific trends. Total schooling and share of adults with any primary school are both higher in districts with more WNLA districts. The difference-in-differences estimates for $\beta_1$ and $\beta_2$ are still positive and statistically significantly different from zero once all controls are included. Including district fixed effects in columns (3) and (7) is particularly important in allowing us to precisely estimate the impact of the mine employment shocks on outcomes for different cohorts within the same district. Region-specific trends account for positive trends in education (particularly post-independence) that might differ by region.

To discuss magnitudes, we focus on the results from models including the most comprehensive set of controls (i.e. columns (4) and (8)). Relative to districts with no WNLA stations, in districts with the mean number of WNLA stations, cohorts eligible for school during the expansion and contraction of mining employment gained an additional 0.25 to 0.37 years of education and were between 2.2 and 3.4 percentage points more likely to have ever been in primary school. These estimates represent a roughly 10 to 15% gain in total years of education and a 5 to 8% gain in the share with any primary school.

Instead of using only two exposed cohort dummies to measure education gaps for these two cohorts across WNLA and non-WNLA areas, relative to a combined older and younger comparison cohort, we can use the full range of age groups across the 1977 and 1998 Census samples to illustrate the long run human capital effects of exposure to international employment shocks. Figures 3 and 4 show how differential exposure to these mining employment shocks enabled cohorts in WNLA districts to overtake exposed cohorts in non-WNLA districts, permanently altering the long run level of human capital in WNLA districts.

Figure 3 presents coefficients from a regression of total years of education on nine age dummies (for each of the five-year age groups between ages 20 and 64) and the interaction of these dummies with a WNLA district indicator, controlling for gender, district and region fixed effects. The 1977 Census contributes data for age groups 40 to 64, while the 1998 Census contributes data for age groups 20 to 44 (that is, the age group 40-44 is the only age group that contains data from both Census years). Vertical lines demarcate cohorts too old for primary school by 1967 (the older comparison cohorts), cohorts eligible for primary school during the labor expansion and labor contraction, and cohorts too young for primary school before 1977 (the younger comparison cohorts). The youngest cohorts in 1977 (ages 20 to 24) and cohorts older than 25 (ages 25 to 44) in 1998 comprise the cohorts eligible for primary school during the mining labor expansion or during the labor ban.[24] The black line in Figure 3 represents regression-adjusted education levels among cohorts from WNLA districts; the dotted line represents regression-adjusted education levels among cohorts in non-WNLA districts.

Several important features of the figure stand out. First, the trends in education appear to be parallel for the pre-1967 comparison cohorts and for the post-1977 cohorts. Second, education is increasing over time in both WNLA and non-WNLA areas, with large increase beginning in the cohort eligible for primary school in the 1967 to 1974 period.[25] To understand the increase in education from just over one year to over four years in non-WNLA areas, we need to recall that even non-WNLA districts were likely affected by the expansion and contraction in foreign employment – just to a lesser extent than WNLA districts. Our identification strategy relies on using WNLA stations to isolate *larger* migration flows in some districts relative to others. The figure clearly shows how this higher level of exposure to the same employment shocks enabled WNLA districts to overtake non-WNLA districts in the accumulation of human capital. While WNLA areas start with *lower* average education than non-WNLA areas, these areas start to overtake average education levels in cohorts 40-44 years and 35-39 years who are eligible for primary school just as employment on the mines begins to expand. Education levels in WNLA areas remains above education levels in non-WNLA areas for the period of labor contraction (cohorts 30-34 years and 25-29 years).

To better illustrate these cohort-specific changes in this education gap between WNLA and non-WNLA areas, we plot the interaction term alone by five-year age cohort in Figure 4, along with confidence intervals for these differences. The black line clear shows education levels in the WNLA areas overtaking education levels in the non-WNLA areas starting with the 40-44 year old cohort, and continuing until the

[24] Recall we cannot perform an exact mapping from age to cohort eligibility as in the sub-district level analysis, since the 1977 Census data reports only five-year age groups. We make the 1998 data comparable by collapsing data to five-year age groups, which makes the assignment to eligibility cohorts is fuzzier than in the previous analysis.
[25] The figure does not control for trend – a trend term would be collinear with age dummies.

youngest cohort (age 20-24), when the differences are no longer statistically significant. Conditional on our identification assumptions, we interpret the results in Tables 4 and 5, along with results in Figures 3 and 4, as showing that in the space of only ten years, exposure to mining employment shocks and concomitant migrant remittances enabled WNLA districts to overtake non-WNLA districts in their total amount of human capital, with long-lasting effect.

### *ii. Robustness: Sub-district level analysis*

Although our preferred estimates use both older and younger comparison cohorts and hence the district-level data, we can still use sub-district level variation in access to any WNLA recruiting station in the 1998 Census to check the robustness of our main results using only the younger comparison cohorts as controls. An added benefit of this approach is that we can use finer measures of age eligibility for primary school, since the 1998 Census data are captured for individual ages rather than in aggregated five-year age groups.

Table 6 presents results from this sub-district analysis. In these regressions, we use an indicator for the presence of any WNLA station in the Traditional Authority to capture lower costs of accessing a mining job (rather than the number of stations in the district – no TA has more than one recruiting station). Since the 1998 Census contains additional education outcomes, we also provide results for the share of adults stating that they are bilingual (English and Chichewa) as a measure of actual skills learned in school. The sample is restricted to adults ages 20-44 in 1998, we use individual ages to classify individuals as age-eligible for primary schooling during different periods, and regressions add in controls for female, age, age squared, the log of historical district population density, region and district fixed effects and region-specific trends.

Across all education outcomes, the difference-in-differences results are again positive, economically large and statistically significant. Since estimates are reasonably robust across specifications, we focus on results from regressions with the full set of controls. Relative to younger comparison groups, eligible cohorts in WNLA sub-districts enjoy between 0.157 and 0.26 more years of schooling than eligible cohorts in non-WNLA sub-districts, are 1.5 to 2.1 percentage points more likely to have ever been to primary school and are 1.7 to 2.7 percentage points more likely to be bilingual. Relative to mean values of outcomes in this 1998 sample, these point estimates represent between 3.8 to 6.4 percent more education, and between 2.3 to 3.2 percent higher enrollment in primary school.

Point estimates are smaller than those estimated in Table 5 since we are using a finer level of variation in exposure to mine employment shocks (within-district variation in access to a WNLA station) and a finer measure of eligibility for schooling during the expansion and contraction years (age, not five-year age groups). We also use only the younger comparison cohorts as controls. Since Figures 3 and 4 showed us how cohorts exposed to mine employment shocks were able to catch up and overtake education levels of non-exposed older cohorts, it is not surprising that controlling for (now reversed) education gaps between younger cohorts in WNLA and non-WNLA areas results in smaller point estimates on our difference-in-differences interaction terms. We next discuss the most important potential threat to validity of our results: Internal migration.

*iii. Sensitivity of results to adjustments for composition effects of internal migration*

As noted in Section 3, one drawback of our data is that neither the 1977 nor 1998 Census indicate district of birth for adults with different levels of education. This implies an error-prone measure of childhood exposure to WNLA recruiting stations for individuals who move across districts after completing their education, but before we see them in the relevant Census year. For an adult in our sample living in an WNLA district during a Census year, their district of birth could have been a non-WNLA district (or vice versa). 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 WNLA stations generates unpredictable biases our estimates.

For example, suppose all districts had the same average level of education before internal migration. Then, suppose that adults with more education move from non-WNLA to WNLA districts while less educated adults move from WNLA to non-WNLA districts. Such internal migration would artificially generate average (positive) differences in adult educational attainment across these districts that we would ascribe to exposure to WNLA stations. Of course, the WNLA dummy in our regressions controls for constant differences in internal migration rates. But, internal migration flows that differ by district as well as by age group could threaten the validity of our results by generating 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 values 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 (we call this the net migration rate, or $NetMigRate_{asd}$), for each district in each five-year age group and gender cell. In our data, this number is always between -0.35 and 0.29.[26] We need to assume that this net migration rate

[26] The net migration rate is the number of net migrants (in-migrants – out-migrants) per person living in the district, and is computed from the 1977 Census as the difference between total in-migrants and total out-migrants divided by

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 the same level of education and therefore we only need to account for the potential education of net migrants, the difference between in- and outmigrants.[27]

We adjust our education variables ($\bar{Y}_{asd}$) measured at district, age and sex level in the following way:

$$\bar{Y}_{asd}^{BOUND,upper\ or\ lower} = \frac{N_{asd}\bar{Y}_{asd} - NetMigrants_{asd} * \bar{Y}_{as}^{m}}{N_{asd} - NetMigrants_{asd}} \qquad (2)$$

$$= \frac{N_{asd}\bar{Y}_{asd} - NetMigRate_{asd} * N_{asd} * \bar{Y}_{as}^{m}}{N_{asd} - NetMigRate_{asd} * N_{asd}}$$

$$= \frac{\bar{Y}_{asd} - NetMigRate_{asd} * \bar{Y}_{as}^{m}}{1 - NetMigRate_{asd}}$$

where $\bar{Y}_{asd}^{BOUND,upper\ or\ lower}$ represents the adjusted mean education outcome at district, age and sex level, $N_{asd}$ is total population in a district-age-sex cell, $\bar{Y}_{as}^{m}$ is either the maximum or minimum value of the relevant education variable across all districts at age and sex level, $\bar{Y}_{asd}^{BOUND,upper}$ is the bound when we impute minimum education for net-migrants and $\bar{Y}_{asd}^{BOUND,lower}$ is the bound when we impute maximum education for net-migrants, and $NetMigrants_{asd}$ is the total number of net migrants in a district-age-sex cell. $NetMigrants_{asd}$ is estimated by multiplying the total population in that district-age-sex cell with the net migration rate ($NetMigRate_{asd}$) for that cell. Each component of (2) comes from the relevant Census wave, except for $NetMigRate_{asd}$ which is computed using 1977 Census data and then applied to both Census waves. We then estimate the main regression specifications for the combined 1977/1998 Census sample using these adjusted education variables, one set for each of the extreme values of $\bar{Y}_{as}^{m}$.

There are two notable features of equation (2). First, the adjustments we make for internal migration imply that $\bar{Y}_{asd}^{BOUND,upper}$ and $\bar{Y}_{asd}^{BOUND,lower}$ provide upper and lower bounds on mean education and average share of adults with any primary school across the entire sample. Second, these adjustments do not imply that the difference-in-differences regressions using these new variables will produce estimates that contain the main estimates in Table 5. 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

---

total current population in the district. Census 1977 counts the number of people in each age, district and sex cell and enumerates how many of these individuals were born in each district. A 0.2 net migration rate means that for every person living in the district, there are 0.2 net in-migrants.

[27] 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. We are experimenting with ways to relax this assumption.

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

$$NetMigRate_{asd} * \left(\bar{Y}_{asd} - \bar{Y}_{as}^{min}\right) > 0 \qquad (3)$$

$$NetMigRate_{asd} * (\bar{Y}_{asd} - \bar{Y}_{as}^{max}) < 0 \qquad (4)$$

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, $\bar{Y}_{asd}^{BOUND,upper} > \bar{Y}_{asd}$ only in receiving districts. This makes sense: in receiving districts, our adjustments take out the low levels of education of net in-migrants to create a higher adjusted mean education variable. Similarly, when we impute the maximum level of education for net migrants, equation (4) will only be satisfied in receiving districts; subtracting high levels of net in-migrant education generates $\bar{Y}_{asd}^{BOUND,lower} < \bar{Y}_{asd}$. For sending districts where $NetMigRate_{asd} < 0$, the inequalities in (3) and (4) are reversed. For such districts, it is possible that $\bar{Y}_{asd}^{BOUND,lower} > \bar{Y}_{asd}$ and $\bar{Y}_{asd}^{BOUND,upper} < \bar{Y}_{asd}$.

Because we have both sending *and* receiving districts in our sample, and because rates of internal migration are different across WNLA and non-WNLA districts (rates of in-migration are higher in WNLA districts, results not shown), our adjustments have different effects on the bounds values in specific WNLA and non-WNLA districts. Even more complicated patterns of net migration that vary across exposed and non-exposed cohorts, as well as across WNLA and non-WNLA areas, imply that any adjustments for internal migration may generate in difference-in-differences estimates that do not bound our main result.[28] Therefore, to argue that selective internal migration does not confound our results, we would like to see similar, positive estimates of the long run education effects of mine employment shocks after making these adjustments for compositional changes.

Table 7 displays results from difference-in-differences regressions estimated using the new adjusted education variables, first including all controls and district fixed effects, and then adding in a region-specific trend term for each outcome. We compare the main coefficients in Table 7 with the coefficients in Table 5. Overwhelmingly, the positive human capital effects of exposure to the employment expansion and contraction in WNLA districts are still evident. All of our estimates are statistically different from zero at the 5 or 10% level.

[28] Crudely, if net migration rates are more likely to be positive in WNLA 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.

If we assume net migrants have the minimum level of schooling in the age-sex cell for a given Census year, the presence of anyone new in a receiving district "drags the average down" and their absence from a birth district artificially inflates that district's average education. Adjusting for these uneducated net migrants, we still see large, positive impacts of exposure to treatment among exposed cohorts: those exposed during the labor expansion have 0.32 more years of education, while those exposed during the labor contraction have 0.52 more years of education. If we instead assume that net migrants are educated, removing them from our outcome measure in receiving districts and adding them back to sending districts reveals similar, large positive impact of exposure to mine employment shocks. The difference-in-differences estimates in columns (3) and (4) imply that exposed cohorts in WNLA areas gained between 0.25 and 0.39 more years of education. These estimates compare favorably to our main results in Table 5: between 0.25 and 0.37 more years of education (Table 5, column 4).

Results are similarly robust for the share with any primary school variable. In Table 5, exposed cohorts from districts with more WNLA stations are 2.2 to 3.4 percentage points more likely to have ever attended primary school. After adjusting for internal migration in Table 7 columns (5)-(8), these exposed cohorts from districts with WNLA stations are between 2.4 and 4.4 percentage points more likely to have ever been to primary school.

Substantial internal migration could have led to a spurious relationship between measures of exposure to mine employment shocks in childhood and education outcomes measured in adulthood. However, even after accounting for the two extreme types of composition effects driven by internal migration, we still find large, positive and statistically significant coefficients on the difference-in-differences interaction terms.

*iv. Discussion of magnitudes*

Our results show overwhelmingly positive net effects of local exposure to large foreign employment shocks on the educational attainment of Malawian children age-eligible for school during these shocks. Our preferred results from the district-level analysis in Table 5 show that these long run education effects were large: a 10-15% increase in years of schooling and a 5 to 8% increase in the share of adults with any primary school. These primary school enrollment effects account for between one third and one half of the total increase in enrollment rates between 1967 and 1978.[29] The evidence suggests that exposure to international labor migration allowed rural Malawian communities with better access to mining jobs to

[29] 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.

invest in education and overtake those communities with less access to mining jobs in terms of human capital accumulation.

Given the size of earnings associated with mine employment in South Africa discussed in Section 2, it seems reasonable that in the wake of labor expansion, investments in human capital would have risen to this extent. The income effects associated with mining employment seemed to have outweighed any substitution effects for child labor at national level.[30] Such large income effects associated with men leaving rural areas of Malawi and remitting earnings homewards might lead us to expect a fall in education among cohorts eligible for schooling during the years of the labor ban in 1974 to 1977. However, as we saw in Figure 2, this labor contraction did not imply an immediate reversal of remittance earnings. The fact that returning miners would have received an equivalent of three years' worth of local income all at once when they lost their jobs means that the income effect of international labor migration lasted a long time.[31]

Of course, the fact that we see educational gains between WNLA and non-WNLA districts disappearing among the youngest cohorts in our sample, those eligible for primary school only after 1977 (see Figure 4), suggests that these income effects do eventually run out. Since the next best alternative to mine work was work on the estate sector, and estate earnings were a little over one tenth of what a miner could earn in the late 1970s, families would never have been able to replace their lost mining income fully, despite the fact that the estate sector was growing rapidly until the late 1970s.[32]

A remaining question is how we understand these economically large and meaningful effects on human capital attainment in light of the fact that over 40% of children report working in 1977? Was it feasible that this investment in human capital stimulated by foreign labor earnings could have come at the expense of child labor? In the next section, we explore evidence for such mechanisms.

v. *Exploring mechanism: Heterogeneous effects across districts with and without estates*

To directly measure whether exposure to mine employment shocks affected the demand for child labor, it would be ideal to have outcomes measuring child labor at the district level in periods before, during and

[30] This finding is consistent with Lu and Treiman (2011) who find that the income effect dominates the substitution effect for households with migrant workers in South Africa. They find that the odds of attending school are 1.3 times higher for children from households with a migrant worker.
[31] We can see this lingering effect among children aged 6 to 15 in 1977 (at the end of the labor ban). Enrollment in school and grade for age attainment is higher among children exposed to the labor ban at school entry age and in WNLA areas (results not shown). Thus, although fathers lost their mining jobs and returned home, the income from saved remittances and involuntary deferred pay enabled children to continue in school during these labor ban years.
[32] Paton (1995, pg 56) notes that while migrants returning from South Africa did move to the estate sector, they were generally too well off to seek immediate employment.

after the 1967 to 1977 period. Unfortunately, such data do not exist. Instead, we use the 1977 Census to provide a snapshot of child labor at the end of the labor ban period and to investigate whether local technologies of production condition the educational investment responses to foreign labor earnings.

Looking back at our summary statistics in Table 3, we see the share of children aged 10 to 19 working for a wage and/or as *mlimi* is lower in WNLA relative to non-WNLA districts and that school enrollment is higher in these WNLA districts. These differences are statistically significant. Note that differences in housework across WNLA and non-WNLA areas do not drive the overall differences in employment shares. Of course, such differences could have existed in earlier periods too, and this is something we cannot account for using a single Census.

Instead, we investigate whether district-level exposure to mine employment shocks (proxied for by the number of WNLA stations in the district) is correlated with lower rates of child labor only in those areas where the opportunity cost of schooling is low. As Section 2 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 demand for child labor would have been much more sensitive to large fluctuations in local male labor supply. In these estate districts, we would expect that removing male labor from the local economy would sharply increase the demand for child labor and raise the opportunity cost of going to school.

Table 8 provides suggestive evidence from the 1977 Census that this was indeed going on. We show correlations between each of the three child labor measures and the number of WNLA stations in the district, for the full sample and for two sets of districts: those with large tobacco and tea plantations and those without.[33] Each regression controls for age group (ages 10 to 14 or ages 15 to 19), gender, the log of population density in 1931 and region fixed effects.

Columns (1), (4) and (7) present regression-adjusted mean differences in rates of child labor that mirror the differences across WNLA and non-WNLA districts shown in Table 3. Overall child labor rates are high, at 40-49%, but 3.2 percentage points fewer children are employed in districts with the mean number of WNLA stations. Splitting the sample into estate and non-estate districts (which means we lose some power to detect statistical differences), we see that these lower child labor rates are only evident in districts without agricultural estates. In contrast, in estate districts, having more WNLA stations is

[33] We take the definition of these districts from Christiansen (1983). We have corroborated these definitions using FAO measures of crop suitability at the district level.

associated with more child labor: child employment in wage work/*mlimi* is almost double in WNLA relative to non-WNLA districts.[34]

These correlations, while not definitive, do suggest something about the mechanisms through which the exposure to mining employment shocks had impacts on long-run educational outcomes. In areas where children were less valuable for work (non-estate areas), the income effect of having more miners in the district easily outweighed the substitution effect, and led to more school enrollment, and consequently higher educational attainment. In estate areas, where children were ready substitutes for male labor, the net effect of more miners in the district should have had smaller effects on school enrollment and thus smaller effects on long run attainment.

Table 9 shows just these relationships. We estimate versions of equation (1) using the district level dataset from the combined 1977 and 1998 Census in Panel A and then the sub-district level dataset from the 1998 Census data in Panel B and split the sample into districts with and without large tea or tobacco estates.

All of the positive long run education effects we estimate for exposed cohorts in exposed districts is driven by variation in access to mine work among districts without large agricultural estates. In these districts, cohorts eligible for primary schooling between 1967 and 1977 have between 0.26 and 0.38 more years of education and are between 2.2 and 3.4 percentage points more likely to be in school (Panel A). Results are similar in Panel B which uses only the younger cohorts as comparisons. These estimates are large, and strongly statistically significant. In contrast, the coefficients estimated for the sample of estate districts are much smaller than for the non-estate sample and never statistically significantly different than zero (although we do run into power concerns in Panel A).

These differential effects of exposure to mine employment shocks on long run educational attainment in estate and non-estate districts links up well with the cross-sectional variation in rates of child labor across estate and non-estate districts. Our results indicate that where child labor was not a good substitute for male labor (i.e. in non-estate districts), the effect of international labor migration and concomitant remittance flows was to substantially increase investments in schooling. Where child labor was a potentially valuable substitute for missing male labor, the effects of this migration on education are much harder to discern, and much more muted in the estate areas of the country. The general lesson we draw from this evidence on mechanisms is that the local technology of production strongly influences the effects that international migration has on families left behind.

[34] Note that examining effects across WNLA and non-WNLA districts in non-estate districts only has the added benefit of controlling for possible confounds coming from economic growth in the estate sector, which was high between 1967 and 1977.

## 7. Conclusions

We have used two waves of complete Census data from 1977 and 1998 to show that the massive and unanticipated expansion and sudden contraction of employment of Malawian men on South African mines over a ten-year period had lasting effects on human capital accumulation in sending districts. This new evidence from Africa shows that the income effects of labor migration on the demand for education outweigh the substitution effects of this migration on the demand for child time in the labor market. Depending on the technology of production– specifically, whether adult and child labor are good substitutes or not – international migration of adults can have long-lasting positive effects on human capital accumulation across generations.

These results have broad relevance for development economists. International migration is considered one of the most immediate ways out of poverty for families, and potentially for whole sending communities. Our paper provides direct evidence on one of the channels through which this migration can positively impact the lives of those left behind, even in contexts where child labor is highly prevalent. Given the centrality of human capital in the development process, our results suggest that better access to foreign work opportunities and foreign incomes could enable poor, rural communities to lay the foundations for economic growth by investing in the education of the next generation.

## References


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

Chamber of Mines of South Africa, 1973. Annual Report, Johannesburg.

Chirwa, Wiseman 1993. "Child and youth labour on the Nyasaland Plantations 1890-1953", *Journal of Southern African Studies*, Vol. 19: 4. December. Pp. 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 2005. "Child Labour in Commercial Agriculture: the Case of Malawi's Tea Industry" Baseline Study Report Prepared for The Employers' Consultative Association of Malawi (ECAM), ILO/ISAMAT Project and ILO/IPEC.

Christiansen, Robert. 1984."The Pattern of Internal Migration in Response to Structural Change in the Economy of Malawi 1966-77", *Development and Change,* Vol. 15: 125 – 151.

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*, 25(3): 83-106.

Coleman, G. 1972 "International Labour Migration from Malawi, 1875-1966", *Journal of Social Science*, Vol. 2: pp. 31-46

Coleman, G. 1977 "Regional and district origins of migrant labour from Malawi", *Journal of Social Science*, Vol. 6: pp. 44-79

Collier, Paul; Hoeffler, Anke; Pattillo, Catherine (2004). "Africa's Exodus: Capital Flight and the Brain Drain as Portfolio Decisions" *Journal of African Economies* **13** (Suppl. 2): ii15–ii54

Crush, Jonathan. S., 1986. "The Extrusion of Foreign Labour from the South African Gold Mining Industry" *Geoforum* Vol17 (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. 2008. "Migrant opportunity and the educational attainment of youth in rural China"

Docquier, Frédéric, and Hillel Rapoport. 2012. "Globalization, Brain Drain, and Development" *Journal of Economic Literature*, 50(3): 681-730

Edmonds, Eric V., and Norbert Schady. 2012 "Poverty Alleviation and Child Labor," *American Economic Journal: Economic Policy*, 4(4): 100-124.

Edmonds, Eric V., 2008. "Child Labor"**,** Handbook of Development Economics, Elsevier

Edwards, Alejandra Cox and Manuelita Ureta, 2003, "International Migration, Remittances and Schooling: Evidence from El Salvador," *Journal of Development Economics* in *72 (2)*: 429–461

Eldring, Line. 2003. *Child Labour in the Tea Sector in Malawi: A Pilot Study,* Fafo, Oslo, 14, 17-20

Foster, Andrew D and Mark R. Rosenzweig 1996. "Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution", *American Economic Review,* American Economic Association, vol. 86(4), pages 931-53, September

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

Government of Nyasaland, 1961. "Aide Memoire on WNLA activities, Provincial Reports, PCN1-21-16", Zomba National Archives

Heyneman, Stephen P. 1980 "The evaluation of human capital in Malawi", *World Bank Staff Working Papers* No. 420.The World Bank, Washington D.C.

Hirschmann, David, 1990 "Malawi's 'Captured' Peasantry: An Empirical Analysis" *The Journal of Developing Areas* Vol.24 (4): 467-488.

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

Kadzamira, Esme Chipo 2003 Where has all the education gone in Malawi? Institute of Development Studies, University of Sussex

Kruger, Diana I., 2007. "Coffee production effects on child labor and schooling in rural Brazil," *Journal of Development Economics*, Vol. 82: 2, pp. 448-463, March.

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

Kydd, Jonathan and Robert Christiansen 1983 “The return of Malawian labour from South Africa and Zimbabwe” *Journal of Modern African Studies* Vol. 21 (2): 311-326

Lu, Yao and Donald J. Treiman 2011 “Migration, Remittances and Educational Stratification among Blacks in Apartheid and Post-Apartheid South Africa”, *Social Forces,* Vol. 89 (4): 1119-1143

Lucas, Robert E. B. 1985 “Mines and migrants in South Africa”, *American Economic Review*, Vol. 75 (5): 1094-1108

Lucas, Robert E. B. 1987 “Emigration to South Africa’s Mines” *American Economic Review* Vol. 77: (3): 313-330

Malawi Ministry of Education 1977 “Major trends in education in Malawi 1974-1976”, UNESCO International Conference on Education

Nyasaland Ministry of Education. 1960. Annual Report. Government Printer, Zomba.

Malawi National Statistical Office, 1972. Malawi Statistical Yearbook. Government Printer, Zomba.

Malawi National Statistical Office, 1995. Malawi Statistical Yearbook. Government Printer, Zomba.

Nyasaland Governor’s Memorandum on Labor Migrancy in Malawi. 1956

Nyasaland Provincial Office Memo, December 7 1961

McKenzie, David and Hillel Rapoport 2011. ”Can migration reduce educational attainment? Evidence from Mexico,” *Journal of Population Economics*, Vol. 24(4): 1331-1358

Otañez, M. G., Muggli, M. E., Hurt, R. D., Glantz, S. A. 2006 “Eliminating child labour in Malawi: a British American Tobacco corporate responsibility project to sidestep tobacco labour exploitation” *Tobacco Control* Vol 15:224–230.

Paton, Bill. 1995. Labour Export Policy in the Development of Southern Africa. Macmillan.

Prothero, R. Mansell.1974. “Foreign migrant labour for South Africa”, International Migration Review, Vol. 8: 3. Pp. 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.

Stark, Oded (2004). “Rethinking the Brain Drain”, *World Development*, Vol. 32: 1: 16.

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)

United Nations Human Development Report, 2013. The rise of the South: Human Progress in a Diverse World. United Nations Development Programme.

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: Annual recruitment of Malawian miners to South African mines, 1950-1994**

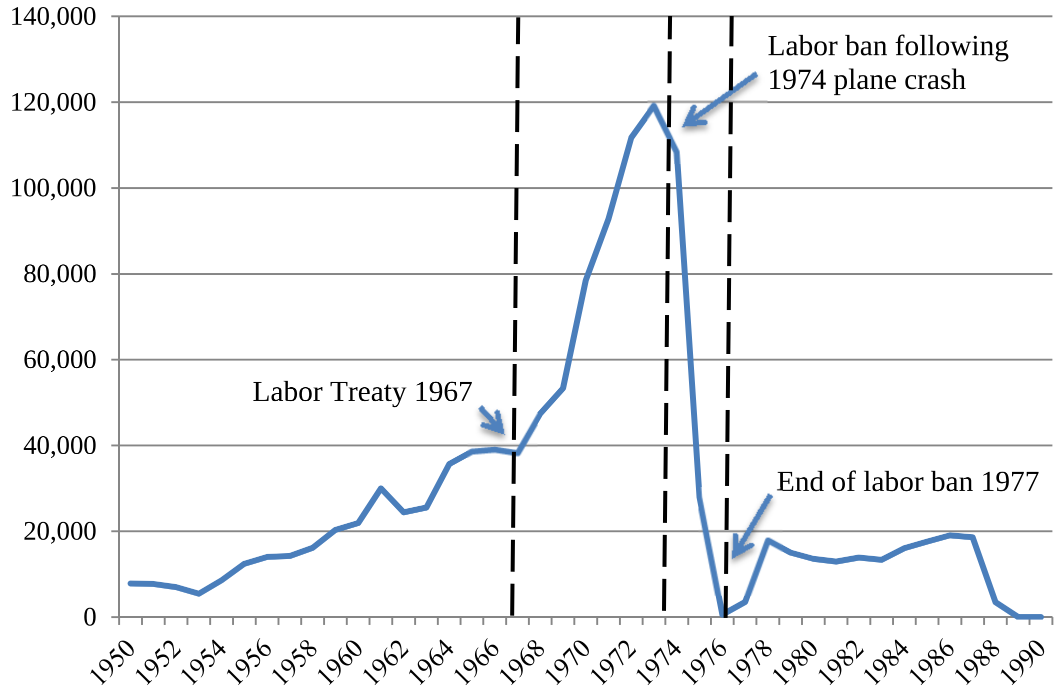
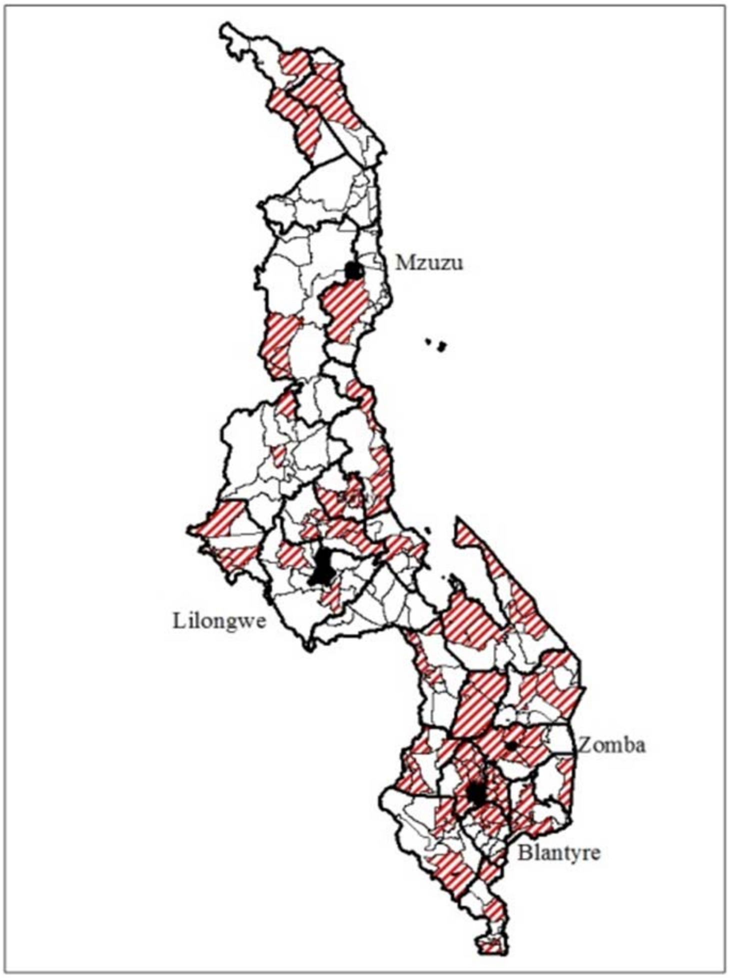



Figure 1 shows number of Malawian workers recruited to work on mines in South Africa in each year, 1950-1994. The three dotted lines represent (from left to right) the initiation of the new recruiting agreement in August 1967, the moratorium on migration to South Africa after the April 1974 Malawian plane crash and the legal resumption of mine migration to South Africa in 1978.

**Figure 2 The spatial distribution of mine recruiting stations across Malawi**




<u>Figure 2</u> shows district boundaries (thick black lines), sub-district/traditional authority boundaries (thinner black lines) and the distribution of WNLA recruiting stations established by 1937 (red hatched areas) across the country. Malawi's four cities (black shaded areas) are excluded from the analyses.

**Figure 3: Mean education by age group and WNLA status of district**

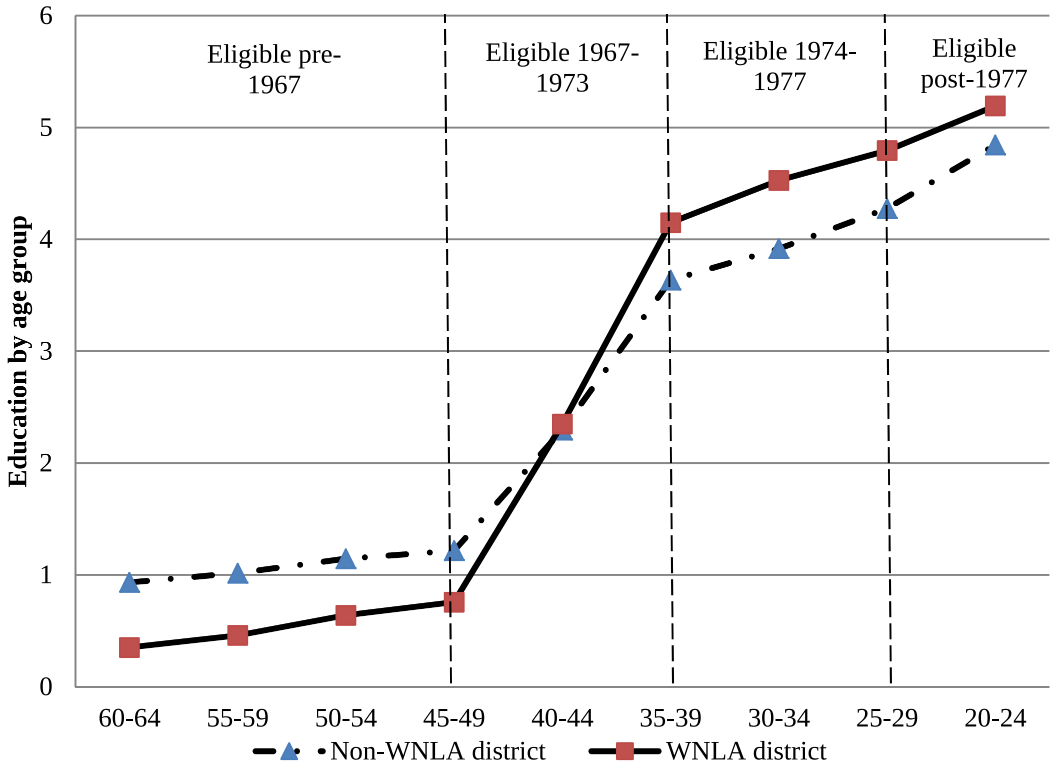



Figure 3 shows age group coefficients from a regression of total years of education on nine age group dummies and their interaction with a WNLA station indicator and other control variables (excluding region-specific trends).

**Figure 4: Age-specific differences in mean education levels across WNLA and non-WNLA districts**

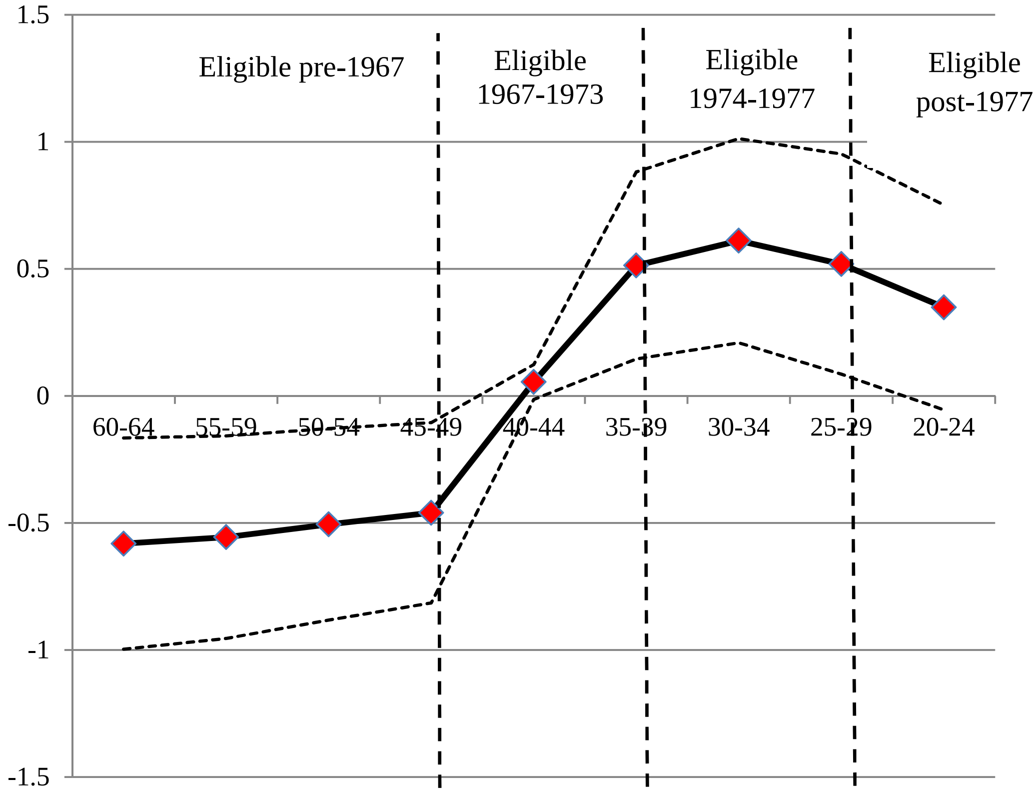



Figure 4: shows estimated interaction term coefficients and 95% confidence intervals from a regression of total years of education on nine age group dummies and their interaction with a WNLA station indicator and other control variables, excluding region-specific trends.

**Table 1: Correlation between WNLA recruiting stations and outmigrant stocks and flows**

| | *Ln Number of Men Ever Worked Abroad, 1977* | | | | *Ln Number of men issued foreign travel documents, 1958-1967* | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Any WNLA station in district | 0.473*** | 0.674** | | | 0.13 | 0.401** | | |
| | (0.163) | (0.264) | | | (0.131) | (0.175) | | |
| Number of WNLA stations | | | 0.0959** | 0.100* | | | 0.0884* | 0.0853** |
| | | | (0.045) | (0.048) | | | (0.050) | (0.034) |
| Log population density 1931 | | 0.366 | | 0.32 | | 0.681*** | | 0.656*** |
| | | (0.296) | | (0.282) | | (0.212) | | (0.202) |
| Estate district | | 0.342 | | 0.305 | | 0.23 | | 0.218 |
| | | (0.497) | | (0.483) | | (0.303) | | (0.295) |
| Central region | 1.005** | 0.712 | 0.907** | 0.646 | 0.851*** | 0.36 | 0.761** | 0.292 |
| | (0.420) | (0.514) | (0.426) | (0.524) | (0.290) | (0.340) | (0.287) | (0.336) |
| Southern region | 1.441*** | 1.175*** | 1.156*** | 0.774* | 1.283*** | 0.708*** | 1.294*** | 0.504* |
| | (0.342) | (0.375) | (0.326) | (0.439) | (0.182) | (0.249) | (0.203) | (0.292) |
| Dummy: 1963-1967 | | | | | -0.202 | -0.202 | -0.21 | -0.202 |
| | | | | | (0.214) | (0.187) | (0.257) | (0.214) |
| N | 23 | 23 | 23 | 23 | 46 | 46 | 46 | 46 |
| R2 | 0.35 | 0.41 | 0.37 | 0.42 | 0.30 | 0.49 | 0.33 | 0.51 |

***p<0.01, **p<0.05, *p<0.1. Robust standard errors in parentheses in all regressions. Unit of observation is the district in columns 1-4; district-year in columns 5-8. District-level data on ever worked abroad is from the 1977 Census; the number of men issued foreign travel documents are reported for 1958-1960 and for 1963-1967 periods in Malawi's Department of Labour Annual Reports. Any WNLA station is a binary indicator for presence of recruiting station in the district in 1937; Number of WNLA stations is a count variable of all stations in the district in 1937.

**Table 2: Correlates of WNLA recruiting stations at district and sub-district level**

| | *Number of recruiting stations in the district* | | | | *Sub-district contains a recruiting station* | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Log population density, Census 1931 | -0.545 | -0.565 | -0.406 | -0.132 | -0.0617** | -0.0862** | -0.063 | -0.044 |
| | (0.332) | (0.392) | (0.447) | (0.514) | (0.029) | (0.038) | (0.039) | (0.049) |
| Estate district | | -0.086 | -0.008 | -0.243 | | -0.084 | -0.062 | -0.097 |
| | | (0.932) | (0.934) | (0.764) | | (0.072) | (0.074) | (0.078) |
| Altitude (meters)^ | | | 0.00141* | 0.000 | | | 0.000142** | 0.000 |
| | | | (0.001) | (0.001) | | | (0.000) | (0.000) |
| Central region | | | | 1.519** | | | | 0.132 |
| | | | | (0.630) | | | | (0.089) |
| Southern region | | | | -1.020 | | | | -0.109 |
| | | | | (1.277) | | | | (0.102) |
| Observations | 27 | 27 | 27 | 27 | 228 | 228 | 228 | 228 |
| R-squared | 0.05 | 0.05 | 0.11 | 0.41 | 0.01 | 0.02 | 0.04 | 0.09 |
| Mean of outcome | 1.74 | 1.74 | 1.74 | 1.74 | 0.19 | 0.19 | 0.19 | 0.19 |

***p<0.01, **p<0.05, *p<0.1. Robust standard errors in parentheses in all regressions, clustered at the sub-district (TA) level in columns 5-8. Outcome is the number of recruiting stations in the district in 1937 or a binary indicator for presence of recruiting station in the subdistrict (TA) in 1937. Log population density and the estate dummy measured at district level. Altitude is average altitude for each TA or district; it stands in as a proxy for malaria risk at the TA or district level.

**Table 3: Summary statistics for Census data**

| | Full sample | | WNLA Recruiting Districts | | Non-WNLA Districts | | *p* value of difference |
|---|---|---|---|---|---|---|---|
| | Mean | s.d. | Mean | s.d. | Mean | s.d. | |
| *Panel A: Outcomes and control variables for District-5 Year Age Group-Sex cells* | | | | | | | |
| *Variables measuring exposure to mining employment shocks* | | | | | | | |
| Number of WNLA stations | 2.00 | 1.96 | 3.20 | 1.52 | 0 | | |
| Any WNLA station | 0.63 | | 1 | | 0 | | |
| Eligible for primary school in 1967-1973 | 0.30 | 0.46 | 0.30 | 0.46 | 0.30 | 0.46 | 0.500 |
| Eligible for primary school in 1974-1977 | 0.20 | 0.40 | 0.20 | 0.40 | 0.20 | 0.40 | 0.500 |
| Younger comparison cohorts: Eligible after 1977 | 0.10 | 0.30 | 0.10 | 0.30 | 0.10 | 0.30 | 0.500 |
| Older comparison cohorts: Eligible before 1967 | 0.40 | 0.49 | 0.40 | 0.49 | 0.40 | 0.49 | 0.500 |
| *Education and child labor variables* | | | | | | | |
| Share of 10-19 year olds enrolled in 1977 | 0.39 | 0.18 | 0.45 | 0.19 | 0.31 | 0.13 | 0.000 |
| Share of 10-19 year olds working in 1977 | 0.42 | 0.21 | 0.38 | 0.21 | 0.47 | 0.19 | 0.018 |
| Share of 10-19 year olds working for wage/Mlimi in 1977 | 0.29 | 0.17 | 0.25 | 0.17 | 0.34 | 0.16 | 0.006 |
| Share of 10-19 year olds working at home in 1977 | 0.13 | 0.10 | 0.13 | 0.10 | 0.13 | 0.10 | 0.836 |
| Total years of education, adults in 1977 and 1998 | 2.56 | 2.20 | 2.85 | 2.36 | 2.08 | 1.80 | 0.000 |
| Share with any primary school, adults in 1977 and 1998 | 0.41 | 0.29 | 0.45 | 0.30 | 0.35 | 0.25 | 0.000 |
| *Control variables* | | | | | | | |
| Female | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.50 | 0.500 |
| Log Population density 1931 | 3.11 | 0.78 | 2.77 | 0.56 | 3.67 | 0.76 | 0.000 |
| Number of observations | 480 | | 300 | | 180 | | |
| Number of districts | 24 | | 15 | | 9 | | |

*…continued*

**Table 3 (continued): Summary statistics for Census data**

| | Full sample | | WNLA Recruiting Districts | | Non-WNLA Districts | | *p* value of difference |
|---|---|---|---|---|---|---|---|
| | Mean | s.d. | Mean | s.d. | Mean | s.d. | |
| *Panel B: Geographic and historical district variables* | | | | | | | |
| Any WNLA station | 0.59 | | 1 | | 0 | | |
| Number of WNLA stations | 1.74 | | 2.94 | | 0 | | |
| Fraction of men ever worked abroad by 1977~ | 0.35 | 0.09 | 0.37 | 0.10 | 0.31 | 0.05 | 0.103 |
| Area (km squared) | 3,437 | 2,284 | 4,033 | 2,484 | 2,570 | 1,707 | 0.041 |
| Altitude: high malaria area=1 | 0.19 | 0.40 | 0.13 | 0.34 | 0.27 | 0.47 | 0.188 |
| Population density in 1931 | 24.35 | 21.25 | 15.64 | 9.27 | 37.02 | 27.34 | 0.009 |
| District contains an estate | 0.22 | 0.42 | 0.25 | 0.45 | 0.18 | 0.45 | 0.683 |
| Fraction of ages 5-18 literate in 1945 | 0.08 | 0.04 | 0.09 | 0.04 | 0.06 | 0.02 | 0.001 |
| Fraction of men working for a wage (farm, cash or other) 1966 | 0.63 | 0.11 | 0.60 | 0.12 | 0.66 | 0.07 | 0.091 |
| Fraction of men with a farm wage 1966 | 0.33 | 0.14 | 0.35 | 0.15 | 0.30 | 0.12 | 0.185 |
| Fraction of men with a cash wage 1966 | 0.16 | 0.09 | 0.13 | 0.08 | 0.20 | 0.09 | 0.015 |
| Fraction of men with another source of wage 1966 | 0.14 | 0.04 | 0.13 | 0.05 | 0.15 | 0.04 | 0.105 |
| Fraction of men not earning any wage 1966 | 0.37 | 0.11 | 0.40 | 0.12 | 0.34 | 0.07 | 0.091 |
| Number of districts | 27 | | 16 | | 11 | | |

Data in Panel A are means from the 1998 micro data and the 1977 aggregate data, reported at the district-5 year age group-sex level. Data in Panel B are district level means from geographic data, aggregate Census data in 1931, 1945, 1966, 1977 and administrative data. ~Fraction of men ever outmigrated by 1977 only available for 23 of 27 districts. *p* values are reported for the test of the difference in means across recruiting and non-recruiting station areas using robust standard errors. Estate is a dummy variable as described in the text.

**Table 4: Mean educational attainment by cohort and exposure to mining employment shocks at district level**

| | *Years of education [Mean=2.56]* | | | *Any primary school [Mean=0.41]* | | |
|---|---|---|---|---|---|---|
| | WNLA recruiting districts | Non-WNLA districts | *Difference* | WNLA recruiting districts | Non-WNLA districts | *Difference* |
| *Cohorts exposed to mine employment shocks* | | | | | | |
| 1) Age eligible for school during labor expansion, 1967-1973 | 3.24 | 2.25 | 0.98*** | 0.50 | 0.38 | 0.12*** |
| | (2.15) | (1.57) | | (0.28) | (0.23) | |
| | 90 | 54 | | 90 | 54 | |
| 2) Age eligible for school during labor contraction, 1974-1977 | 4.97 | 3.62 | 1.35*** | 0.74 | 0.58 | 0.16*** |
| | (1.73) | (1.39) | | (0.16) | (0.17) | |
| | 60 | 36 | | 60 | 36 | |
| *Comparison cohorts not exposed to mine employment shocks* | | | | | | |
| 3) Combined younger and older comparison cohorts | 1.76 | 1.35 | 0.41*** | 0.29 | 0.24 | 0.05*** |
| Age eligible for school before 1967 or after 1977 | (2.04) | (1.66) | | (0.26) | (0.23) | |
| | 150 | 90 | | 150 | 90 | |
| *Difference in differences:* | | | | | | |
| (Eligible 1967 to 1973) - (Combined comparison cohorts) (1) - (3) | | | 0.58*** | | | 0.07*** |
| (Eligible 1974 to 1977) - (Combined comparison cohorts) (2) - (3) | | | 0.94*** | | | 0.11*** |

Means, standard deviations and number of observations reported in each block of cells. Unit of observation is the district-5 year age group-sex cell. Difference-in-differences are estimated with no additional controls, standard errors are robust and clustered at the district level. $p<0.1$*, $p<0.05$**, $p<0.01$***. Eligible cohorts include individuals aged 6-12 during the 1967-1973 expansion of mine employment and those age 6-12 during the contraction of employment. The comparison cohort includes younger cohorts eligible for primary school only after 1977 and older cohorts eligible for primary school before 1967. Adult sample is restricted to ages 20-44 in both Census years.

**Table 5: Long run effects of mine employment shocks on education: Difference-in-differences at district level**

| | *Total years of schooling attained* | | | | *Share with any primary school* | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Eligible 1967-1973*Num. WNLA stations | 0.172 | 0.216 | 0.234** | 0.252* | 0.016 | 0.022 | 0.022* | 0.022* |
| | (0.178) | (0.130) | (0.128) | (0.132) | (0.020) | (0.012) | (0.012) | (0.012) |
| Eligible 1974-1977*Num. WNLA stations | 0.284 | 0.314 | 0.334* | 0.378* | 0.028 | 0.032* | 0.032* | 0.034* |
| | (0.210) | (0.184) | (0.184) | (0.184) | (0.020) | (0.018) | (0.018) | (0.016) |
| Eligible 1967-1973 | 1.098*** | -0.661*** | -0.679*** | -0.697*** | 0.169*** | -0.087*** | -0.088*** | -0.089*** |
| | (0.182) | (0.137) | (0.136) | (0.147) | (0.023) | (0.013) | (0.013) | (0.014) |
| Eligible 1974-1977 | 2.589*** | -0.358* | -0.377* | -0.419* | 0.376*** | -0.034* | -0.036* | -0.037* |
| | (0.250) | (0.195) | (0.196) | (0.221) | (0.029) | (0.018) | (0.018) | (0.020) |
| Num WNLA Stations | 0.186** | 0.254*** | | | 0.018* | 0.016* | | |
| | (0.068) | (0.054) | | | (0.010) | (0.008) | | |
| Additional controls | N | Y | Y | Y | N | Y | Y | Y |
| District FE | N | N | Y | Y | N | N | Y | Y |
| Region trends | N | N | N | Y | N | N | N | Y |
| N | 480 | 480 | 480 | 480 | 480 | 480 | 480 | 480 |
| R2 | 0.37 | 0.87 | 0.90 | 0.93 | 0.35 | 0.89 | 0.93 | 0.94 |
| Mean of outcome variable | 2.56 | 2.56 | 2.56 | 2.56 | 0.41 | 0.41 | 0.41 | 0.41 |

***$p$ <0.01, **$p$ <0.05, *$p$ <0.1. Robust standard errors clustered at the district level. Unit of observation is the district-5 year age group-sex cell. Vector of controls includes female, age group, a Census year indicator, two region fixed effects and the log of district-level population density in 1931. Number of WNLA stations in the district is a count variable, marginal effects evaluated at the mean of this variable. Sample includes adults ages 20 to 44 in 1977 and 1998 census.

**Table 6: Long run effects of mine employment shocks on education: Difference-in-differences at sub-district level**

| | *Total years of schooling attained* | | | | *Share with any primary school* | | | | *Share bilingual* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Eligible 1967-1973*WNLA station | 0.269* | 0.212* | 0.294*** | 0.261** | 0.026 | 0.017 | 0.033*** | 0.021* | 0.027 | 0.021 | 0.029*** | 0.027** |
| | (0.152) | (0.118) | (0.102) | (0.111) | (0.017) | (0.014) | (0.011) | (0.011) | (0.018) | (0.013) | (0.011) | (0.012) |
| Eligible 1974-1977*WNLA station | 0.169* | 0.142** | 0.172*** | 0.157** | 0.017** | 0.013* | 0.020*** | 0.015** | 0.018* | 0.015** | 0.018*** | 0.017** |
| | (0.088) | (0.065) | (0.062) | (0.066) | (0.008) | (0.007) | (0.006) | (0.006) | (0.010) | (0.007) | (0.007) | (0.007) |
| Eligible 1967-1973 | -1.529*** | -0.420*** | -0.393*** | -0.385*** | -0.172*** | -0.063*** | -0.059*** | -0.056*** | -0.154*** | -0.031*** | -0.028*** | -0.028*** |
| | (0.082) | (0.049) | (0.043) | (0.047) | (0.009) | (0.006) | (0.005) | (0.005) | (0.009) | (0.005) | (0.005) | (0.005) |
| Eligible 1974-1977 | -0.814*** | -0.136*** | -0.132*** | -0.130*** | -0.089*** | -0.014*** | -0.015*** | -0.014*** | -0.079*** | 0.002 | 0.002 | 0.002 |
| | (0.053) | (0.033) | (0.031) | (0.033) | (0.005) | (0.003) | (0.003) | (0.003) | (0.006) | (0.003) | (0.003) | (0.003) |
| WNLA Station | 0.231 | -0.024 | -0.264 | -0.250 | 0.038 | 0.001 | -0.025 | -0.021 | 0.025 | -0.003 | -0.027 | -0.026 |
| | (0.244) | (0.193) | (0.181) | (0.185) | (0.025) | (0.022) | (0.016) | (0.016) | (0.029) | (0.021) | (0.020) | (0.020) |
| Additional controls | N | Y | Y | Y | N | Y | Y | Y | N | Y | Y | Y |
| District FE | N | N | Y | Y | N | N | Y | Y | N | N | Y | Y |
| Region trends | N | N | N | Y | N | N | N | Y | N | N | N | Y |
| N | 7,413 | 7,413 | 7,413 | 7,413 | 7,413 | 7,413 | 7,413 | 7,413 | 7,413 | 7,413 | 7,413 | 7,413 |
| R2 | 0.40 | 0.62 | 0.73 | 0.73 | 0.35 | 0.56 | 0.74 | 0.74 | 0.38 | 0.65 | 0.75 | 0.75 |
| Mean of outcome variable | 4.03 | 4.03 | 4.03 | 4.03 | 0.64 | 0.64 | 0.64 | 0.64 | 0.40 | 0.40 | 0.40 | 0.40 |

***$p$ <0.01, **$p$ <0.05, *$p$ <0.1. Unit of observation is the sub-district -age-sex level. Robust standard errors clustered at the sub-district level. Vector of controls includes: female, age, age squared, two region fixed effects and the log of district-level population density in 1931. WNLA station is an indicator variable that identifies which subdistrict (TA) had a WNLA recruiting station in 1937. Sample includes adults ages 20 to 44 in 1998.

**Table 7: Long run effects of mine employment shocks on education: Bounds for internal migration at district level**

| | *Total years of education* | | | | *Share with any primary school* | | | |
|---|---|---|---|---|---|---|---|---|
| Assumptions about education levels of net migrants: | Minimum years of schooling | | Maximum years of schooling | | Lowest share with primary school | | Highest share with primary school | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Eligible 1967-1973*Num. WNLA stations | 0.304* | 0.328* | 0.246* | 0.274** | 0.028* | 0.028* | 0.024* | 0.024* |
| | (0.166) | (0.178) | (0.134) | (0.132) | (0.014) | (0.014) | (0.012) | (0.012) |
| Eligible 1974-1977*Num. WNLA stations | 0.472* | 0.526* | 0.328* | 0.39** | 0.044** | 0.044* | 0.032* | 0.034** |
| | (0.238) | (0.262) | (0.190) | (0.174) | (0.020) | (0.022) | (0.018) | (0.016) |
| 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 | 0.90 | 0.92 | 0.89 | 0.92 | 0.93 | 0.94 | 0.92 | 0.93 |
| Mean of outcome variable | 2.59 | 2.59 | 2.53 | 2.53 | 0.41 | 0.41 | 0.41 | 0.41 |

***$p$ <0.01, **$p$ <0.05, *$p$ <0.1. Robust standard errors clustered at the district level. Unit of observation is the district-5 year age group-sex cell. Vector of controls includes female, age group, a Census year indicator, two region fixed effects and the log of district-level population density in 1945. Number of WNLA stations in the district is a count variable, marginal effects evaluated at the mean of this variable. Sample includes adults ages 20 to 44 in 1977 and 1998 census. Comparison cohorts are older and younger cohorts. 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-sex cell. Details of variable construction are explained in the text.

**Table 8: Correlation between mine employment shock exposure and child labor shares in 1977 - Heterogeneity across estate and non-estate districts**

| | *Share ages 10-19 working at all* | | | *Share ages 10-19 working for wage or as Mlimi* | | | *Share ages 10-19 working at home* | | |
|---|---|---|---|---|---|---|---|---|---|
| | All districts | Districts with No Estates | Districts with Estates | All districts | Districts with No Estates | Districts with Estates | All districts | Districts with No Estates | Districts with Estates |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Number of WNLA stations in district | -0.032** | -0.042*** | 0.16 | -0.032** | -0.032*** | 0.338*** | 0.000 | -0.01 | -0.18** |
| | (0.010) | (0.012) | (0.116) | (0.006) | (0.008) | (0.106) | (0.008) | (0.008) | (0.078) |
| N | 108 | 80 | 28 | 108 | 80 | 28 | 108 | 80 | 28 |
| R2 | 0.82 | 0.82 | 0.91 | 0.77 | 0.77 | 0.91 | 0.60 | 0.62 | 0.65 |
| Mean of outcome | 0.42 | 0.39 | 0.48 | 0.29 | 0.26 | 0.36 | 0.13 | 0.13 | 0.13 |

Notes: Robust standard errors, p<0.1*, p<0.05**, p<0.01***. Share working includes all work outside the home as well as work inside the home; total wage work includes Mlimi (farm labor) and wage work for a firm. Sample restricted to rural areas. Other controls: age dummy for ages 15 to 19, log population density in 1931, Central and Southern region indicators. Mean number of WNLA stations in district=2

**Table 9: Long run effects of mine employment shocks on education: Heterogeneous effects in estate and non-estate districts**

| | *Total years of education* | | *Share with any primary school* | |
|---|---|---|---|---|
| | Districts with No Estates | Districts with Estates | Districts with No Estates | Districts with Estates |
| | (1) | (2) | (3) | (4) |
| | *Panel A: Combined older and younger comparison cohorts, district level* | | | |
| Eligible 1967-1973*Num. WNLA stations | 0.266* | 0.120 | 0.022* | 0.020 |
| | (0.136) | (0.150) | (0.012) | (0.022) |
| Eligible 1974-1977*Num. WNLA stations | 0.384** | 0.174 | 0.034* | 0.024 |
| | (0.178) | (0.182) | (0.016) | (0.028) |
| N | 360 | 120 | 360 | 120 |
| R2 | 0.93 | 0.93 | 0.94 | 0.94 |
| Mean of outcome | 2.64 | 2.32 | 0.42 | 0.39 |
| | *Panel B: Younger comparison cohorts only, sub-district level* | | | |
| Eligible 1967-1973*Any WNLA station | 0.306** | 0.177 | 0.026* | 0.006 |
| | (0.132) | (0.209) | (0.014) | (0.019) |
| Eligible 1974-1977*Any WNLA station | 0.157* | 0.157 | 0.014** | 0.012 |
| | (0.082) | (0.111) | (0.007) | (0.009) |
| N | 2,258 | 5,155 | 2,258 | 5,155 |
| R2 | 0.71 | 0.75 | 0.76 | 0.73 |
| Mean of outcome | 4.08 | 4.00 | 0.65 | 0.64 |

***$p$ <0.01, **$p$ <0.05, *$p$ <0.1. Robust standard errors clustered at the district level. Unit of observation is the district-5 year age group-sex cell in Panel A and the sub-district -age-sex cell in Panel B. Estate denotes those districts which have substantial presence of tobacco and sugar estates, as described in the text. All regressions control for female, age and age squared, the log of district-level population density in 1931, a full set of district fixed effects and region-specific trends. WNLA station is an indicator variable. Mean number of WNLA stations by district is 2; marginal effects in Panel A evaluated at this mean.

**Appendix Table 1: Distribution of treatment and control cohorts**

| Birth cohort | Age in 1967 | Age in 1974 | Age in 1977 | Age in 1998 | Treatment and control cohorts |
|---|---|---|---|---|---|
| 1933-1953 | 14-34 | 21-41 | 24-44 | 45-65 | Older comparison cohorts eligible for primary schooling before 1967 (adult outcomes from 1977 Census) |
| 1954-1961 | 6-13 | 13-20 | 16-23 | 37-44 | Cohorts exposed eligible for primary school during expansion of labor demand 1967-1973 (adult outcomes from 1977 or 1998 Census) |
| 1962-1971 | 0-5, not yet born | 3-12 | 6-15 | 27-36 | Cohorts eligible for primary school during contraction of labor demand 1974-1977 (adult outcomes from 1998 Census) |
| 1972-1978 | not yet born | not yet born 0-2 | not yet born, 0-5 | 20-26 | Younger comparison cohorts eligible for primary schooling after 1977 (adult outcomes from 1998 Census) |

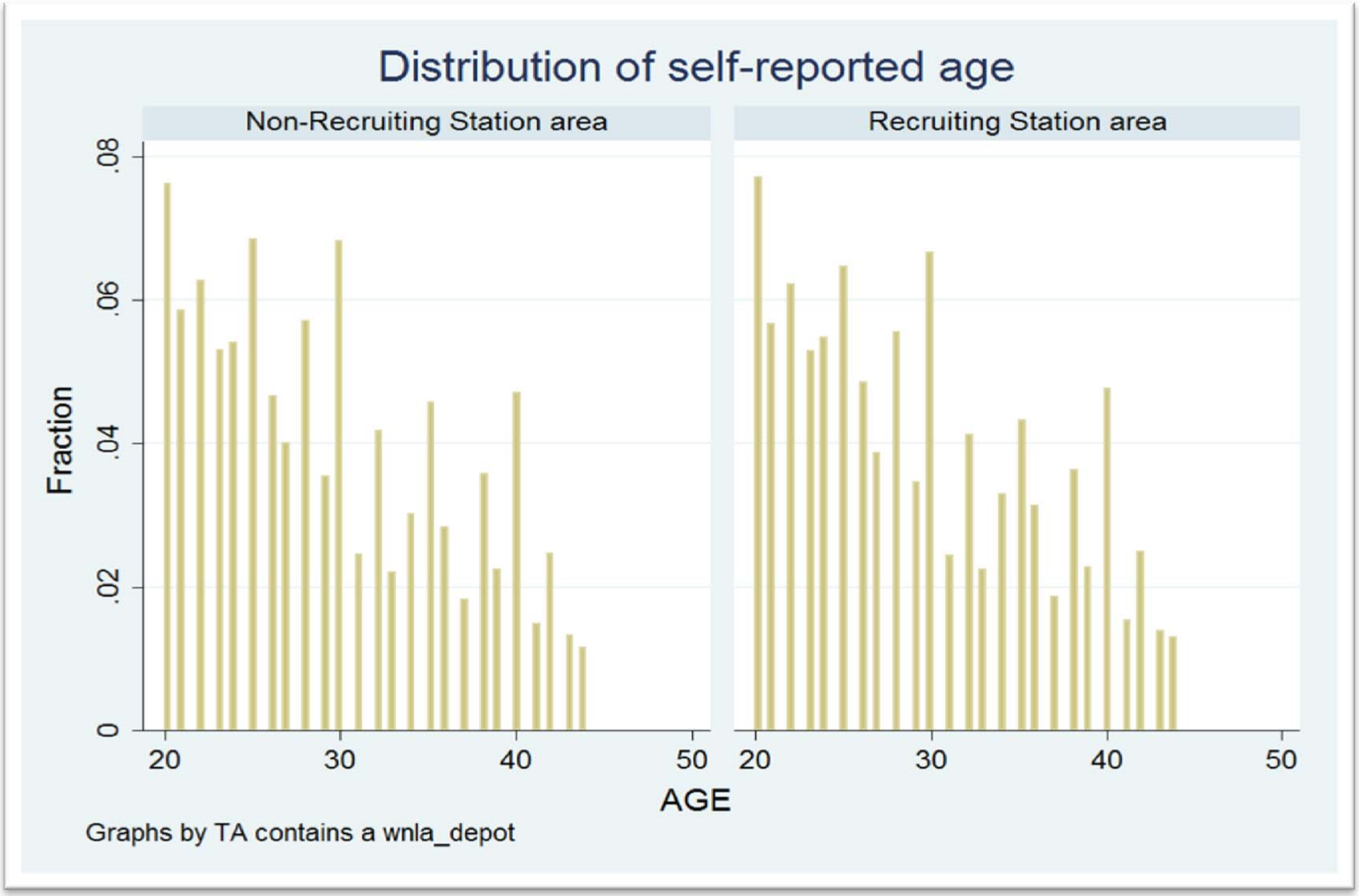
Notes: This table explains which cohorts were age eligible for primary school during each of the 1967-1973 and 1974-1977 shocks, and which cohorts were age ineligible. The main analysis sample from the 1998 Census includes individuals aged 20 to 44 in 1998, i.e. born between 1954 and 1978 (blue shaded area). We use cohorts aged 20 to 44 in the 1977 Census to build the older comparison cohorts (orange shaded area) and to add to the older treated cohorts (part of the blue shaded area).

**Appendix Table 2: Mean educational attainment by cohort and exposure to mining employment shocks at sub-district level**

| | *Years of education [Mean=4.02]* | | | *Any primary school [Mean=0.64]* | | | *Bilingual [Mean=0.40]* | | |
|---|---|---|---|---|---|---|---|---|---|
| | WNLA area | Non-WNLA area | *Difference* | WNLA area | Non-WNLA area | *Difference* | WNLA area | Non-WNLA area | *Difference* |
| *Cohorts exposed to mine employment shocks* | | | | | | | | | |
| 1) Age eligible for school during labor expansion, 1967-1973 | 3.606 | 3.098 | 0.508* | 0.600 | 0.535 | 0.065** | 0.359 | 0.306 | 0.053* |
| | (1.725) | (1.671) | | (0.203) | (0.216) | | (0.201) | (0.189) | |
| | 434 | 1,287 | | 434 | 1,287 | | 434 | 1,287 | |
| 1) Age eligible for school during labor contraction, 1974-1977 | 4.220 | 3.826 | 0.395** | 0.673 | 0.619 | 0.054** | 0.424 | 0.382 | 0.042 |
| | (1.725) | (1.780) | | (0.182) | (0.204) | | (0.193) | (0.197) | |
| | 725 | 2,345 | | 725 | 2,345 | | 725 | 2,345 | |
| *Comparison cohorts not exposed to mine employment shocks* | 4.853 | 4.629 | 0.223 | 0.745 | 0.707 | 0.037 | 0.484 | 0.460 | 0.024 |
| 3) Younger comparison cohorts | (1.668) | (1.772) | | (0.165) | (0.176) | | (0.196) | (0.194) | |
| Age eligible for school after 1977 | 578 | 2,044 | | 578 | 2,044 | | 578 | 2,044 | |
| *Difference in differences:* | | | | | | | | | |
| Eligible 1967 to 1973 - Younger comparison cohort (1) - (3) | | | 0.285*** | | | 0.028*** | | | 0.029*** |
| Eligible 1974 to 1977 - Younger comparison cohort (2) - (3) | | | 0.171*** | | | 0.017*** | | | 0.018*** |

Means, standard deviations and number of observations reported in each block of cells. Unit of observation is the sub-district-age-gender cell. Difference-in-differences are estimated with no additional controls, and standard errors are robust and clustered at the sub-district (TA) level. *p* <0.1*, *p* <0.05**, *p* <0.01***. Eligible cohorts include individuals aged 6-12 during the 1967-1973 expansion of employment and those age 6-12 during the contraction of employment. The comparison cohort includes younger cohorts eligible for primary school only after 1977.

**Appendix Figure 1: Distribution of self-reported ages by recruiting station area**




Histograms show the fraction of the sample reporting each age, by whether the traditional authority has a recruiting station or not.