Labor Migration, Capital Accumulation, and the Structure of Rural Labor Markets*

Taryn Dinkelman † Grace Kumchulesi ‡ Martine Mariotti §

First version: November 2015
This version: September 2022

Abstract

Between 1967 and 1974, a bilateral labor treaty boosted circular labor migration from Malawi to South Africa by 200%, triggering the return of over 53 million USD in migrant earnings to origin communities. A deadly migrant worker plane crash in 1974 ended the flow of money and all migrants were repatriated. We investigate how local labor markets responded to the temporary increase in the local supply of capital generated by this natural experiment. In districts receiving more migrant capital as a result of the plane crash, we find that workers -– particularly women – moved out of farming and into more capital-intensive non-farm service sectors in thirty years following the end of migration. We show that investments in non-farm physical capital and in human capital contribute to these sectoral labor shifts. Our results demonstrate that large, temporary increases in the local supply of capital can trigger persistent changes in rural labor markets over the long run. [156 words]

.

Keywords: international migration, capital accumulation, structural change, Africa
JEL Codes: J21, J24, O15, O16, O47, O55

*This project was funded by the UK Department for International Development (DFID) and the Institute for the Study of Labor (IZA) for the benefit of developing countries (GA-C2-RA4-181). Views expressed are those of the authors and not necessarily those of DFID or IZA. Freed Kumchulesi, Spencer Lambdin, Anwita Mahajan, Annelise Sauter-Ortiz, Mahnum Shazad, Dustin Sheehan, Khwima Singini, Ashley Wong, Lucy Xie and Zheng-Yi Yang provided excellent research assistance. We thank officials at the Malawi National Archives and Lucy McCann at the Rhodes House Library for their assistance in data collection. The paper has benefited from comments at the 2016 AEA Meetings, the 2016 Australasian Development Economics Workshop, the 2017 Advances in Micro-Development conference at the Barcelona GSE Summer Forum, Boston University, IFPRI, the First and Second GLM-LIC Conferences in Washington DC October 2015 and Oxford UK October 2016, the 10th International Conference on Migration and Development (France), the 2016 SOLE meetings, the 2017 Spring BREAD meetings, the 2017 Agri-Dev conference at Yale University, Pontificia Universidad Catolica de Chile, the School of Foreign Service at Georgetown University, the Third Annual Workshop on Natural Experiments in Economic History: Labour, Development and Health at Deakin University, Duke University (Sanford), Johns Hopkins University (SAIS), the Universities of Colorado-Denver, Notre Dame, Sydney, Virginia (Darden), Stanford, Stockholm, and Wellesley College.

†Corresponding author: Dinkelman, University of Notre Dame, NBER, BREAD and IZA. tdinkelm@nd.edu

‡Kumchulesi: National Planning Commission of Malawi. gkumchulesi@npc.mw

§Mariotti: Australian National University and Stellenbosch University. martine.mariotti@anu.edu.au

1 Introduction

What factors facilitate a structural shift of employment out of agriculture? What do workers in poor, rural areas require in order to move away from low productivity farming and towards more productive jobs in manufacturing and services? Such questions are at the heart of classic theories of economic development (Lewis, 1954; Rosenstein-Rodan, 1943) and continue to motivate economists to explore why sectoral productivity gaps persist, and what factors might trigger structural change in low-income countries.[1]

Theories of structural transformation typically rely on exogenous productivity growth or income shocks to generate incentives for labor to move across sectors. Recent empirical evidence points to specific technology and trade shocks as important triggers for structural transformation in the labor market. For example, in Brazil, new labor-saving hybrid seeds released workers from low- productivity agriculture, shifting them towards industrial jobs (Bustos, Caprettini and Ponticelli, 2016; Bustos, Garber, and Ponticelli, 2020). In Vietnam, trade liberalization reallocated farm workers towards higher productivity non-farm enterprises by changing relative prices of goods across these sectors (McCaig and Pavcnik, 2018). And in China, accession to the WTO reduced tariff uncertainty to a greater extent in the non-farm relative to the farm sector, resulting in employment shifts out of agriculture(Erten and Leight, 2021).

Our paper provides new empirical evidence on an alternative trigger for long-run structural change: a temporary increase in the local supply of capital provided by international migrant workers. Analyzing a natural experiment in Malawi that dramatically changed opportunities for international migration over a short time period, we show that a temporary migrant capital surge led to persistent changes in the structure of local labor markets that emerged over the following thirty years.[2]

[1]For example, Gollin (2002) and Gollin, Lagakos and Waugh (2014) document and analyze productivity gaps across sectors in developing countries. McMillan, Rodrik and Verduzco-Gallo (2014) document patterns of structural change in African countries since the 1990s.

[2]In prior work Dinkelman and Mariotti (2016), we exploited some features of this natural experiment to investigate how these international migration shocks affected contemporaneous human capital attainment.

We exploit spatial variation in migrant money injected into local labor markets in Malawi between 1974 and 1975 following large, unanticipated shocks to international migration opportunities. In 1967, lucrative opportunities to migrate to the South African gold mines were opened up with a new treaty between the two countries facilitating legal, temporary migration contracts. This opportunity ended unexpectedly in the mid-1970s, with the banning and subsequent repatriation of all Malawian migrants in response to a mineworker plane crash that killed miners in transit in 1974. Migrants at different points in their time-limited contracts at the time of this plane crash were repatriated with different amounts of money, generating exogenous spatial variation in migrant capital per worker.

Between 1967 and 1975, legal temporary migration to South Africa rose by 200% and then fell to zero. This short, sharp shock to international migration was associated with a large capital injection: Malawi received over 53 million USD in compulsory migrant remittances during this time; three times larger than the country's foreign aid from the US in 1974. Figure 1 illustrates this temporal variation in labor migration (on the left hand side y-axis) and in migrant money (on the right hand side y-axis). One quarter of the total remittance flows over the entire 1967-1977 period occurred in the 19 months after the plane crash. We hone in on this 1974-1975 period to leverage variation in migrant capital across districts generated by the exogenously-timed plane crash.

To study how these migrant capital injections affected local labor markets, we build a panel dataset of money flows and labor market outcomes by digitizing archival material on district-level remittances and matching it to six waves of Malawian population census data. The main treatment variable is the amount of migrant capital per worker received by each district immediately after the migration ban, when all migrants who were still away had returned home, and were paid out their accumulated deferred pay.

Our empirical strategy compares subsequent changes in labor market outcomes across districts that receive more versus less migrant money per returned worker, controlling for district and decade fixed effects, and allowing for time-varying effects of baseline covariates.

The key identification assumption is that conditional on district and decade fixed effects, and controlling for differential trends related to baseline covariates, districts receiving more or less capital per worker after the plane crash would have evolved similarly in the absence of this injection of capital.

After the end of the international migration episode, we find that districts with larger capital injections per migrant experienced a greater share of workers shifting out of farming and into non-farm service activities. Sectoral shifts began slowly in the first ten years following the end of migration, once all migrants had returned, and persisted into the second and third decades post-shock, long after the migrant capital surge dissipated. Women and men shifted into the service sector, specifically into construction (men) and into retail sector jobs (both). Impacts are greater, and most robust, for women.

Our structural change results are unlikely to be driven by miners themselves acquiring more human capital or changing their aspirations while abroad. Mine work was extremely low-skilled work and miners were constrained to live in mining compounds, physically segregated from the South African economy during apartheid. Moreover, our results are strongest in subsequent decades for women, who would not have been migrant mineworkers.

The biggest threat to identification is the possibility that the structure of work in high migrant capital per worker districts might have changed even in the absence of the temporary migrant capital infusion. Because disaggregated employment data do not exist for Malawi prior to the 1970s, we cannot directly test for parallel pre-trends in employment outcomes across high and low migrant capital per worker districts. Instead, we test for parallel trends in population growth – another typical measure capturing the level of development in a district – prior to the plane crash. We cannot reject parallel trends in population growth prior to 1967, suggesting that high and low migrant capital per worker districts were not on differential growth paths before the migrant capital injection. However, we show that the migrant capital injection triggers population growth across districts, as well as a spatial shift in population towards market towns and urban areas within a district. These changes in the distribution of population within a district are additional indicators of structural change in the rural economy.

The evidence in our paper is consistent with migration-induced capital accumulation triggering a slow but steady expansion of the non-farm service sector in rural Malawi. Economic theory suggests this could have happened through some combination of a temporary positive income shock and an increased supply of capital invested in the farm and/or non-farm sectors (Herrendorf, Rogerson and Valentinyi, 2014).

To investigate mechanisms for the persistence of changes in the labor market, we assemble data on farm and non-farm investment goods from population and agricultural censuses, household income and expenditure surveys, and demographic and health surveys, before and after the migration shock. While our data are unable to tease out all possible mechanisms of influence, we provide evidence that both human capital accumulation and non-farm physical capital (rather than farm capital) account for some of the long-run changes we identify. Human capital gains in high migrant capital districts are larger for females than for males.[3] And, fifteen years after the end of the temporary migration episode, households in districts that received more migrant capital were significantly wealthier and held more private assets. A key result we show is that these wealth gains are not solely driven by younger cohorts, who would have benefited from human capital gains funded by migrant earnings. Instead, older cohorts also benefit from being in higher migrant-capital districts, suggesting that education gains are only a part of the story behind the structural changes we measure in our data.

The primary contribution of our paper is to empirically demonstrate that a once-off capital shock, facilitated by temporary international migration, can trigger persistent changes in labor market structure over the long-run. Our paper contributes to a rapidly growing literature on the broader impacts of remittances in developing countries.[4] Clemens and McKenzie (2018) note that the cross-country literature on effects of remittances on home

[3]The education result echoes findings in our prior paper Dinkelman and Mariotti (2016), which uses a related identification strategy but does not split out education impacts by gender.

[4]There is a long tradition of estimating how migrant money affects migrant households, e.g. Dustmann and Kirchkamp (2002), Woodruff and Zeneto (2007), Yang (2008), Kinnan, Wang and Wang (2016).

country outcomes is largely inconclusive, and that economists remain "surprisingly unsure of its (remittances) broad development effects". Addressing this deficit in the literature, a handful of recent empirical studies use natural experiments and exogenous shocks to migrant demand or to remittances *within* countries to uncover the broader, market-level development impacts of international migration. For example, Theoharides (2018) shows that increased (exogenous) demand for migrants from the Philippines stimulates higher levels of schooling attainment at local labor market level. Theoharides (2020) also finds that shutting down the largest Filipino migration channel to Japan reduced income, and increased labor force participation and child labor in sending labor markets of the Philippines. These two papers demonstrate that shocks transmitted through international migration can have profound effects on sending country economies.

Our paper is most closely related to Khanna, Murathanoglu, Theoharides, and Yang (2022), who show that an exchange rate-induced shock to migrant remittances to the Philippines has positive long-run impacts on local incomes in sending labor markets. However, where their findings are based in a setting with long-standing and continued migration flows and focus on outcomes two decades after an exchange rate shock, we show that even once-off episodes of temporary migration can have long-lasting (up to three decades), persistent effects in the sending economy. Both our study, and theirs, demonstrate that migrant capital matters for long-run outcomes in local labor markets, and that decades of data are required to detect impacts on employment and incomes.

A second key contribution of our paper is to highlight the potential importance of temporary international migration and associated migrant capital for Africa, the region with perhaps the greatest potential for structural change. Migration is an important feature of labor markets in many low-income countries. Remittance flows are now the largest source of external funds in developing countries, exceeding total foreign aid flows and foreign direct investment (The World Bank, 2019). Only a handful of papers estimate the developmental impacts of seasonal and guest worker programs on origin countries, all of them outside of Africa, e.g. Gibson, McKenzie and Rohorua (2014), Gibson and McKenzie (2014), and Kosack (2019). The experience of Malawi suggests that managed circular labor migration that channels earnings back to sending communities may offer a practical tool for triggering longer-run structural change in communities where industrial, agricultural, and trade revolutions have been slow to arrive. There is still vast scope to learn more about how these tools should work in practice. For example, the changes we measure in Malawian labor markets occur after the end of international migration. As a result, our setting is not well designed to tell us whether the sectoral employment shifts in response to migrant money would have been even larger, or would not have occurred at all, if international migration had been allowed to continue. This remains a set of questions for future work.

Finally, our results provide new evidence on the triggers of structural transformation. Similar to Bustos, Garber, and Ponticelli (2020), we highlight the role that an increased supply of capital plays in facilitating structural change, although the increase in capital here is transitory, being driven by an increase in earnings from temporary labor migration (not from agricultural technology shocks as in Bustos, Garber, and Ponticelli (2020)), and we measure impacts over a longer horizon. To the best of our knowledge, the structural change literature has not yet considered international migration as a source of local capital that could facilitate reallocation of labor across sectors.[5]

2 Linking temporary capital injections with long-run changes in the labor market

Before presenting our evidence on how the international migration shock affected the Malawian economy over the long run, it is worth considering a more general question: How might a temporary injection of migrant money affect the types of work that people are able to do

[5]Foster and Rosenzweig (2008) model how national rural-to-urban labor migration and remittances facilitate structural change in rural labor markets.

in a rural economy?

First, and most directly, the return of migrant money increases local incomes of migrants and migrant households. The local demand shock generated by returning migrants increases the demand for farm- and non-farm goods. When preferences are non-homothetic, the demand for – and hence viability of – goods and services produced by the non-farm sector will increase faster than the demand for food (Herrendorf, Rogerson and Valentinyi, 2014; Gollin, Jedwab and Vollrath, 2016). Returns to work change across sectors, bidding up the price of non-farm relative to farm work, encouraging workers to move out of farming. Yet these changes are likely to be transitory, unless there are mechanisms for capital accumulation to occur in response to the temporary increase in migrant capital. This is where the following channels of impact become important.

In addition to boosting consumption in migrant households, migrant capital provides liquidity for investment in all sectors and across a broader range of households. In the farm sector, farmers may choose to invest in farm capital such as seeds, fertilizer, or farming equipment. If the economy is closed, as is the case for Malawian districts, more farm capital lets farmers meet minimum food production requirements with less labor, and excess labor can be released to the non-farm sector.[6] In the non-farm sector, migrant capital can enable individuals to overcome fixed start-up costs of opening businesses, and/or allow them to expand production in this higher productivity sector (as in Woodruff and Zeneto (2007) in Mexico, and Yang (2008) in the Philippines). Even if migrants are not starting businesses themselves, they may still act as local financial intermediaries that lend money to entrepreneurs, thereby allowing the non-farm sector to expand and pulling workers out of the farm sector.[7]

[6]Robinson (2016) shows wide variation in maize prices across markets in Malawi, suggesting it is reasonable to assume that districts in Malawi function like closed economies. Researchers have also documented large frictions in markets for land and other agricultural inputs (seeds, fertilizers) in Malawi (Beegle, Galasso and Goldberg, 2017). Investment in some types of farm capital may be less effective at triggering structural shifts in employment if other complementary inputs are hard to source.

[7]Buera, Kaboski and Shin (2013) show how more capital in the economy can be a mechanism for structural change by facilitating more entrepreneurial activity; Kaboski and Townsend (2012) provide evidence for this in Thailand. In Malawi, the almost complete lack of financial infrastructure means that most borrowing activity is local in scope. As late as 2014 (The World Bank, 2014), there were only 4.85 ATMS and 3.2 bank branches per 100,000 people respectively.

Investment of migrant earnings in human capital (education or health) can also occur. However, these changes feed into the labor market with a lag, with investments in education or health directly improving worker productivity of the next generation. If literate or healthier workers have higher returns in non-farm relative to farm sectors, an increase in the share of skilled workers could lead to shifts in the sectoral allocation of labor over time (e.g. as in Porzio, Rossi, and Santangelo (2021)).

Underlying each of these investment channels is the idea that there is insufficient capital in the local economy to break out of a poverty trap (Balboni, Bandiera, Burgess, Ghatak, Heil , 2021). Incomes are so low that savings cannot be sustained, and people cannot afford to buy much more than food. Increasing the supply of capital in the economy even temporarily can lead to a big push through some combination of the above channels: some households spend more on non-food items (services, physical capital, human capital investments); other households start up small businesses to supply this new demand. Importantly, in order for temporary increases in migrant capital to trigger long-term impacts on the structure of the labor market, the shock to migrant capital should translate into higher savings and investments at the level of the local labor market, not just at the level of the individual migrant. As savings and investment accumulate within and across migrant households, the impact of the initial income shock grows and becomes visible at district level over time.

In Section 7, we assemble evidence on these channels of persistence and discuss why migrant capital injections may be slow to effect change at the level of the local labor market.

3 Context: Rural labor markets in Malawi

Malawi was and still is a predominantly agricultural economy, although the importance of agriculture for work and output has changed over time. In the 1960s, agriculture accounted for half of Malawian GDP. By 2015, this contribution had shrunk to 30%. Employment in agriculture declined much more slowly from 84% in 1977 to 61% by 2008. The share of manufacturing in GDP fell from an already low 13% in the 1960s to 10% in 2015, with employment in this sector shifting from 5% in 1977 to 9% in 2008. Over five decades, the share of services in GDP increased from around 40% to almost 55% (The World Bank, 2016) with the share of employment in the sector rising three-fold from 8% in 1977 to 24% in 2008. These shifting patterns of employment from farm to non-farm sector, with non-farm jobs concentrated in services rather than manufacturing, strongly resemble patterns of structural change in other African labor markets over the last three decades (e.g. Gollin, Jedwab and Vollrath (2016) and McMillan, Rodrik and Verduzco-Gallo (2014)).

Table 1 illustrates these broad patterns of change for men and women separately. Between 1977 and 2008, the share of workers in the Malawian agricultural sector fell from 94% to 70% for women and from 76% to 53% for men. Over the same period, the share working in services increased dramatically. In 1977, 2.8% of women and 12% of men were in service sector jobs and by 2008, these shares had risen to 21% for women, and 28% for men. A Herfindahl index of sectoral employment concentration shows a fall from 0.89 to 0.55 for women's employment and from 0.62 to 0.36 for men's employment signifying a decline in employment concentration over time.

It is worth noting that labor force participation rates for men and women are high in Malawi. Between 84% and 96% of working adult adults were working or looking for work in different decades. Nonetheless, there is good evidence of substantial underemployment, and excess labor supply in rural Malawi (for example, Goldberg (2016) and Dillon, Brummund, and Mwabu (2019)).

To fix ideas about the nature of non-farm work in rural Malawi, we graph the percentages of workers employed in non-farm industries and occupations, for each of the most prevalent non-farm categories of work. Figure 1a shows that retail and wholesale trade tops the list of non-farm industries: 45% of women and 32% of men work in this sector, with the next largest categories being the public sector (e.g. teachers, medical staff) and hotels and restaurants for women, and manufacturing, construction, and the public sector (including defense) for men.

Smaller shares of women work in manufacturing, transport and communications, and other services, and 6% of men work in business services or transportation/communications/storage. Occupational patterns in Figure 1b follow industry patterns in Figure 1a. Almost 30% of women working in the non-farm sector are working proprietors (self-employed). 17% are sales and shop assistants, 16% are teachers or nurses, and 6% are food and beverage producers. About 18% of men report being working proprietors and 11% are shop assistants; 19% are teachers, nurses, or work in security while 12% work as brick-makers or carpenters. The non-farm sector is dominated by work in retail trade, the public sector, and small-scale personal and general services, often run by the self-employed.

We note several facts about how services are produced in Malawi. First, small businesses are common in rural areas. While all rural households farm, data from the mid-1990s shows that around one in five households also owns a non-farm enterprise (Alvarez-Cuadrado, Amodio, and Poschke, 2021). Second, although many of the service sector jobs documented in Table 1 and Figure 1 are in small-scale household enterprises, these activities tend to use more capital on average than farming does; farming in Malawi is not at all capital-intensive (Chen, Restuccia, and Santaeulalia-Llopis, 2017). Third, labor tends to be more productive on average in the non-farm sector than in farming (Gollin, Lagakos and Waugh, 2014). Even though the types of jobs that people hold in the non-farm sector are low-skilled, the average worker in this sector may still be more productive in such a job relative to the farm sector.

With these stylized facts about Malawi's rural economy in hand, we next describe the shocks to international labor migration.

4 The natural experiment: Temporary migration shocks and migrant capital

4.1 Historical circumstances of the natural experiment

Off and on during the 20th century, mine work in South Africa provided a feasible sector of work for Malawian men willing to migrate. These work opportunities were organized between the Malawian government, whether colonial or independent, and the South African mine labor recruiting agency, the Witwatersrand Native Labour Association (*Wenela*). Workers were employed on fixed two-year contracts at set rates that were higher than any local wage work, which itself was limited. Selection into mine work on the basis of observable characteristics was limited to minimum age and weight requirements. There were no education or skill requirements and about 60% of the Malawian miners in South Africa had no education at all. Most miners were between the ages of 20 and 30 (author calculations, South African Census data 1970 and 1980). Prior to 1967, legal circular migration was limited by national recruiting quotas that were held at less than 2% of the working age male population.

Then in 1967, a new labor treaty (Treaty, 1967) removed the old quota and migration expanded from 40,000 to over 120,000 men in five years, as seen in Figure 1. This short-lived surge in migration ended abruptly in April 1974, when a *Wenela* plane returning to Malawi crashed, killing all miners aboard. Then-president Banda banned all labor recruiting in Malawi and recalled all miners from South Africa. The number of Malawians working on South African mines fell to zero in the two years following 1974. By 1977, all miners were back home in Malawi.[8]

Between the labor treaty (1967) and the labor ban (April 1974), the prevalence of mine work expanded dramatically across all districts, facilitated by low entry barriers and wages

[8]By 1977, Banda had realized that mining capital was a crucial source of foreign reserves for the country and rescinded the ban on migration. Migrant flows never returned to prior levels because by the 1980s, *Wenela* had redirected the bulk of recruiting towards the South African labor market. See Mariotti (2015) for an analysis of how this shock affected outcomes in South Africa.

that promised 2.5 times more than the next best wage-earning job at home.[9] By the late 1970s, approximately one in three Malawian adult men had ever worked as an international migrant (see Appendix Table 1). Crucially for our paper, because these miners were contractually obliged to receive 60% of their earnings deferred pay upon repatriation, the majority of migrant earnings returned to Malawi. This deferred pay formed the basis of their ability to accumulate capital for use in rural sending regions.[10]

Flows of money paralleled the ramp up of migration in the late 1960s and the equally dramatic decline of migration in the mid-1970s. Between 1966 and 1975, total migrant capital flows rose by about 20%, driven by increases in migration and by increases in the mining wage offered in South Africa (see Figure 1).[11] Total deferred pay inflows over the entire period were 53 million USD. At peak migration in early 1973, Malawi received 2.75 million USD from miner earnings in a single month, or almost 115,000 USD on average per district. Each migrant returning from a completed two-year contract would have received between 130 USD and 295 USD, depending on when he left. As a benchmark, average per capita GDP in 1970 in Malawi was only 63 USD.

Figure 1 shows that money flows spiked after the plane crash (indicated by the second vertical line), when all miners were repatriated. This spike lasted just under two years. Depending on how one discounts, between one fifth and one quarter of the total amount of migrant capital from the entire migration episode returned to Malawi in the 19 months after the plane crash. The latter part of the migration period (1974-1975) represents the period of largest, coordinated capital flows back to rural districts in Malawi. Because not all miners would have completed their two-year contract at the time of the plane crash, the average amount of money per migrant returned to the country after April 1974 was 87.6 Kwacha, or 73 USD per migrant.

[9] These wage jobs were typically in farm labor on agricultural estates growing tobacco or sugarcane.

[10] We refer the reader to our prior paper Dinkelman and Mariotti (2016) for a more detailed discussion of the history of organized legal mine migration from Malawi to South Africa.

[11] Figure A.1 in the Online Appendix shows the trend in miner real and nominal wage rates in Malawian Kwacha between 1966 and 1975 alongside the global gold price in USD, the driver of increased wages.

4.2 Exploiting the variation in capital per migrant generated by the natural experiment

Capital flows per worker through the Treaty and ban period fluctuated across districts, we argue, somewhat randomly. The total amount of capital per migrant that flowed back to a district was driven by the number of migrants contracting at specific wage rates (with higher wages being paid over time) and also, at the end of the Treaty period, by how much of his contract each miner had completed upon repatriation (with those having completed more of their contracts having more paid out). This variation in capital per returned migrant driven by timing differences in contracts is what we exploit to identify effects on sectoral employment shares in the labor market.

One possible approach would be to use the cross-district variation in migrant capital per worker received between 1967 and 1975. However, such comparisons could be confounded by unobserved factors that drive different levels and time profiles of migration across districts. Assuaging concerns about such potential confounders would require data to show a lack of pre-existing differential trends in labor market variables across high and low migrant capital per worker districts before 1967. Unfortunately, the British colonial Census instruments of the 1940s and 1950s and the first independence-era Census of the 1960s failed to collect sector of work data (nationally or at district-level) for any African workers.

An alternative approach, and the one we adopt here, is to focus only on cross-district variation in migrant capital per worker (K/L) induced by the timing of the plane crash – we call this the *K/L shock* variation. This variation is appealing because the plane crash is clearly an exogenous shock and allows us to pinpoint a much finer source of variation in K/L per district. Remittances triggered as a result of the plane crash would vary at district-level only because workers contracted at different wages (whether in 1972, 73 or 74) and started at only slightly different times, leading to different total months of contract completion at the time of the crash. For example, a district with more migrants closer to the end of their 24 month work period would experience a larger flow of money back; and a district with more workers who had started slightly later could also receive higher flows because of the higher wage. The differences in the composition of these contracts around the time of the plane crash is likely to be random. It is this variation, the capital per worker returned after the ban (*K/L shock*) that forms the basis of our identification strategy.

Table 2 examines the district-level correlates of the *K/L shock*, and of *K/L* received over the entire period. We regress each of these outcome measures on historical and geographic district characteristics, measured prior to the 1967 *Wenela* treaty. These characteristics include whether the district is in a malaria area or not, the log of population density in 1945, the youth literacy rate in 1945, whether the district has any agricultural estates (offering an alternative source of paid employment), the share of males (or females) not earning a wage in 1966, the region the district is in, and the historical number of *Wenela* recruiting stations in the district. The first two columns of Table 2 report the results for *K/L shock* – capital per worker received after the plane crash – while the last two columns report the results for capital per worker over the entire treaty period.

It is worth noting that our cross-sectional variation comes from only 24 districts. In Table 2, we report heteroskedasticity-robust standard errors and statistical significance using p-values from the small sample t distribution. We also report wild bootstrapped p-values in square brackets.[12] In columns (2) and (4) of the table, we report the result of the joint test of all slopes equal to zero excluding recruiting stations, because we do not use recruiting stations in our empirical specification.

In columns (1) and (3), we see that districts in the South, with a lower probability of malaria, and with a higher share of men without a cash wage are predicted to have higher *K/L shock* and *K/L* when we use regular robust standard errors, but not when we compute wild bootstrap p-values. We also cannot reject that the set of covariates together have no significant correlation with *K/L shock*.

In Dinkelman and Mariotti (2016), we argue that the prevalence of a recruiting station in

[12]Wild bootstrap p values are appropriate here because we have only one observation per district, and the number of districts is small (Roodman et al, 2019; Liu, 1988; Davidson and MacKinnon, 2010).

a district is a key determinant of migration, and that these stations were as good as randomly allocated to districts. Columns (2) and (4) of Table 2 show the importance of the number of recruiting stations in predicting both *K/L shock* and *K/L*. The prevalence of recruiting stations explains an additional 30% of the variation in the *K/L shock*, and of *K/L*.

Given the arguably random locations of recruiting stations that were established in the early 1920s across the country it is tempting to consider using them to instrument for *K/L shock*. However, Table 2 shows us that recruiting stations strongly predict both *K/L* and *K/L shock*. Since we focus on the effects of more migrant capital per worker returning after the exogenously-timed plane crash, the remainder of the paper proceeds without a consideration of recruiting station prevalence.

5 Empirical strategy and Data

5.1 Estimation

To isolate the persistent effects of migrant capital per worker at market-level, we specify the following empirical model for labor market outcomes Y_{dt}:

$$Y_{dt} = \sum_t \alpha_t \frac{K_d}{L_d} Shock * Decade_t + \kappa_t + \delta_d + W_d Trend_t \lambda + \epsilon_{dt} \tag{1}$$

where Y_{dt} is the share of economically active adults in agriculture, manufacturing or services, d is the district, t is the decade (1977, 1987, 1998 or 2008), $\frac{K_d}{L_d} shock$ is the amount of deferred pay per returned migrant received by district d between 1974 and 1975, $Decade_t$ is a set of decade dummies for one, two and three decades after the end of migration (1987, 1998 and 2008 respectively), κ_t is a decade fixed effect, δ_d is a district fixed effect, $W_d Trend_t$ is a vector of baseline district covariates interacted with a linear trend term, and ϵ_{dt} is an idiosyncratic error term. Regressions are estimated separately for men and women.

Equation (1) allows us to estimate the effect of the capital shocks per returned migrant on differential changes (rather than level differences) in employment outcomes across districts after 1977. δ_d controls for constant average differences in labor markets across districts (e.g. districts with lake access are always able to support work in fishing industries). These controls also standardize for district population size. κ_t controls for aggregate changes in the labor market that affect all workers equally, for example, a nationwide drought that occurred in the early 1990s and which likely affected agriculture everywhere. Trend interactions flexibly allow districts with different initial population densities, literacy rates, marriage rates, malaria risk and potential for cash wages to evolve differently over time.

Our identification assumption is that conditional on these district and decade fixed effects, and controlling for differential trends related to baseline observables, districts receiving more or less capital per worker after the plane crash would have evolved similarly in the absence of this injection of capital.

Both sign and significance of the α_t parameters are informative. α_{1987} tells us the percentage point change in the relevant employment outcome (e.g. share in farming) between 1977 and 1987, the first decade after the plane crash. α_{1998} and α_{2008} provide the same parameter estimates for later decades. For the migrant capital injection to have had any effect on the local economy, we should consider α_{1987}. To look for evidence of persistence of the initial shock over time, we should see non-zero estimates of α_{1998} and α_{2008}.

In addition to employment outcomes, we examine what happens to log population and urbanization rates within districts in the wake of the capital shock. These secondary outcomes are common proxies for economic development. Because we can measure these outcomes in earlier years of Census data (back to 1945 for population density and back to 1966 for urban share of the district), we can estimate equation 1 using a difference-in-differences design. The analysis of population and urbanization outcomes has the added benefit of demonstrating that there were no differences in pre-shock trends in the economies of high and low *K/L shock* districts. These tests play an important role in building the case that our main results are not simply picking up changes across districts that would have occurred in the absence of the migrant capital inflows.

A potential threat to validity is that some other shock to local labor markets occurs in high *K/L shock* districts in years after the plane crash. Our reading of the economic history literature suggests that a key candidate confounder is the tobacco liberalization which occurred in the early 1990s, and allowed subsistence farmers to start producing tobacco for export. We find no evidence that including controls for tobacco suitability of the district interacted with decade dummies changes our main findings for the effect of $\frac{K_d}{L_d} shock$ (see Online Appendix Table A.5).

5.2 Data and summary statistics

We measure labor market outcomes, population outcomes, and district-level covariates using six waves of Census data from 1945 to 2008. We digitize historical Census data available at the district-gender level from 1945, 1966 and 1977 and match this with Census data from 1987 (the 10% sample), 1998 (100% sample) and 2008 (10% sample). Details of variable construction are in the Online Data Appendix.

Key labor market variables are defined for men and women using labor market questions that are generally the same across survey instruments. Broad industry of work is available for all economically active individuals 10 years and above. We use these broad measures of industry – agricultural, manufacturing and services – to look at labor reallocation across sectors, as well as a finer breakdown of industries in the non-farm sector: general manufacturing, retail, transport and communication, and all other services, including personal services and government employment. Using the broader measures of industrial sector, we construct a Herfindahl index that captures the diversity of employment within the district. Data on total population and population by gender are available for each district from 1945 onwards and on urban shares of population from 1966 onwards.

We also collect and digitize material from *Wenela* administrative records showing district-month level remittance flows from October 1967 to November 1975. Three categories of monies were recorded by the mine labor recruiting agency: deferred pay, remittances and other deposits. 89% of all monies returning to Malawi were in the form of compulsory deferred pay; this is the measure of migrant capital we use in our analysis. Amounts are aggregated over time to the district-level.

We construct estimates of the number of migrants returning to each district in different years by combining information in two different variables from the 1977 Census data. We estimate the district-specific number of migrants for the entire migration period (1967-1977) by multiplying the district-specific number of men who report ever migrating at all in the 1977 Census by the national share of these ever migrants who report returning to the country between 1967 and 1977.[13] To estimate the number of migrants returning to each district after the plane crash, between 1974 and 1977, we make a similar calculation: multiplying the district-specific number of men who report ever migrating at all in the 1977 Census by the national share of these ever migrants who report returning to the country between 1974 and 1977. Figure 1 shows that the number of returnees peaked in 1974 and 1975, just after the plane crash.

Means and standard deviations for the main covariates in our analysis dataset are in Online Appendix Table A.1. An important takeaway from this table is that international migration was common in Malawi. By 1977, on average one in three adult men in a district had worked abroad. However, only between 2.5 and 9% of men in each district were actually absent from Malawi at the time of the plane crash. This was because contracts were time-limited to two years, and migrants were forced to return home at the end of the contract.

Focusing on the post-plane crash variation in *K/L shock*, we see in Online Appendix Table A.1 that the average district had 87.6 Kwacha (the equivalent of 73 USD) returning per migrant worker after the plane crash. This represents more than one year's worth of earnings as measured by GDP per capita at the time.

[13]Exact year of return is only recorded at national level in 1977.

6 Main results

6.1 Impacts of *K/L shock* on broad sector of work

Table 3 presents our main results for the broad categories of employment in agriculture, manufacturing and services, and the Herfindahl measure of employment diversification. Panel A shows estimates for women, Panel B for men. The unit of observation in each regression is the district-gender-decade. For each outcome, we present estimates of α_t from equation (1), including all controls. Regressions are weighted using 1977 population weights. Robust standard errors are clustered at district level and we report statistical significance using p-values from the small sample t distribution to account for the small number of districts. Following Young (2019), who argues that p-values generated through randomization inference are more reliable than inference based on robust clustered standard errors in situations with few clusters, we also show p-values generated using randomization inference in square brackets. In all tables, we evaluate the coefficients at the mean amount of capital per returned migrant for the post-plane crash period (multiplying $\hat{\alpha}$ by 87.6, which translates into a mean of *K/L shock* of 73 USD).

Table 3 shows that districts with larger capital flows per returned migrant experienced more structural change away from agriculture into the non-farm sector, with women and (to a lesser extent) men shifting into services and some decline in industrial concentration of employment for women. For all sectors of work in the table, point estimates are larger and more significant for women.

Specifically, for each additional 73 USD per worker that flowed back to a district by 1977, the share of women working in agriculture fell by 0.52 percentage points in the first decade following the shock (87.6*0.00006*100), by 2.4 percentage points in the next decade, and by 2.8 percentage points by the third decade after the shock.[14] Column (2) shows smaller shifts of women into manufacturing (0.4 percentage points or lower) and column (3)

[14] To interpret the coefficient estimates in equation (1) as percentage point changes in the employment outcomes, we multiply each point estimate by the mean *K/L shock* (87.6) and by 100.

shows larger shifts into services (0.2-3.1 percentage points by three decades after the shock). The agricultural and service sector employment results for women are significant even using randomization inference p-values and especially in the twenty and thirty years after the end of the migration episode. Women's employment concentration also declines significantly.

Panel B shows similar patterns of structural change for men, although these results are smaller and not consistently significant using either regular or randomization inference p-values. The share of men working in agriculture fell by 0.7-1.2 percentage points in the first two decades after the shock (in a district receiving the mean value of migrant capital per worker), and rose by 0.48 percentage points in the third decade. Shifts into the service sector were positive. Male employment in services increased by a significant 1.13 percentage points in 1998, although the effects in the first and third decades were more muted at 0.14-0.17 percentage points.

In Online Appendix Table A.3, we show that our main results – especially for women – are stronger when we omit the baseline controls interacted with trend terms, and point estimates are similar (and larger) when we exclude the district containing the capital, Lilongwe. Our results are also robust to including controls for the number of migrants returning 1967-1974. Online Appendix Table A.4 shows that our results are also robust to alternative specifications of the treatment variable that use the full migrant capital inflow between 1967 and 1977, and to rescaling the capital shock by baseline district-level population in 1966, instead of by the number of migrant workers.

Overall, the employment shifts in response to the capital inflows are positive, somewhat lagged, and persistent. They suggest some measure of structural change, especially for women's work, facilitated by exposure to labor migration opportunities. We return to a discussion of the lagged and persistent nature of the employment impacts in Section 7.

We can use Table 3 to construct a back-of-the envelope estimate of how much migrant capital contributed to overall structural change over time. In the four decades before 2008, employment shares in agriculture fell by around 24 percentage points for women, from 94%

to 70% (Table 1). Summing up our estimates of the percentage points from Table 3, we estimate that the migrant capital shock in an average district accounted for about 25% of the structural reallocation of female labor out of farming and into non-farm work. The contribution of migrant capital to the reallocation of male labor across sectors is smaller, at around 11%.[15]

6.2 Impacts of *K/L shock* on narrow sector of work

What types of non-farm work developed in high capital shock districts, in the thirty years after this migration episode? Table 4 presents a finer breakdown of sector of work for the non-farm sector: general manufacturing and construction, general services, retail, and transport and communications. General services include personal services (e.g. guards, domestic workers and cooks), business services (advertising, or insurance, banks and engineers, legal services, accountants) and other services (e.g. barbers, tailors, typists, public sector workers).[16] The retail sector includes wholesale and retail trade of food, fuel and other goods, hotels and restaurants, car repairs etc. Transport includes transport of goods and/or people, including using buses, taxis, boats, bikes, warehousing, and telecommunications. All specifications follow the form of equation (1).

For women, we see the most significant and sustained impacts on employment in the retail and transport and communications sectors. For each 73 USD per returned migrant received, the share of women working in retail rose by between 1.13 and 2.27 percentage points. The joint test for the shift of women into retail can reject zero at conventional levels of significance. The share of women working in transport and communications also rose, by a much smaller extent, in the first two decades following the capital injection, and the share in construction rose by about 0.19 percentage points in the third decade.

[15] An average district had 58,000 women and 57,000 men. The percentage increases in employment in services translate into a 5 percentage point increase in employment among 58,000 women, or 2,900 more jobs. For men, the equivalent impact is 821 jobs.

[16] We omit mining as a separate category since shares working in the local mining sector are so low. Most of those in business services classify themselves as working proprietors.

For men, more capital in the district shifted work out of manufacturing (0.34 - 0.67 percentage points) and transport (0 - 0.4 percentage points) and into construction (0.37 - 1.04 percentage points) and retail (0.35 - 1.5 percentage points). The largest relative shifts for men were towards construction and retail.

6.3 Impacts of *K/L shock* on type of worker

Table 5 investigates whether the *K/L shock* affects the types of workers and work arrangements found in local labor markets. We estimate equation (1) for outcomes describing the share of workers (in all sectors, columns (1)-(3), or in the service sector only, columns (4)-(6)) that are self-employed, working without pay in family businesses or farms, or working for a wage. Panel A shows how access to migrant capital affected the employment situation of women, Panel B for men. For these outcomes, we can only examine changes between 1987 and 2008 due to changes in the way this information was captured in other years. There are 72 observations in each regression.

In districts receiving more migrant capital, men and women shift towards working on family farms and in family businesses (the data do not allow us to break up these categories). Women shift out of self-employment: this likely reflects the move out of agriculture, since many farmers report being self-employed. In columns (4) through (6), we see that conditional on working in the service sector at all, districts with more capital inflows three decades after the shock have a smaller share of (male and female) wage workers, and at the same time, a larger share of women working in family businesses. These patterns are consistent with the capital injections from returning migrants enabling households to start up new businesses, and shift some of their labor into off-farm activities.

6.4 Impacts of *K/L shock* on population growth and urbanization

Next, we investigate how migrant capital inflows affected population growth and urbanization within districts. We estimate versions of equation (1) using a set of population variables P_{dt}

measured at district-year-age group level. We group the population into children (under age 5), youth (ages 5-18), and adults (ages 19 and older). For urbanization outcomes, the unit of observation is the district-year cell. t now includes six years of Census data from 1945 to 2008 for population outcomes and five years of data (excluding 1945) for the urbanization measure. Regressions are unweighted, all controls (except population density for the population outcomes) are included, and standard errors are estimated as before (clustered at district-level, and also generated through randomization inference procedures).

Figures 3a and 3b plot estimates of α_t, the relationship between the amount of migrant capital per worker received by each district between 1974 and 1975 and district-level population growth and urbanization outcomes before and after the migration surge. Standard error bars are included, and the omitted category for population outcomes is 1945, and 1966 for the urbanization outcomes. Each point on the line represents the marginal impact of receiving 73 USD of migrant capital between 1974 and 1975 on the log of the age-group specific population counts in the district in each Census year (Figure 3a) or on the share of the population living in urban areas in the district (Figure 3b). The regression table underlying these figures is in Online Appendix Table A.2.

The figure shows that districts that were going to receive 73 USD in migrant capital between 1974 and 1975 did not look significantly different (population-wise) compared with districts that were about to receive less migrant capital, prior to 1966. Between 1966 and 1977, this pattern diverges. Districts receiving more migrant capital per worker by 1977 experience significant population growth by 1977. This growth occurs immediately, is sustained over the next decades, and is not just mechanically picking up returning migrants. The largest significant effects are seen for children under age 5, perhaps suggestive of a small baby boom post-migration shock. Thirty years after the labor migration episode, total population in the high *K/L shock* districts had increased by around 4.3%.

In Figure 3b, we see there was also a spatial reorganization of population within districts in the wake of the shock, where the comparison year is 1966. By 1977, districts that received

73 USD more per worker than other districts had 1.1 percentage points more of their population living in an urban area. This gap in the share of population living in urban areas persists, and is significant until 1998. Internal population rearrangements within districts are likely to contribute to the structural change in these rural labor markets.

7 Understanding the timing and persistence of accumulation

The results in Section 6 demonstrate that the *K/L shock* impacted the structure of work in rural labor markets, but this impact took time to be felt in any magnitude, with muted effects in the first decade after the shock. Although population growth effects appear immediately after the shock, the shifts in employment across sectors take more time to manifest.

Section 2 discussed theoretical channels through which temporary migrant capital injections might trigger sectoral shifts in employment shares, and why district-level shifts may only be seen over the long-run. The three main channels of impact included: (1) an initial positive demand shock increasing the consumption of farm and non-farm goods; (2) an increase in the supply of liquidity facilitating farm and non-farm investments; and (3) an increase in human capital investments. These three channels may not occur at all, may occur independently, or may be related, in that the initial consumption boost may have an immediate impact on demand that becomes a persistent effect when coupled with the second and third channels. The third channel additionally complements the second channel in the future, when the next generation of workers enters the labor market.

In this section, we link our results from section 6 to the hypothetical channels of impact in section 2 and show that the long-run results are consistent with several of the mechanisms outlined in section 2. Throughout most of this section, we estimate regressions that take the form of equation (1). For the outcomes considered in this section, we have the added benefit of using pre-shock data on outcomes so that we can estimate more standard difference-in-differences type specifications.

7.1 Evidence on increased demand for farm and non-farm goods and physical capital investments

Ideally we would have data to show short-run purchases of food and non-farm goods increasing in response to the *K/L shock*. Unfortunately, in the absence of annual data on any of these outcomes, we can only observe outcomes for non-food outcomes from 1987 onwards. However, using these longer-run outcomes, we can test for persistent changes in the consumption of farm and non-farm goods.

In Table 6 we examine whether there is an increase in purchases of both farm and non-farm goods. Outcomes for Panel A, farm goods, are from the National Sample Survey of Agriculture in 1968 and the National Household Income and Expenditure Survey in 1998. We weight up to district-level using sample weights (see Data Appendix for details). Outcomes for Panel B, non-farm goods, are from the national Census in various years. In both panels, we estimate how farm and non-farm goods' ownership changes in districts exposed to high versus low *K/L shock*, in the decades after the plane crash, controlling for differences in ownership across these districts in the decades before the labor ban.

Panel A presents estimates of the effect of exposure to the migrant capital shock on productive farm asset investments. We measure the average number of hoes and pangas per household before, and twenty years after, the migration shock; and do the same for the share of households owning any cattle. Oxcart ownership can only be measured in 1987, 1998 and 2008.

Districts receiving more migrant capital per worker do not seem to be strongly investing in farm-specific capital in the longer run. Although the coefficients on *K/L shock* in 1987 are negative for hoes and pangas, these estimates are not statistically significant using the randomization inference p-values.[17] We do not place much weight on the differential decline

[17] One caveat to our results is that data on inputs like fertilizer, hybrid seeds, or type of crop planted

in oxcart use in high *K/L shock* districts, since this effect is so small.

In contrast to these results on farm-specific capital, there is some evidence that non-farm specific capital increased, post-migration shock. In Table 6, Panel B, we examine changes in (district-level) ownership of assets that are used in non-farm work, or for household consumption. We measure the share of households in the district in a given year that own a radio, have a household with durable walls, a durable roof, and/or both durable walls and a durable roof. Panel B indicates that ten years after the plane crash, districts receiving more migrant capital per worker experience significant increases in the share of households with durable walls and roofs, and durable roofs alone. The share of houses constructed with better quality materials increased by 4.8 percentage points (0.000555*87.6), or by almost 40%, in the decade following migrant capital inflows.

With little evidence of persistent changes to the ownership of farm inputs in Table 6, it is not likely that the temporary migrant capital shock affected the way farming is conducted in Malawi. The agricultural sector in high *K/L shock* districts does not seem to have undergone any transformation in the production process, and increases in demand for farm investment goods are probably not the main drivers of employment shifts. The results in Table 6 do indicate that ten years after the shock, there is a meaningful change in housing quality that may well have resulted from an increase in the demand for non-farm goods. Improved housing represents increased consumption, and as well as increased investment in premises required to provide non-farm goods and services (e.g. spaza shops, village bars and food stalls, hairdressing premises etc). Any positive demand shock triggered by the migrant capital inflows post-plane crash would likely have been for non-farm goods and services, and would have led to increased investment in these non-farm sectors.

We complement the analysis of asset ownership in the Census data with the 1992 Malawi Demographic Health Survey (DHS) which has a more comprehensive set of household assets. If our hypothesis about the migrant capital shock relaxing liquidity constraints is correct,

where we may see a longer term impact are not available prior to the early 2000s.

then we should expect to see increases in ownership across a wide range of assets, some of which can arguably be said to help with production in a household business, while others may generate more consumption benefits.

The 1992 Malawi DHS is a cross-sectional dataset, so we estimate the cross-sectional correlation between measures of household assets in 1992 and the *K/L shock* received by each district in the mid-1970s. We ask whether districts that received more capital after the migration shock exhibited more physical asset ownership across a wider range of assets.

We estimate the following for households h in district d:

$$HHAssets_{hd} = \gamma_0 + \gamma_1 \frac{K_d}{L_d} shock + W_d \sigma + \mu_r + \omega_{hd} \quad (2)$$

where the W_d variables are as before. Since this data is at the cross sectional level we cannot include district fixed effects. We must assume that baseline district-level controls and region fixed effects (μ_r) adequately account for any differences in initial wealth conditions across districts. We also control for the number of migrants returning to each district prior to 1974 (in the 1967-1973 window), to compare districts with the same migration "propensity". ω_{hd} an idiosyncratic error term. Our outcomes $HHAssets_{hd}$ include the DHS wealth index (units are in standard deviations), the count of household-level assets and indicators capturing ownership of specific assets like electricity, radios and cars.[18]

Table 7 column (1) shows clearly that districts receiving more *K/L shock* by 1977 have higher asset holdings 15 years later, as measured by a higher asset index, more holdings of durable houses and roofs, greater access to electricity, and higher radio and car ownership. This evidence is consistent with increases in investment over time contributing to the long-run structural transformation that we see. These results indicate that changes in household fortunes occur after a shorter time period than the 20 years for which we see some of the main results in the Census data. Beyond offering more support to our main employment

[18] The DHS wealth index is an index of assets constructed used principal components analysis. For more detail on how the variable is constructed, see `https://dhsprogram.com/programming/wealth%20index/DHS_Wealth_Index_Files.pdf`

findings, this table helps us understand that some of the “delay” in employment results is a function of not having outcomes data in the intervening Census years.

The DHS data allow us to further investigate whether all of the gains from the migrant capital shock are limited to the older generation in 1992; those who would have been adults at the time of the shock. This investigation helps us understand the ways in which the migrant capital shock persisted. In Table 7 columns (2) and (3), we split the results from (2) into household heads that were adults at the start of the 1967 Treaty and those who would have been children, or not yet born, at the time of the migration shock.

Columns (2) and (3) illustrate that both age groups enjoyed higher asset ownership in high *K/L shock* districts. Specifically, 15 years after the shock, the older household heads from higher *K/L shock* districts, have more assets than older household heads in lower capital districts. The same is true for the younger household heads in high versus low *K/L shock* districts. The results in these two columns imply that the migration shock not only shifted employment shares but also had tangible and persistent effects on household wealth.

The fact that older household heads enjoy higher asset ownership when they are in high versus low *K/L shock* districts gives us direct evidence of physical capital accumulation that is independent of the next generation's potential human capital accumulation. Column (2) of Table 7 indicates that one of the channels driving structural change in Malawi after the shock is indeed physical capital investment, unrelated to educational gain. However, the fact that younger household heads also enjoyed higher asset ownership suggests that human capital accumulation may have also played a role in shifting employment across sectors.

7.2 Evidence on human capital accumulation

One of our key findings in Tables 3, 4, and 5 was that the sectoral employment shifts post-plane crash took some time to manifest at district-level. One channel that could contribute to a delayed response is a human capital channel. Only the very oldest of those age eligible to receive education during the shock would have entered the labor force by 1987. But by

1998 and 2008, more of those whose educational attainment might have benefited from the shock would have entered the labor force.

In a final piece of analysis, we investigate whether the increase in district-level *K/L shock* led to improvements in educational attainment. We know from prior work (Dinkelman and Mariotti (2016)) that districts more exposed to the total migration shock 1967-1977 invested more in the schooling of those who were children at the time of the migration shocks.

In Table 8, we relate education profiles of different districts and cohorts to the inflow of post-plane crash migrant money per worker. We focus on total accumulated education of adults aged 20 to 65 in the 1998 census, some of whom would have been eligible for schooling during parts of the migration surge. We ask whether mean levels of human capital are higher for cohorts living in districts that received the largest *K/L shock* during their years of primary school eligibility.

We estimate the following regression for the average education outcomes of cohort c in district d ($Educ_{cd}$), separately by gender:

$$Educ_{cd} = \gamma_0 + \gamma_1 \frac{K_d}{L_d} * Shock_c + \gamma_2 \frac{K_d}{L_d} * Postshock_c + \phi_c + \mu_d + W_d Trend_c \sigma + \omega_{cd} \quad (3)$$

We exploit an additional piece of time variation (related to timing of birth) to check whether cohorts eligible for primary schooling in the narrow 1974-1977 period ($Shock_c$ cohorts) and for those eligible right after the end of the shock, 1977-1980 ($Postshock_c$ cohorts), had more education in the long run if they were in high capital inflow districts.[19] Individuals are grouped into five-year age bins. Primary school age eligibility in Malawi runs from age 6 to age 15. Our control cohorts (omitted cohort category) are those eligible for primary schooling before 1974. All other controls in W_d are the same as in equation (1), including district fixed effects, μ_d.[20].

[19] Effects on education found in the post-shock cohort would lend support to our claim that the migrant capital shock did have immediate impacts, but that we just can't measure them in any other data other than in long-run education outcomes. See Dinkelman and Mariotti (2016) for a longer discussion of why migrant earnings would have been important for relaxing liquidity constraints for parents sending children to school.

[20] The trend term here is a series of cohort level dummies, hence the notation $Trend_c$

Results in Table 8 show that for each additional 73 USD received by returning migrants in a district, female education rose by a significant 0.17 years for the shock-era cohorts (87.6*0.00195), and 0.16 years in post-shock cohorts (87.6*0.00176). Male education rose by a smaller and non-significant 0.1 years and 0.06 years in the same cohorts. Each of these effects is relative to the control cohort of older adults ineligible for schooling at the time of the labor ban.[21] The average effect of migrant capital on female level of education is an 8 - 9% increase (significant), while for males the increase is 1.6 - 3% (insignificant). Exposure to more migrant capital also raises female enrollment in primary school by 3.7 - 4%, with negligible changes in enrollment for males.

The results of Table 8 are important for several reasons. First, together with the results in Tables 6 and 7, they indicate that the migrant capital coming back to Malawi in the wake of the plane crash was not just consumed, but invested in assets. Second, the timing of schooling investments means that individuals who gained the most education in response to more *K/L shock* (including younger females) would only have entered the work force in early adulthood around the mid to late 1990s. This is exactly the decade in which we start to see larger and significant impacts of the *K/L shock* on sectoral employment shifts, and more so for women than men. Finally, it is likely that the human capital investments triggered by migrant capital contributed to the structural shifts of labor out of agriculture and towards services. This is in line with other micro-economic evidence that suggests that people with more schooling are more likely to work off-farm relative to on-farm (e.g. Huffman (1980) for the historical US, Yang (1997) for China and Fafchamps and Quisumbing (1999) for Pakistan).

[21]Some of the control cohorts would have been eligible for schooling in the 1967-1973 window. This means our estimates in Table 8 likely underestimate the impact of total migrant capital per worker on total educational attainment at district level.

7.3 Discussion

The three sets of results on mechanisms presented in this section offer support to our accumulation hypothesis and explain why the sectoral employment shifts take time to gain traction at the level of the district, and why the initial impact of the remitted migrant capital is increasing over time.

First there are small direct impacts on the consumption of non-farm goods (radios, household construction). This is consistent with our claim that under non-homothetic preferences, more income will translate into increased demand for non-food items. This increase in demand may cause some of the initial shift we see in employment shares.

We claimed in Section 2 that for this shift in demand to be permanent we need workers in the market positioned to take advantage of the demand shock through access to capital that they either lend out or borrow to invest in their businesses. We see evidence of an increase in physical asset ownership that is consistent with this second channel fostering accumulation. Twenty years after the shock, the employment share effect has grown as a result of both the initial trigger to demand as well as the subsequent increased investment.

While these initial liquidity constraints were being resolved, workers were also investing in the human capital of the next generation. This generation would have joined the labor market anywhere between 10 and 25 years after the shock and contributed to the growth in the effect sizes we see two and three decades after the shock. The larger human capital responses of female education are consistent with the larger shifts we see in female employment out of agriculture and into services in the last two decades after the shock and may go some way to explaining the greater structural changes in women's work than for men over time.

We do not distinguish between the share of employment reallocating across sectors that is caused by each channel, rather we note that all three together provide evidence consistent with the conceptual channels in Section 2. More importantly, all three are consistent with the speed at which we see the employment share shifts taking place.

A final question is why we see employment shifts out of agriculture and into services that are larger for women than men. We note that differences between men's and women's work through the structural transformation of an economy are not well understood in African settings. This makes it difficult to know whether we should expect larger responses to the labor migration shock from men or from women. While women tend to be marginal or added workers in agricultural models of household production (e.g. Schultz (2001) and LaFave and Thomas (2016), working less in non-farm family businesses, it is not clear how we should expect female employment to change at a more aggregate level, as economies grow richer. Reviewing the literature on structural transformation reveals a significant gap in the literature on this front.

Although we have come to expect a U-shaped pattern of female labor force participation that is seen through the structural transformation (Goldin, 1995), empirically, these patterns are mainly observed in historically rich countries, in which manufacturing has been an important sector of work through the structural transformation. As we noted in section 2, Malawi's structural transformation resembles patterns in the rest of Africa (Rodrik, 2016), where manufacturing has not been as important a sector of work. If women do have comparative advantage in services (e.g. Ngai and Petrongolo (2017) and Ngai, Olivetti, and Petrongolo (2021)), and if the way that African economies develop is to skip over manufacturing, then it does not seem unreasonable that female employment shifts out of agriculture and into low-skilled services in response to the labor migration shock in our Malawi setting. We leave exploration of this topic to future work.

8 Conclusion

This paper marshals historical data from Malawi to exploit a natural experiment that shocked access to international labor migration in order to demonstrate that migrant capital can contribute to changes in the structure of rural labor markets in the long run. In places receiving larger inflows of migrant capital per worker, employment shifts out of agriculture and towards the service sector. This is particularly the case for women. Jobs in construction, retail, general services, and transport and communications increased and employment became more diverse in those districts that received more capital per worker after the migration shock. Even after the end of migration, accumulation persisted at higher rates. Districts with more migrant capital invested more in physical, non-farm capital and in human capital (especially for girls) over the long run. They are also wealthier.

Our work sheds light on a relatively understudied period in Malawi's economic history and is broadly relevant to African labor markets in the past and present. Many southern African countries were affected by similar fluctuations in worker flows to the South African gold mines. Structural change could also have occurred in these other countries as a result of capital accumulated through international labor migration.

Although historical, our work is policy-relevant for countries considering temporary or seasonal labor migration programs. Our results suggest that legal, time-limited migration might present a practical way for communities to accumulate capital in labor-rich, resource-poor countries, with important implications for women's work in these countries. When such migration flows are widespread, and accompanied by large return flows of money, impacts on the local labor market can be persistent, with positive long-term consequences for labor allocation across sectors at home.

References

Alvarez-Cuadrado, Francisco and Francesco Amodio and Markus Poschke, 2021. "Selection and Absolute Advantage in Farming and Entrepreneurship", Working Paper.

Balboni, Clare, Oriana Bandiera, Robin Burgess, Maitreesh Ghatak and Anton Heil, 2022. "Why Do People Stay Poor?", *The Quarterly Journal of Economics* 137 (2): 785-844.

Beegle, Kathleen, Emanuela Galasso and Jessica Goldberg. 2017. "Direct and Indirect Effects of Malawi's Public Works Program on Food Security" *Journal of Development Economics* 128: 1-23.

Buera, Francisco J., Joseph P. Kaboski and Yongseok Shin. 2013. "Macro-perspective on asset grants programs: Occupational and wealth mobility." *American Economic Review Papers and Proceedings* 104 (5): 159–164

Bustos, Paula, Bruno Caprettini and Jacopo Ponticelli. 2016. "Agricultural Productivity and Structural Transformation. Evidence from Brazil", *American Economic Review*, 106(6): 1320-1365

Bustos, Paula, Gabriel Garber, and Jacopo Ponticelli. 2020. "Capital accumulation and Structural Transformation", *Quarterly Journal of Economics*, 135 (2): 1037–1094.

Chen, Chaoran and Diego Restuccia and Raul Santaeulalia-Llopis. 2017. "Land Misallocation and Productivity", *NBER Working Paper* No. 23128. Doi: 10.3386/w23128.

Clemens, Michael and David McKenzie. 2018. "Why don't remittances appear to affect growth?", *The Economic Journal* 128(612) $https://doi.org/10.1111/ecoj.12463$

Davidson, R. and James G. MacKinnon. 2010. "Wild bootstrap tests for IV regression", *Journal of Business and Economic Statistics* 28: 128–144.

Dillon, Brian, Peter Brummund, and Germano Mwabu. 2019. "Asymmetric Non-separation and Rural Labor Markets", *Journal of Development Economics* Vol. 139: 78-96.

Dinkelman, Taryn and Martine Mariotti. 2016. "The long-run effects of labor migration on human capital formation in communities of origin", *American Economic Journal: Applied Economics*, 8 (4): 1-35.

Dustmann, Christian, and Oliver Kirchkamp. 2002. "The Optimal Migration Duration and Activity Choice after Re-migration", *Journal of Development Economics* 67 (2): 351-72.

Erten, Bilge and Jessica Leight. 2021. "Exporting out of agriculture: The Impact of WTO Accession on Structural Transformation in China",*Review of Economics and Statistics* 103 (2): 364–380.

Fafchamps, Marcel and Agnes Quisumbing. 1999. "Human Capital, Productivity, and Labor Allocation in Rural Pakistan" *Journal of Human Resources*, 34(2)

Foster, Andrew D. and Mark R. Rosenzweig. 2008. "Economic development and the decline of agricultural employment" in Handbook of Development Economics Vol.4: Chapter 47. Elsevier.

Gibson, John, David McKenzie and Halahingano Rohorua. 2014. "Development impacts of seasonal and temporary migration: A review of evidence from the Pacific and Southeast Asia." *Asia and the Pacific Policy Studies* 1(1): 18-32

Gibson, John and David McKenzie. 2014. "The Development Impact of a Best Practice Seasonal Worker Policy", *Review of Economics and Statistics*, 96(2): 229-243

Goldberg, Jessica. 2016. "Kwacha Gonna Do? Experimental Evidence about Labor Supply in Rural Malawi" *American Economic Journal: Applied Economics* 8: 1

Goldin, Claudia. 1995. "The U-Shaped Female Labor Force Function in Economic Development and Economic History" in: Schultz T.P. 1995. Investment in Women's Human Capital and Economic Development. University of Chicago Press. pp. 61-90

Gollin, Douglas, Stephen L. Parente, and Richard Rogerson. 2002. "The Role of Agriculture in Development" *American Economic Review Papers and Proceedings* 92: 2

Gollin, Douglas, David Lagakos and Michael E. Waugh. 2014. "The agricultural productivity gap", *Quarterly Journal of Economics*, 129(2): 939-993

Gollin, Douglas, Remi Jedwab and Dietrich Vollrath. 2016 "Urbanization with and without Industrialization", *Journal of Economic Growth* 21 (1): 35-70

Herrendorf, Berthold, Richard Rogerson and Àkos Valentinyi. 2014. "Growth and structural transformation", in (eds.) Philippe Aghion and Steven N. Durlauf, Handbook of Economic Growth Volume 2, Elsevier.

Huffman, William. 1980. "Farm and off-farm work decisions: The role of human capital", *Review of Economics and Statistics*, 62(1): 14-34

Kaboski, Joseph P. and Robert M. Townsend. 2012. "The impact of credit on village economies", *American Economic Journal: Applied Economics* 4(2), pages 98-133, April.

Khanna, Gaurav, Emir Murathanoglu, Caroline Theoharides, and Dean Yang 2022. "Abundance from abroad: Migrant earnings and economic development in the Philippines", (NBER Working Paper No. 29862).

Kinnan, Cynthia, Shing-Yi Wang and Yongxiang Wang. 2016. "Access to migration for rural households", *American Economic Journal: Applied Economics*, 10 (4): 79–119.

Kosack, Edward. 2021. "Guest Worker Programs and Human Capital Investment: The Bracero Program in Mexico, 1942-1964" *Journal of Human Resources*, 56 (2): 570–599.

LaFave, Daniel and Duncan Thomas. 2016. "Farms, families, and markets: New evidence on completeness of markets in agricultural settings" *Econometrica* 84 (5): 1917–60.

Lewis, W. Arthur. 1954. "Economic Development with Unlimited Supplies of Labor", *The Manchester School*, 22(2): 139-91

Liu, R. Y. 1988."Bootstrap procedures under some non-I.I.D. models", *Annals of Statistics* 16: 1696–1708

Mariotti, Martine. 2015. "Fathers' Employment and Sons' Stature: The Long-Run Effects of a Positive Regional Employment Shock in South Africa's Mining Industry", *Economic Development and Cultural Change*, 63(3): 485-514.

McCaig, Brian and Nina Pavcnik. 2018. "Export markets and labor allocation in a low-income country", *American Economic Review* 108(7): 1899-1941.

McMillan Margaret, Dani Rodrik and Íñigo Verduzco-Gallo. 2014. "Globalization, Structural Change, and Productivity Growth, with an Update on Africa." *World Development* 63: 11–32.

Ngai, L. Rachel, Claudia Olivetti and Barbara Petrongolo. 2021. "Structural transformation and U-shaped female employment". Working Paper.

Ngai, L. Rachel and Barbara Petrongolo. 2017. "Gender gaps and the rise of the service economy" *American Economic Journal: Macroeconomics* 9 (4): 1-44.

Porzio Tommaso and Federico Rossi and Gabriella Santangelo. 2021. "Does schooling cause structural transformation?", Conditionally accepted, *American Economic Review*. Accessed May 2019.

Robinson, Amanda Lea. 2016. "Internal Borders: Ethnic Diversity and Market Segmentation in Malawi", *World Development* 87:371–384.

Rodrik, Dani. 2016 "Premature deindustrialization" *Journal of Economic Growth* 21: 1–33.

Roodman, David, James G. MacKinnon, Morten Ørregaard Nielsen, and Matthew D. Webb. 2019. "Fast and wild: Bootstrap inference in Stata using boottest", *The Stata Journal* 19 (1): 4—60 10.1177/1536867X19830877

Rosenstein-Rodan, P. N. 1943. "Problems of Industrialisation of Eastern and South-Eastern Europe"', *Economic Journal*, 53 (202): 202-211

Schultz, T. Paul. 2001 "Women's Roles in the Agricultural Household: Bargaining and Human Capital Investments," in Handbook of Agricultural Economics, Vol. 1, ed. by B. Gardner and G. Rausser. Amsterdam: North-Holland.

Theoharides, Caroline. 2018. "Manila to Malaysia, Quezon to Qatar: International Migration and its Effects on Origin-Country Human Capital", *Journal of Human Resources*, Vol. 53. No. 4, Fall 2018, pp. 1022-1049.

Theoharides, Caroline. 2020. "The Unintended Consequences of Migration Policy on Origin-Country Labor Market Decisions", *Journal of Development Economics*, 142

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.

Woodruff, Christopher, and Rene Zenteno. 2007. “Migrant Networks and Microenterprises in Mexico.”, *Journal of Development Economics*, 82 (2): 509-28.

World Bank. 2014. “Global Financial Development Database (GFDD)”, downloaded from *http : //data.worldbank.org/data − catalog/global − financial − development* on August 9, 2016

World Bank. 2016. “World Bank national accounts data, and OECD National Accounts data files”, data downloaded from *http : //data.worldbank.org/indicator/NV.SRV.TETC.ZS?locations = MW* on September 2, 2016

World Bank. 2019. “Migration and Remittances: Recent Developments and Outlook”, *Migration and Development Brief Num. 31*, April *http : //www.knomad.org/sites/default/files/2019 − 04/Migrationanddevelopmentbrief31.pdf*

Yang, Dennis. 1997. “Education and Off-Farm Work”, *Economic Development and Cultural Change* 45 (3): 613-632.

Yang, Dean. 2008. “International Migration, Remittances and Household Investment: Evidence from Philippine Migrants’ Exchange Rate Shocks.” *Economic Journal* 118 (528): 591–630.

Young, Alwyn. 2019. “Channeling Fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results” *Quarterly Journal of Economics* 134 (2):557–598.

Appendix Figures and Tables

Figure 1: Capital Flows and Migrant Workers Over Time

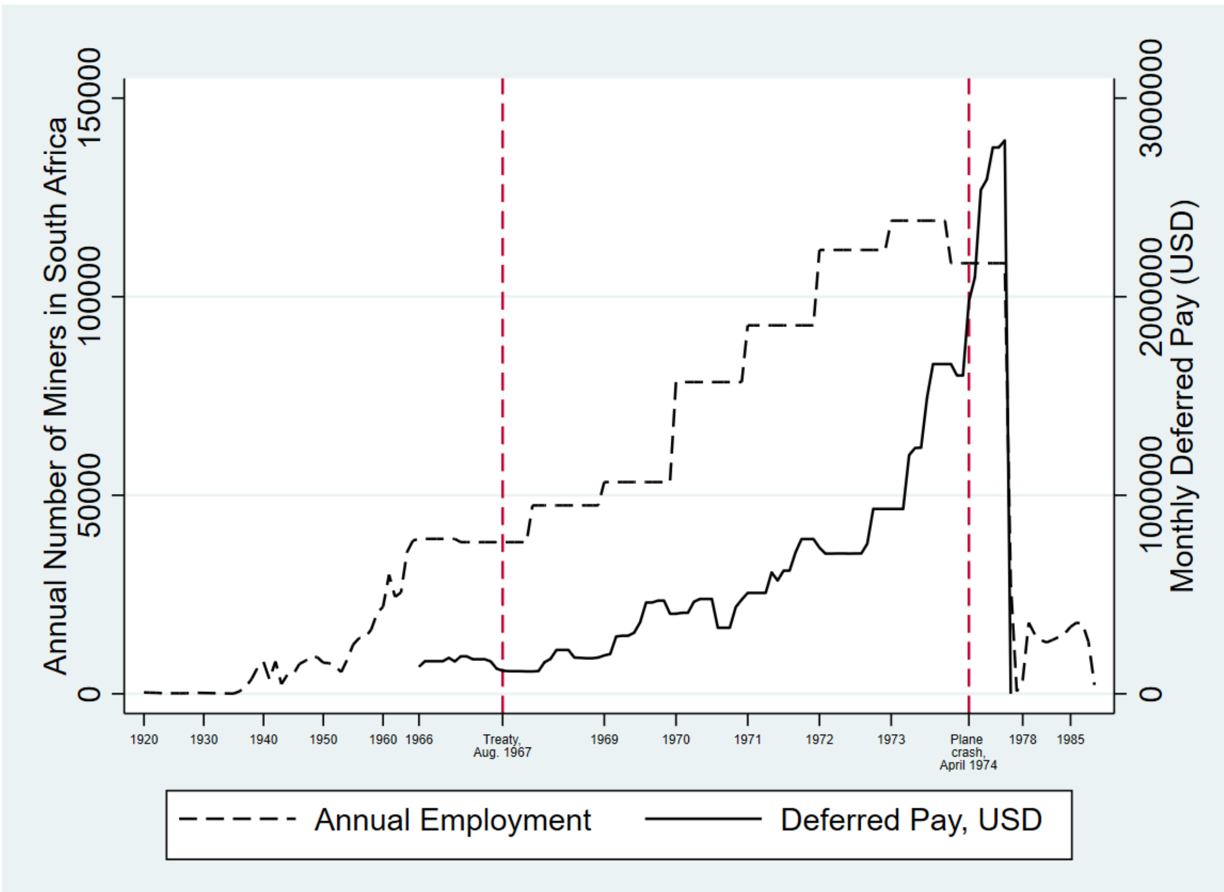
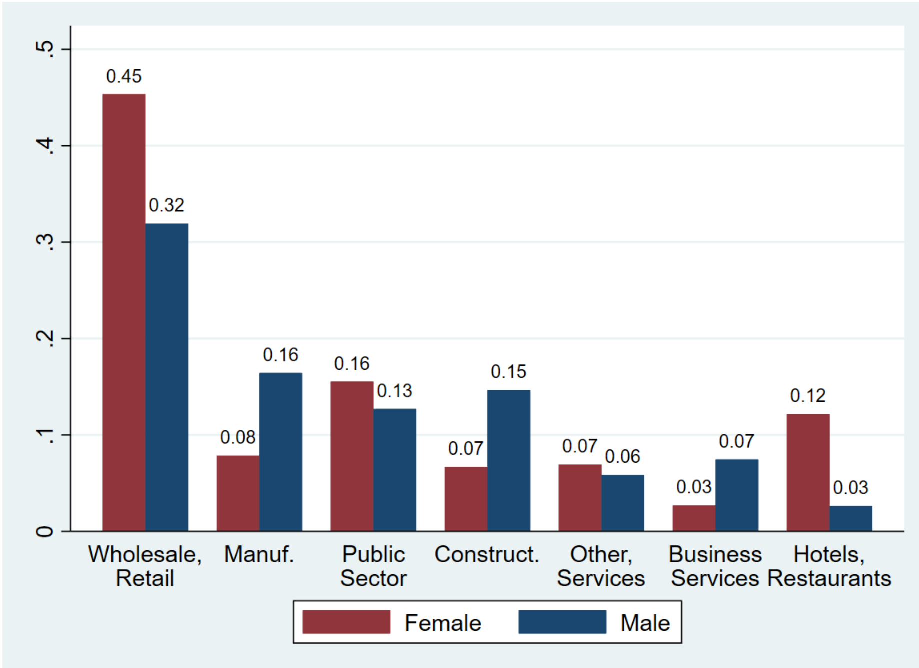


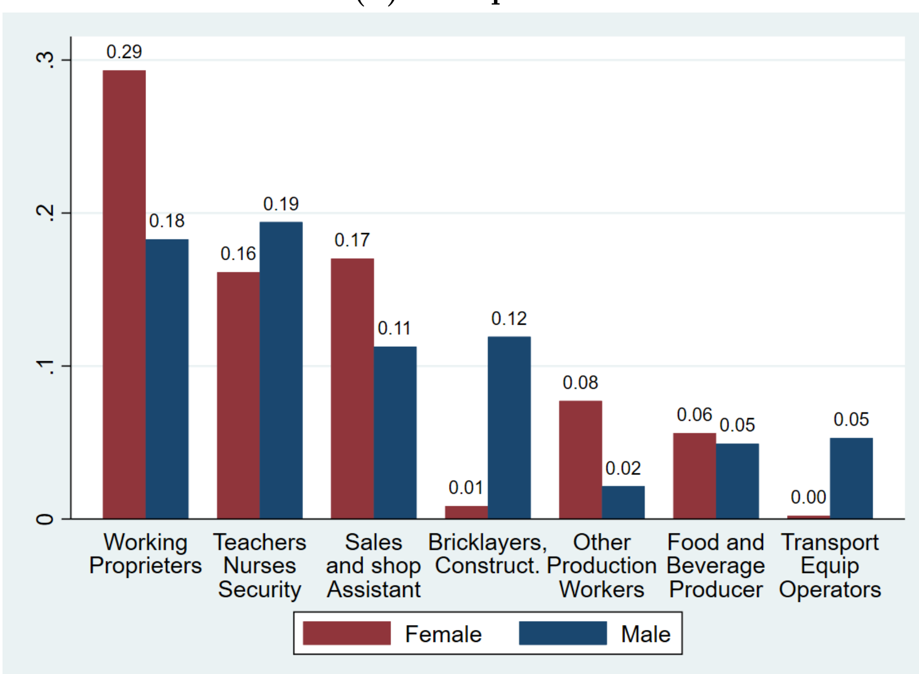
Figure 1 plots annual number of Malawian migrants contracted to work on South African mines on the Y-axis on the left, and the monthly deferred pay amounts in USD returned to Malawi on the Y-axis on the right. The two vertical dashed lines represent (from left to right) the abolition of labor quotas in August 1967 and the moratorium on migration after the April 1974 Malawian plane crash.

Figure 2: Most Prevalent Industries and Occupations among Non-farm Workers by Gender, 2008

(a) Industry of Work



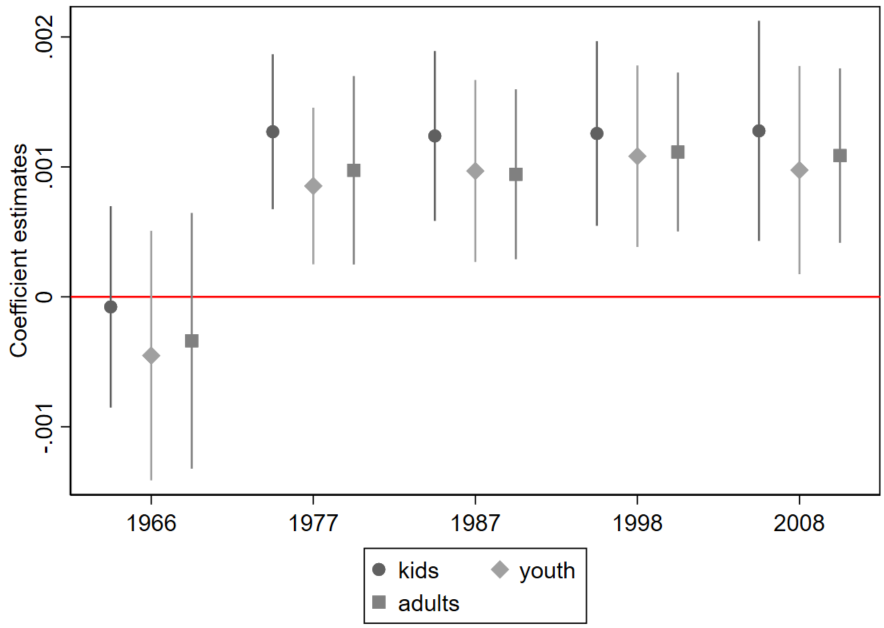
(b) Occupation



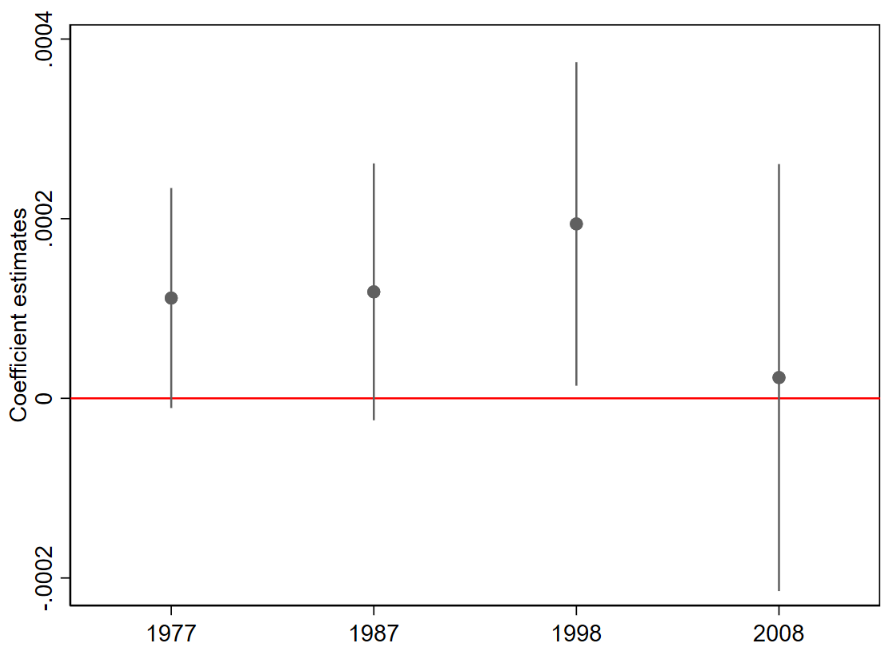
Figures indicate the share of men and women employed in non-farm industries (left) or non-farm occupations (right) using two digit industry and occupation classifications in the 2008 Malawi Census.

Figure 3: Effects of K/L Shock on District-level Population Growth and Urbanization

(a) Cohort-specific population density



(b) Urbanization



Figures plot the coefficients from equation (1), for equations where the outcomes are the log of the age-specific population totals in the district (on the left) or the share of the urban population in the district (on the right). Age groups are: adults (19+), youth $(5 - 18)$ and children (under age 5). Points shown are coefficients on the interaction of Census year dummies with the district-level K/L Shock variable. Base year is 1945. Mean value of K/L Shock is 87.6 Kwacha (73 USD). Full regression results are in Online Appendix Table A.2.

Table 1: **Employment Shares by Sector and Gender Over Time**

	1977	1987	1998	2008
Sector of work: Women				
Agriculture	0.943	0.949	0.888	0.708
Manufacturing	0.016	0.012	0.012	0.037
Services	0.028	0.032	0.067	0.198
Industrial Concentration Index	0.893	0.905	0.805	0.550
Sector of work: Men				
Agriculture	0.760	0.764	0.731	0.532
Manufacturing	0.093	0.076	0.074	0.132
Services	0.120	0.134	0.171	0.279
Industrial Concentration Index	0.618	0.625	0.589	0.357

Population-weighted shares of adults in each sector of work and employment category from Census data. Information on the industrial sector of work for the economically active population (workers and unemployed) 10 years and older are collapsed to district-gender cells. 24 observations per cell. Home workers are excluded from these definitions. Industrial Concentration Index is a Herfindahl index of sector of work; larger values imply more concentration of work sector in the district. Data appendix contains details of dataset construction. Totals do not sum to 1 because of residual 'not stated' categories for industry of work.

Table 2: **Correlates of District-Level K/L Shock and K/L Total**

	K/L Shock		K/L 1967-1977	
	(1)	(2)	(3)	(4)
Malaria area [0/1]	-217.9**	-84.47	-290.4**	-101.1
	(88.35)	(50.66)	(125.68)	(74.96)
	[0.118]	[0.154]	[0.136]	[0.228]
Log Pop Density 1945	-71.05	-24.01	-90.93	-24.17
	(46.68)	(21.29)	(66.40)	(32.10)
	[0.332]	[0.312]	[0.390]	[0.470]
Youth Literacy Rate 1945	-4.628	249.9	-23.44	337.8
	(854.97)	(521.67)	(1224.77)	(792.46)
	[0.996]	[0.694]	[0.988]	[0.736]
Estate District [0/1]	-73.99	-16.45	-103.1	-21.5
	(58.34)	(35.19)	(83.51)	(50.53)
	[0.268]	[0.774]	[0.270]	[0.778]
Male Share No Wage 1966	-2,552**	-698.2	-3,320*	-689.1
	(1144.91)	(497.46)	(1633.52)	(728.58)
	[0.508]	[0.234]	[0.566]	[0.416]
Female Share No Wage 1966	1,854**	447.1	2,462*	465.6
	(814.84)	(362.03)	(1161.85)	(531.87)
	[0.398]	[0.338]	[0.444]	[0.476]
Central Region [0/1]	98.37	113.7	153.6	175.2
	(98.95)	(77.11)	(143.87)	(112.68)
	[0.362]	[0.260]	[0.328]	[0.218]
Southern Region [0/1]	53.59	305.8***	59.67	417.6***
	(110.94)	(84.10)	(160.59)	(124.77)
	[0.588]	[0.004]	[0.690]	[0.010]
Num. Recruiting Stations		42.24***		59.93***
		(8.85)		(12.61)
		[0.002]		[0.004]
Observations	24	24	24	24
R2	0.67	0.92	0.63	0.90
p-val joint null†	0.21	0.00	0.28	0.00
Wild bootstrap p-val†	0.75	0.19	0.86	0.27
Y Mean	87.60	87.60	129.41	129.41

Robust standard errors in parentheses. Significance levels ***p< 0.01, **p< 0.05, *p< 0.1 where critical values are taken from the small sample t-distribution. Wild bootstrap p-values for small samples are shown in square brackets. Results are district population weighted (1977 weights). Unit of observation is the district. Num. recruiting stations is a count of historically-placed recruiting stations in each district. † Joint tests are conducted for all variables excluding recruiting stations. All other controls are measured at district level prior to 1967.

Table 3: **Long-run Impacts of K/L Shock on Sectoral Employment Shares**

	Agric.	Manuf.	Services	Industry Concen.
Panel A: Share of Women in Each Industry				
(K/L)Shock*3 Decades Post	-.00032***	.000046	.00036***	-.000022
	(.0001)	(.000028)	(.00012)	(.00012)
	[0.045]	[0.064]	[0.005]	[0.057]
(K/L)Shock*2 Decades Post	-.00028***	.000019	.00017**	-.00027**
	(.000074)	(.00002)	(.000076)	(.0001)
	[0.028]	[0.017]	[0.004]	[0.104]
(K/L)Shock*1 Decade Post	-.00006	.000019*	.000027	-.0001**
	(.000037)	(.000011)	(.000041)	(.000044)
	[0.009]	[0.410]	[0.001]	[0.347]
Y Mean	0.878	0.019	0.076	0.796
$\alpha_1 = \alpha_2 = \alpha_3 = 0$	0.005	0.055	0.004	0.002
R.I. p-values	0.190	0.410	0.162	0.184
$\alpha_1 + \alpha_2 + \alpha_3 = 0$	0.003	0.148	0.021	0.118
R.I. p-values	0.038	0.298	0.116	0.332
Panel B: Share of Men in Each Industry				
(K/L)Shock*3 Decades Post	.000056	7.0e-06	.000016	.00021
	(.00015)	(.000078)	(.000074)	(.00021)
	[0.097]	[0.007]	[0.000]	[0.489]
(K/L)Shock*2 Decades Post	-.00014	.000032	.00013*	-.00011
	(.00011)	(.000052)	(.000057)	(.00014)
	[0.031]	[0.021]	[0.000]	[0.101]
(K/L)Shock*1 Decade Post	-.000081	4.8e-06	.000019	-.00012
	(.000056)	(.000029)	(.000029)	(.000076)
	[0.007]	[0.033]	[0.000]	[0.367]
Y Mean	0.728	0.087	0.152	0.577
$\alpha_1 = \alpha_2 = \alpha_3 = 0$	0.000	0.559	0.000	0.000
R.I. p-values	0.070	0.728	0.122	0.062
$\alpha_1 + \alpha_2 + \alpha_3 = 0$	0.602	0.779	0.306	0.966
R.I. p-values	0.670	0.834	0.482	0.966

***p< 0.01, **p< 0.05, *p< 0.1. Robust s.e. clustered at district level in parentheses. Randomization inference p-values in square brackets. K/L SHOCK is total deferred pay (in Kwacha) returning to a district after April 1974 divided by number of miners returning to Malawi at that time (mean 87.6). Baseline year is 1977. District-year cell is unit of observation, $N = 96$ in all regressions, N. Districts= 24. Controls include: year and district fixed effects and time trend interactions with baseline district controls (1945 adult literacy, 1945 population density, a malaria dummy, share of men and women married in 1966, share of men and women earning no cash income in 1966, two region dummies). Industrial concentration is a Herfindahl index measuring how concentrated work is in any one sector.

Table 4: **Long-run Impacts of K/L Shock on Non-farm Sectors**

	General Manuf.	Constr.	General Services	Retail	Trans. and Comms.
Panel A: Share of Women in Each Industry					
(K/L)Shock*3 Decades Post	.000023	.000022*	.000091*	.00026***	4.1e-06
	(.000027)	(.000013)	(.000047)	(.000071)	(2.4e-06)
	[0.067]	[0.208]	[0.028]	[0.090]	[0.047]
(K/L)Shock*2 Decades Post	8.4e-06	.000011	.00004	.00013***	3.2e-06*
	(.000019)	(8.4e-06)	(.000035)	(.000046)	(1.6e-06)
	[0.028]	[0.407]	[0.016]	[0.053]	[0.027]
(K/L)Shock*1 Decade Post	.000011	8.1e-06*	.000014	8.9e-06	4.2e-06***
	(.00001)	(4.5e-06)	(.000022)	(.000021)	(8.4e-07)
	[0.123]	[0.408]	[0.001]	[0.077]	[0.003]
Y Mean	0.014	0.005	0.029	0.046	0.001
$\alpha_1 = \alpha_2 = \alpha_3 = 0$	0.233	0.232	0.109	0.000	0.000
R.I. p-values	0.558	0.530	0.430	0.046	0.014
$\alpha_1 + \alpha_2 + \alpha_3 = 0$	0.450	0.106	0.156	0.007	0.025
R.I. p-values	0.494	0.254	0.280	0.066	0.120
Panel B: Share of Men in Each Industry					
(K/L)Shock*3 Decades Post	-.000052	.000068	-.000082**	.00014***	-.00004***
	(.000062)	(.000046)	(.000031)	(.000047)	(.000011)
	[0.299]	[0.065]	[0.025]	[0.189]	[0.017]
(K/L)Shock*2 Decades Post	-.000077*	.00012***	7.3e-07	.00017***	-.000045***
	(.000044)	(.00003)	(.000025)	(.000038)	(8.3e-06)
	[0.093]	[0.089]	[0.015]	[0.006]	[0.007]
(K/L)Shock*1 Decade Post	-.000039	.000043**	-.000011	.00004**	-9.1e-06*
	(.000024)	(.00002)	(.000019)	(.000016)	(5.2e-06)
	[0.102]	[0.451]	[0.000]	[0.112]	[0.007]
Y Mean	0.049	0.036	0.074	0.065	0.013
$\alpha_1 = \alpha_2 = \alpha_3 = 0$	0.034	0.000	0.000	0.000	0.000
R.I. p-values	0.310	0.038	0.138	0.036	0.006
$\alpha_1 + \alpha_2 + \alpha_3 = 0$	0.202	0.023	0.193	0.002	0.001
R.I. p-values	0.342	0.128	0.338	0.086	0.056

***p< 0.01, **p< 0.05, *p< 0.1. Robust s.e. clustered at district level in parentheses. Randomization inference p-values in square brackets. K/L SHOCK is total deferred pay (in Kwacha) returning to a district post-plane crash divided by number of miners returning to Malawi post-plane crash (mean 87.6). Baseline year is 1977. District-year cell is unit of observation, $N = 96$ in all regressions, N. Districts= 24. Controls include year and district fixed effects and time trend interactions with baseline district controls (as in Table 3).

Table 5: **Long-Run Impacts of K/L Shock on Type of Employer**

	Workers in All Sectors			Workers in Service Sectors		
	Self Emp.	Family Bus. or Farm	Wage Work	Self Emp.	Family Bus. or Farm	Wage Work
Panel A: Share of Women						
K/L Shock*3 Decades Post	-.00016	.00015***	.000016	-.000012	.00016***	-.00013
	(.00012)	(.00002)	(.000052)	(.00011)	(.000055)	(.00011)
	[0.680]	[0.644]	[0.295]	[0.038]	[0.059]	[0.259]
K/L Shock*2 Decades Post	-.00028***	.00015***	.00004	.000038	.000058	-.00016**
	(.000078)	(.000027)	(.000029)	(.000068)	(.000048)	(.000072)
	[0.680]	[0.644]	[0.298]	[0.038]	[0.079]	[0.261]
Y Mean	0.876	0.028	0.046	0.073	0.042	0.869
α_1=α_2=α_3=0	0.001	0.000	0.055	0.459	0.012	0.074
R.I. p value	0.078	0.026	0.284	0.580	0.164	0.256
Panel B: Share of Men						
K/L Shock *3 Decades Post	.00014	.000063**	-.000076	.000018	.000096	-.000062
	(.00013)	(.000026)	(.00013)	(.000041)	(.000057)	(.000057)
	[0.689]	[0.658]	[0.299]	[0.044]	[0.029]	[0.299]
K/L Shock *2 Decades Post	-.000058	.000057***	.00003	.000023	-.000011	-.00003
	(.000068)	(.000015)	(.000067)	(.000037)	(.000026)	(.000048)
	[0.689]	[0.658]	[0.305]	[0.044]	[0.043]	[0.301]
Y Mean	0.682	0.031	0.219	0.102	0.026	0.859
α_1=α_2=α_3=0	0.000	0.001	0.064	0.828	0.001	0.532
R.I. p value	0.068	0.076	0.320	0.888	0.074	0.692

***p< 0.01, **p< 0.05, *p< 0.1. Robust s.e. clustered at district level in parentheses. Randomization inference p-values in square brackets. K/L SHOCK is total deferred pay (in Kwacha) returning to each district (May 1974-November 1975) divided by total number of migrants returning to the district in this period (mean = 87.6). Data are from 1987 (one decade post), 1998 (two decades post) and 2008 (three decades post). Unit of observation is the district-gender cell (N= 72). Total districts = 24. Controls include district and year fixed effects and interactions of a linear trend term with baseline district-level variables (as in Table 3).

Table 6: **Long-run Impacts of K/L Shocks on Farm and Non-farm Investments**

Farm Capital	Pangas(*)	Hoes(*)	Any Cattle	Oxcart
(K/L)Shock*3 Decades Post				-.000027**
				(.00001)
				[0.080]
(K/L)Shock*2 Decades Post	-.00044**	-.00075***	.000052	-.000014
	(.00019)	(.00026)	(.000092)	(.00001)
	[0.424]	[0.216]	[0.210]	[0.153]
Y-mean	0.282	1.463	0.094	0.015
Years of Data in Sample	68, 77	68, 77	68, 77	87, 98, 08
N	46	46	46	69
α_1=α_2=0	-	-	-	0.048
R.I. p-value	-	-	-	0.000

Non-Farm Capital	Radio	Durable Walls	Durable Roof	Durable Roof and Walls
(K/L)Shock*3 Decades Post	-.000063*	.00037**	.00054***	.00056***
	(.000036)	(.00017)	(.000065)	(.000064)
	[0.137]	[0.302]	[0.087]	[0.093]
(K/L)Shock*2 Decades Post	.000033			
	(.000052)			
	[0.346]			
(K/L)Shock*1 Decade Post	.000042	.0007***	.00055***	.00055***
	(.000031)	(.00022)	(.000046)	(.000042)
	[0.519]	[0.084]	[0.041]	[0.041]
(K/L)Shock*Post	.000041**			
	(.000019)			
	[0.155]			
Y-mean	0.277	0.389	0.133	0.124
Years of Data in Sample	all	69, 87, 08	69, 87, 08	69, 87, 08
N	115	69	69	69
All $\alpha's$=0	0.000	0.008	0.000	0.000
R.I. p-value	0.000	0.000	0.000	0.000

***p< 0.01, **p< 0.05, *p< 0.1. Robust s.e clustered at district level in parentheses. Randomization inference p-values in square brackets. Outcome is share of households owning each asset; (*) is mean number of hoes and pangas per household. Unit of observation is district-year cell. Total districts with data in all Census years including 1966: 23. Base year is 1968 in all regressions, except for Oxcarts, where base year is 1987. Controls include year and district fixed effects and interactions of a trend term with baseline district variables (as in Table 3). Regressions are weighted.

Table 7: **Does a Temporary K/L Shock Lead to Greater Wealth 15 years Later?**

	Full Sample	Oldest Households: Head $\geq$ 38 y.o	Youngest Households: Head $<$ 38 y.o
DHS Wealth Index (s.d)	0.00114***	0.00128***	0.000930***
	(0.000)	(0.000)	(0.000)
Count of assets	0.00187***	0.00179***	0.00196*
	(0.001)	(0.001)	(0.001)
Durable Roof	0.000576***	0.000570***	0.000567***
	(0.000)	(0.000)	(0.000)
Durable Floor	0.000491***	0.000537***	0.000420***
	(0.000)	(0.000)	(0.000)
Electricity	0.0000917***	0.000143***	0.00000552
	(0.000)	(0.000)	(0.000)
Radio	0.000294*	0.000345**	0.000249
	(0.000)	(0.000)	(0.000)
Car	0.0000681***	0.000103***	0.0000315
	(0.000)	(0.000)	(0.000)
Motorbike	0.0000124	0.0000208	0.00000331
	(0.000)	(0.000)	(0.000)
Bike	-0.000180	-0.000213**	-0.0000868
	(0.000)	(0.000)	(0.000)
Improved Toilet	0.000154	0.0000218	0.000357
	(0.000)	(0.000)	(0.000)
Improved Water Source	0.0000693	-0.00000698	0.000128
	(0.000)	(0.000)	(0.000)

Sample size = $5,323$ in full regressions, $3,035$ in older cohorts and $2,288$ in younger cohort regressions. Younger cohorts are those where the household head is younger than 38 years old; older cohorts are those for whom the household head is 38 years or older in 1992. Robust standard errors clustered on district (24 districts); regressions are weighted using DHS sample weights. Each block of coefficients is taken from a regression of the relevant asset outcome variable on the (K/L) SHOCK variable. All regressions control for region fixed effects and baseline district-level variables (as in Table 3). All outcomes are binary variables except for number of assets (the count of asset measures per household) and the DHS wealth index, which is a composite index of household assets constructed using principal components analysis.

Table 8: **Long-Run Impacts of K/L Shock on Investment in Human Capital of the Next Generation, by Gender**

	Yrs of Completed Education		**Any Prim. Schooling**	
	Females	Males	Females	Males
(K/L) Shock*Shock Cohorts	.002***	.0011	.00016*	.000018
	(.00063)	(.00091)	(.000078)	(.0001)
	[0.456]	[0.496]	[0.454]	[0.562]
(K/L) Shock*Post Shock Cohorts	.0018***	.00056	.00014**	-.000039
	(.00041)	(.00076)	(.000053)	(.000098)
	[0.126]	[0.174]	[0.066]	[0.258]
N	240	240	240	240
Y mean	1.881	3.231	0.347	0.474
$\gamma_1=\gamma_2=0$	0.00121	0.474	0.0206	0.522
R.I. p-value	0.0260	0.622	0.0400	0.686

Robust standard errors clustered at the district level. Randomization inference p values in square brackets. Significance levels ***p< 0.01, **p< 0.05, *p< 0.1. Data are from Census 1977 and 1998. Unit of observation is the district-gender-cohort cell; 240 observations in each regression. Shock cohorts are those age-eligible for primary school during 1974-1977; Post-shock cohorts are those age-eligible for primary school 1977-1980. Other controls include a trend term interacted with baseline district variables: adult literacy in 1945, population density in 1945, a malaria dummy, and region dummies. All regressions contain district fixed effects and cohort dummies. Regressions are unweighted.

Online Appendices: Not for Publication

Figure A.1: Annual Wages for African Mineworkers and Global Gold Prices

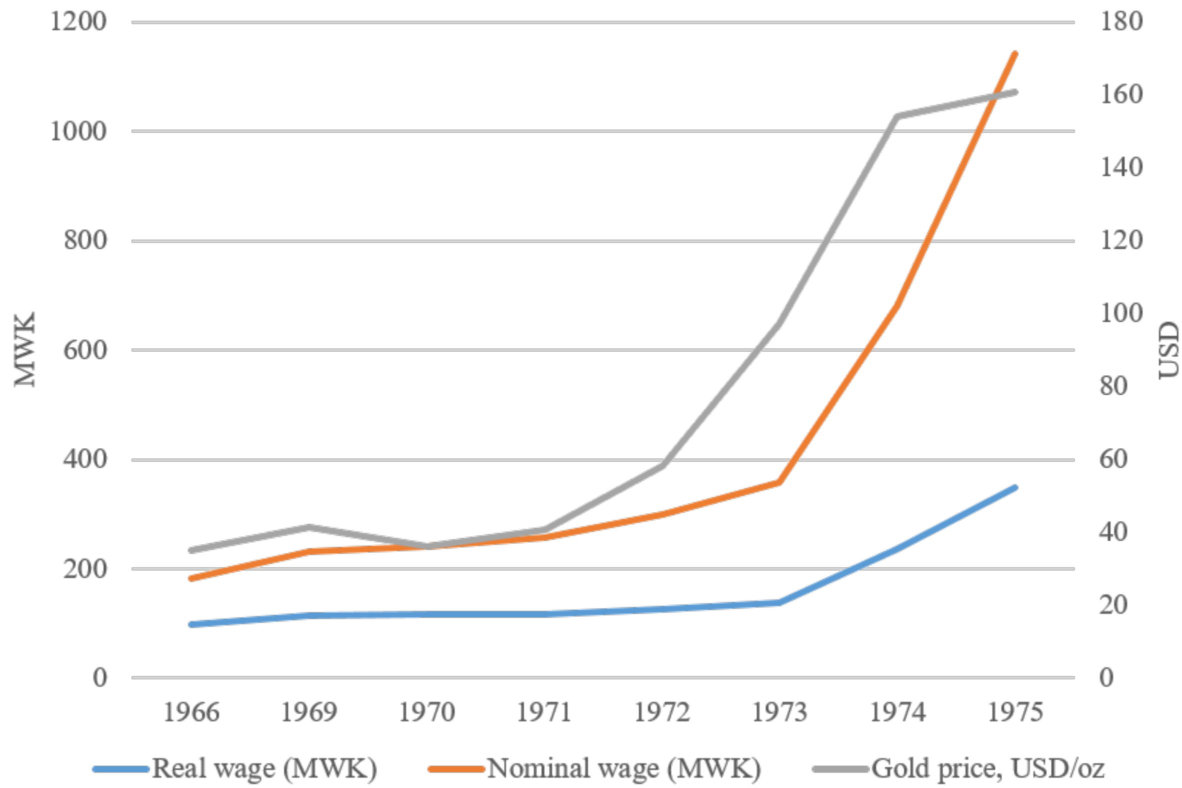


Figure plots annual average wages for mineworkers from Wilson (1972) for the years 1966 and 1969, and from Crush et al (1991) from 1970 onwards. Gold prices per ounce are from the National Mining Association, retrieved 17 March 2022: `https://nma.org/wpcontent/uploads/2016/09/historic_gold_prices_1833_pres.pdf`

Table A.1. **Summary statistics, district-level data**

	Mean	SD	Min	Max
Components of migration shock				
N. men, 1977	51,326	36,938	10,559	167,531
N. men ever migrant by 1977	19,557	15,421	4,232	75,324
Δ N. migrants, 1966-1977	13,642	10,667	2,816	50,121
Migrant share, 1974-1997	0.06	0.01	0.03	0.09
Tot. USD (mill.)/district, 1966-1975(*)	2.25	3.53	0.00	16.29
Tot. USD/person (mill.), 1966-1975	24.04	55.40	0.00	275.68
K/L Shock, post-plane crash	87.60	132.23	0.00	675.31
K/L, pre-plane crash, 1967-1974	204.48	274.40	0.00	1,394.02
Baseline District-level Variables				
Northern Region	0.21	0.41	0.00	1.00
Central Region	0.38	0.49	0.00	1.00
Southern Region	0.42	0.50	0.00	1.00
Population 1945	71,262	60,353	5,919	230,891
Pop. density per km^2, 1945	30.61	26.61	5.10	109.05
Share literate youth, 1945	0.08	0.04	0.03	0.14
Districts in high malaria area	0.28	0.35	0.00	1.00
Districts with any agricultural estate	0.46	0.51	0.00	1.00
Share men w. no cash income, 1966	0.37	0.10	0.22	0.59
Share women w. no cash income, 1966	0.48	0.14	0.28	0.72
Districts with any recruiting stations	0.63	0.49	0.00	1.00

Data for the first set of outcomes are district-level data collected from administrative records and from Census 1977. Adult men are aged 15-64. Data for the second set of outcomes comes from 1945 Census data, from geographic files for Malawi, and from author-collected data on the location of migrant recruiting stations. Literacy is in English and the vernacular. Raw means (unweighted) computed over 24 districts. Agricultural estate is a dummy variable indicating whether a district contains any cash crop estates (e.g. for tobacco or sugar).

Table A.2. **Impacts of K/L Shock on Population Growth and Urbanization**

	Ln pop	Ln Male Pop	Ln Fem. Pop	Ln Pop Age < 5	Ln Pop Age 5- 18	Ln Pop Age 18+	Share Urban
(K/L)Shock*	.0011***	.0013***	.00088**	.0013***	.00098**	.0011***	.000023
3 Decades Post	(.00035)	(.00036)	(.00034)	(.00041)	(.00039)	(.00032)	(.00011)
	[0.069]	[0.246]	[0.075]	[0.067]	[0.071]	[0.070]	[0.055]
(K/L)Shock*	.0012***	.0014***	.00091***	.0013***	.0011***	.0011***	.00019**
2 Decades Post	(.0003)	(.00031)	(.00029)	(.00034)	(.00034)	(.0003)	(.000087)
	[0.064]	[0.154]	[0.038]	[0.059]	[0.044]	[0.041]	[0.049]
(K/L)Shock*	.001***	.0013***	.00078**	.0012***	.00097***	.00094***	.00012*
1 Decade Post	(.00031)	(.00031)	(.00031)	(.00032)	(.00034)	(.00032)	(.000069)
	[0.121]	[0.132]	[0.043]	[0.122]	[0.062]	[0.062]	[0.036]
(K/L)Shock*	.001***	.0013***	.00071**	.0013***	.00085***	.00097**	.00011*
End of Migration	(.0003)	(.00032)	(.00028)	(.00029)	(.00029)	(.00035)	(.000059)
	[0.126]	[0.099]	[0.054]	[0.128]	[0.062]	[0.061]	[0.055]
(K/L)Shock*	-.00031	-7.6e-06	-.00062	-.000077	-.00045	-.00034	
1 Decade Before	(.00044)	(.00047)	(.0004)	(.00037)	(.00046)	(.00048)	
	[0.382]	[0.338]	[0.359]	[0.385]	[0.398]	[0.377]	-
All $\alpha's = 0$	0.002	0.000	0.013	0.001	0.002	0.002	0.000
R.I. p-value	0.120	0.206	0.162	0.126	0.540	0.562	0.044
All $\alpha's = 0$	0.001	0.000	0.008	0.001	0.006	0.002	0.167
R.I. p-value	0.198	0.268	0.070	0.194	0.102	0.102	0.050

Robust standard errors clustered at district level, in parentheses. Randomization inference p-values for each coefficient are in square brackets. Significance levels ***p< 0.01, **p< 0.05, *p< 0.1. K/L SHOCK is the total deferred pay returning to each district between May 1973 and November 1975 in millions of Kwacha divided by the number of miners returning to Malawi in those years. Mean of K/L SHOCK is 87.6. Outcomes data are from Census 1945 (for all population outcomes except urbanization, it is the baseline); Census 1966 (the baseline for urbanization outcome); Census 1977 (end of migration), 1987 (one decade post), 1998 (two decades post) and 2008 (three decades post). Unit of observation is the district-year cell. Total districts = 24, except when Lilongwe is dropped. All regressions include year and district fixed effects. Except for columns (2) and (6), all regressions also include interactions of a time trend with baseline district controls (as in Table 3).

Table A.3. **Robustness: LR Impacts of K/L shock on Sectoral Employment**

Share of Women	Agriculture [Mean = 0.88]				Services [Mean = 0.08]			
(K/L)Shock*	-.00032***	-.0003***	-.0004***	-.00023***	.00036***	.00037***	.00044***	.00026***
3 Dec. Post	(.0001)	(.000049)	(.000039)	(.00004)	(.00012)	(.000064)	(.000071)	(.000052)
	[0.045]	[0.021]	[0.025]	[0.034]	[0.005]	[0.027]	[0.039]	[0.044]
(K/L)Shock*	-.00028***	-.00026***	-.00033***	-.00022***	.00017**	.00018***	.00023***	.00011***
2 Dec. Post	(.000074)	(.000042)	(.000032)	(.000036)	(.000076)	(.000032)	(.000049)	(.000038)
	[0.028]	[0.021]	[0.013]	[0.009]	[0.004]	[0.027]	[0.037]	[0.049]
(K/L)Shock*	-.00006	-.000054***	-.000085***	-.000031	.000027	.000031***	.000053	-5.9e-06
1 Dec. Post	(.000037)	(.000013)	(.000029)	(.000031)	(.000041)	(5.8e-06)	(.000033)	(.00003)
	[0.009]	[0.049]	[0.003]	[0.001]	[0.001]	[0.006]	[0.015]	[0.022]
Controls	All	No trends	No city†	$\Delta Migs^*$	All	No trends	No city†	$\Delta Migs^*$
N	96	96	92	96	96	96	92	96
All $\alpha' s = 0$	0.005	0.000	0.000	0.000	0.004	0.000	0.000	0.000
R.I. p-value	0.190	0.058	0.006	0.062	0.162	0.070	0.058	0.092
All $\alpha' s = 0$	0.003	0.000	0.000	0.000	0.021	0.000	0.000	0.003
R.I. p-value	0.038	0.038	0.000	0.012	0.116	0.028	0.014	0.034

Share of Men	Agriculture [Mean = 0.73]				Services [Mean = 0.15]			
(K/L)Shock*	.000056	.00013**	.000016	.000081	.000016	.000068*	.000083	-.000018
3 Dec. Post	(.00015)	(.000055)	(.00015)	(.00014)	(.000074)	(.000034)	(.000054)	(.000032)
	[0.097]	[0.067]	[0.045]	[0.062]	[0.000]	[0.002]	[0.067]	[0.067]
(K/L)Shock*	-.00014	-.000086**	-.00018*	-.00012	.00013**	.00016***	.00018***	.0001***
2 Dec. Post	(.00011)	(.000041)	(.0001)	(.000093)	(.000057)	(.000026)	(.000039)	(.000027)
	[0.031]	[0.090]	[0.022]	[0.021]	[0.000]	[0.001]	[0.002]	[0.011]
(K/L)Shock*	-.000081	-.000055	-.0001*	-.000071	.000019	.000037*	.000048**	7.2e-06
1 Dec. Post	(.000056)	(.000033)	(.000054)	(.000049)	(.000029)	(.000019)	(.000022)	(.000014)
	[0.007]	[0.115]	[0.007]	[0.005]	[0.000]	[0.000]	[0.054]	[0.043]
Controls	All	No trends	No city†	$\Delta Migs^*$	All	No trends	No city†	$\Delta Migs^*$
N	96	96	92	96	96	96	92	96
All $\alpha' s = 0$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
R.I. p-value	0.070	0.084	0.026	0.028	0.122	0.108	0.082	0.132
All $\alpha' s = 0$	0.602	0.947	0.381	0.695	0.306	0.001	0.009	0.155
R.I. p-value	0.670	0.960	0.426	0.722	0.482	0.078	0.076	0.310

Robust s.e. clustered at district level, in parentheses. Randomization inference p-values are in square brackets. ***p< 0.01, **p< 0.05, *p< 0.1. (K/L) SHOCK is total deferred pay returning to each district between May 1973 and November 1975 in millions of Kwacha divided by the number of miners returning to Malawi in those years (mean is 87.6). Unit of observation is the district-year cell. Total districts=24, except in the No City (†) specification when Lilongwe is dropped. All regressions include year and district fixed effects and interactions of a time trend with baseline district controls (as in Table 3). Columns (2) and (6) exclude these trend interactions. (*) Columns (4) and (8) also control for year dummies interacted with number of migrants from each district returning between 1967 and 1973 (pre-plane crash).

Table A.4. **Robustness: Weighting the K Shock by 1966 District Population Instead of by Migants**

	Agric.	Manuf.	Services	Industry Conc. Index
Panel A: Share of Women in Each Industry				
K Shock/Pop 66*3 Decades Post	-.0019***	.00024	.0022***	.00025
	(.00059)	(.00023)	(.00072)	(.00072)
	[0.049]	[0.135]	[0.000]	[0.079]
K Shock/Pop 66*2 Decades Post	-.0018***	.0001	.0011**	-.0017***
	(.00042)	(.00016)	(.00046)	(.0006)
	[0.032]	[0.080]	[0.000]	[0.123]
K Shock/Pop 66*1 Decade Post	-.00028	.0001	.000075	-.00048
	(.0002)	(.000091)	(.00022)	(.00031)
	[0.020]	[0.434]	[0.000]	[0.438]
Y Mean	0.878	0.019	0.076	0.796
$\alpha_1 = \alpha_2 = \alpha_3 = 0$	0.000	0.074	0.000	0.000
R.I. p-value	0.052	0.578	0.026	0.036
$\alpha_1 + \alpha_2 + \alpha_3 = 0$	0.003	0.361	0.024	0.234
Randomization inference p value	0.052	0.550	0.152	0.504
Panel B: Share of Men in Each Industry				
K Shock/Pop 66*3 Decades Post	.00077	-.00025	-.00006	.0019
	(.00088)	(.00045)	(.00045)	(.0013)
	[0.098]	[0.002]	[0.000]	[0.429]
K Shock/Pop 66*2 Decades Post	-.00065	.000015	.00076*	-.00038
	(.00065)	(.00032)	(.00036)	(.00093)
	[0.031]	[0.019]	[0.000]	[0.157]
K Shock/Pop 66*1 Decade Post	-.00033	-.00011	.00006	-.00052
	(.00031)	(.00016)	(.00018)	(.00044)
	[0.008]	[0.015]	[0.000]	[0.458]
Y Mean	0.728	0.087	0.152	0.577
$\alpha_1 = \alpha_2 = \alpha_3 = 0$	0.000	0.291	0.000	0.000
R.I. p-value	0.122	0.596	0.050	0.080
$\alpha_1 + \alpha_2 + \alpha_3 = 0$	0.911	0.713	0.441	0.706
Randomization inference p value	0.908	0.806	0.592	0.772

N = 96 in all regressions. Robust standard errors clustered at district level, in parentheses. Randomization inference p-values in square brackets. Significance levels ***p< 0.01, **p< 0.05, *p< 0.1. K SHOCK/Pop 66 is the total deferred pay returning to each district between May 1973 and November 1975 in millions of Kwacha divided by the district-level population in 1966. Mean of K SHOCK/Pop 66 is 7.81. Unit of observation is the district-gender cell. Total districts = 24. All regressions include district and year fixed effects, and interactions of a trend term with baseline district-level variables (as in Table 3).

Table A.5. **Robustness: Controlling for Tobacco Suitability of District**

	Agriculture		**Services**	
	Main Specification	No trend controls	Main Specification	No trend controls
Panel A: Share of Women				
(K/L)Shock*	-.00032**	-.00034***	.00036**	.000053*
3 Decades Post	(.0001)	(.000088)	(.00012)	(.000019)
	[0.045]	[0.043]	[0.005]	[0.005]
(K/L)Shock*	-.00028**	-.00029***	.00017*	.000023
2 Decades Post	(.000074)	(.000064)	(.000076)	(.000014)
	[0.028]	[0.027]	[0.004]	[0.003]
(K/L)Shock*	-.00006	-.000066*	.000027	.000021*
1 Decade Post	(.000037)	(.000031)	(.000041)	(9.0e-06)
	[0.009]	[0.009]	[0.001]	[0.001]
N	96	96	96	96
$\alpha_1 = \alpha_2 = \alpha_3 = 0$	0.00513	0.00153	0.00376	0.0289
R.I. p-value	0.190	0.152	0.162	0.160
$\alpha_1 + \alpha_2 + \alpha_3 = 0$	0.00321	0.000423	0.0213	0.0218
R.I. p-value	0.0380	0.0200	0.116	0.0820
Panel B: Share of Men				
(K/L)Shock*	.000056	.000012	.000016	.000031
3 Decades Post	(.00015)	(.00011)	(.000074)	(.000068)
	[0.097]	[0.081]	[0.000]	[0.000]
(K/L)Shock*	-.00014	-.00017**	.00013**	.000048
2 Decades Post	(.00011)	(.000077)	(.000057)	(.000046)
	[0.031]	[0.028]	[0.000]	[0.000]
(K/L)Shock*	-.000081	-.000095**	.000019	.000013
1 Decade Post	(.000056)	(.000045)	(.000029)	(.000027)
	[0.007]	[0.006]	[0.000]	[0.000]
N	96	96	96	96
$\alpha_1 = \alpha_2 = \alpha_3 = 0$	0.000	0.000	0.000	0.491
R.I. p-value	0.0700	0.0460	0.122	0.134
$\alpha_1 + \alpha_2 + \alpha_3 = 0$	0.602	0.282	0.306	0.507
R.I. p-value	0.670	0.378	0.482	0.344

Robust standard errors clustered at district level, in parentheses. Randomization inference p-values for each coefficient are in square brackets. Significance levels ***p< 0.01, **p< 0.05, *p< 0.1. K/L SHOCK is the total deferred pay returning to each district between May 1973 and November 1975 in millions of Kwacha divided by the number of miners returning to Malawi in those years. Mean of K/L SHOCK is 87.6. Total districts=24, total observations is 96 in each regression. All regressions include year and district fixed effects, and interactions of a time trend with baseline district controls (as in Table 3). Regressions in columns (2) and (4) additionally control for interactions of decade dummies with a measure of tobacco suitability of the district. Tobacco suitability is the district-level mean of the FAO measure of soil suitability for growing tobacco.

Online Data Appendix: Not for Publication

1 Census data

Our main datasets are constructed from Census data collected in 1977, 1987, 1998 and 2008. The 1977 Census data were digitized from aggregate Census reports. The 100% microdata from the 1998 Census was obtained from the Malawi National Statistics Office. IPUMSI (`https://international.ipums.org/international/`) provides provides access to the 10% sample for 1998. The 1987 and 2008 Census data are 10% samples from the IPUMSI repository.

We also use data for some outcomes from earlier Census data in 1966, 1945 and 1931. We digitized all relevant tables from aggregate Census reports in these years (Malawi National Statistical Office, 1969; Nyasaland Governor, 1946, 1931).

1.1 District boundary crosswalk: 1931 to 2008

We created a district boundary crosswalk that links district boundaries over time, through name changes and boundary changes. We use the districts existing in 1977 as the sample of districts. We consolidated information in variables from districts that had split in later years into their origin districts in 1977. For districts in earlier years that had split by the late 1970s, we apportioned the earlier cell totals to 1977 district boundaries using land area weights.

1.2 Labor market outcomes

We create three categories of labor market variables: broad sector of work variables, narrow sector of work variables, and economic activity status variables.

Broad sector of work: We define work in the agriculture, manufacturing, or service sector for each Census, using the number of people who are currently economically active (those employed and currently unemployed) in the denominator. Houseworkers and other inactive people (students, pensioners, other dependents) are excluded from both numerator and denominator of these variables. In each year, a small share of those in the labor force do not report an industry (most of these are unemployed people who have not worked before), so shares across the three broad sectors do not sum to one. For a more detailed definition of sector of work within the nonfarm sector, we disaggregate all non-agricultural employment into mining, manufacturing, retail, transport and communications, and all other services (business services, household services, and other non-specified services).

To create a summary measure of employment diversity in the district, we construct a Herfindahl index for (broad) industrial sector of work. The smaller the value of this index, the more evenly people are distributed across sectors. The larger the value of this index, the more people are concentrated within one of the three sectors.

Economic activity variables We define these variables for the sample that includes everyone in the relevant age group in a given district:

- In the labor force: working, unemployed, or doing home production
- Working: working or doing home production
- Subsistence: working as mlimi (subsistence farmer) or doing home production
- Family business worker: working in a non-farm family business
- Self-employed: working in a non-farm business for themselves
- Wage worker: working for someone else for a wage or salary
- Employer: employs other workers in a business

Economic activity variables and sector of work variables differ because the economic activity variables capture activity shares in the entire population, not just those in the labor force. Home production workers (mostly women) are excluded from sector of work variables

Table A.1: **Occupation and Industry Questions in Malawi National Census**

Census 1977	Census 1987	Census 1998	Census 2008
Sample: 10 years + answering yes to Qn. O	Sample: 10 years +, not inactive	Sample: 10 years and male, or female and not inactive (If inactive person is female, do not ask B18 and B19)	Sample: 10 years +, and ever worked (currently, or before) and currently available to work
Q: What is your occupation?	N: What is your occupation?	B18: What is this person's main occupation?	P25. What was [the respondent's] main occupation during the last 7 days or the last time he/she worked? P26. What is [the respondent's] status in the occupation? (Employer, self employed, public sector, private sector, family farm/business, other)
R: What is your industry of work?	O: What is your industry of work?	B19: What is this person's main trade or business (industry)?	P27. What is the main product, service or activity of [the respondent's] place of work?

but included in the economic activity variables. Our data show that the majority of family business workers, self-employed, wage workers and employers work in the non-farm sector.

In Tables A.1 and A.2, we compare the wording of Census questions across years. For the most part, it is possible to create a consistent set of definitions of each of the above variables, using combinations of different Census questions.

1.3 Population density and urbanization variables

We digitized population data from the 1945, and 1966 Nyalasand Census and the 1977 Malawi Census. These data were reported at district level, sometimes separately for men and women in different age groups. We combined these data with district data from the 1987, 1998 and 2008 Census, and constructed population densities at district level using the area of the district. We also measure population totals, for men and women separately, and the share of population in urban areas within the district.

1.4 Migrants at district-level

In Census 1977, the total number of men who report ever migrating from Malawi is reported at district level (Census 1977, Table 4.8) while the share of miners who returned between 1966 and 1977 is reported in national aggregate data (Census 1977, Table 4.11). To construct district-specific numbers of migrants returning between 1966 and 1977, we multiplied the share of workers who had returned to Malawi in the last 10 years (out of all ever migrants who returned to Malawi) by the total men in each district who had ever migrated for work by 1977. Because of the labor ban, all migrants had returned to Malawi by 1975 and so would have been present in the 1977 Census.

1.5 Baseline district covariates from Census data

We used Census data to generate the following variables:

Historical literacy rates: we digitized data on the district-specific share of adults who were literate from the Report on the Census of 1931 (Nyasaland Protectorate, Table 6)

Share of married men and women in 1977: we digitized data on the share of men and women married from Census 1977 (Table 2.1)

Share of men and women with no cash incomes in 1966: we digitized the district-specific rates of men and women earning no cash income from the Malawi 1966 Population

Table A.2: **Economic Activity Status Questions in Malawi National Census**

Census 1977	**Census 1987**	**Census 1998**	**Census 2008**
Sample: 10 years and older	Sample: 10 years and over	Sample: 10 years and over	Sample: Non-visitors, 6 years and over
O: Did you work last week (Y/N)?	M: Activity status in last seven days? Active: Mlimi, Employee, Family business worker, Self employed, Employer, Unemployed (Worked before and seeking/not seeking work, or never worked and seeking/not seeking work). Inactive: Home worker, Student, Dependent, Independent, Other	B17: What was X doing in the last 7 days? Active: Mlimi, Employee, Family business worker, Self-employed, Employer, Unemployed (worked before, seeking/not seeking work, never worked before/seeking work). Inactive: Non-worker: never worked before and not seeking work, home-worker, student, other	P20. Aside from his/her own housework, did X work during the last 7 days? (Y/N)
			P21. Why did X not work during the last 7 days? Inactive: Home-worker, Non-worker (never worked), On leave with job, Retired, Student, Other
P: What was your activity? Active: Mlimi, Employee, Family business worker, Self-employed, Employer, Unemployed (worked before and seeking/not seeking work; never worked before and seeking/not seeking work). Inactive: Home worker, student, dependent, independent, other			P22. Did X do one of the following activities during the last 7 days? Active: Farming/rearing animals/fishing, Production/services/selling, House worker at someone's house, Homeworker at own house, nothing
			P23. Is S available to work? (Y/N) P24. Has X been seeking work during the last 7 days? (N, Y-first job, Y-new job)

Census Final Report (Malawi National Statistics Office, Zomba: Table 21)

1.6 Physical and human capital investments and asset ownership

We measured investments in different ways, based on what information was available in at least two datasets. We used data from the 1977, 1987, 1998 and 2008 Census data as described above, and from the 1968/9 National Sample Survey of Agriculture (NSSA). The NSSA data were collected from around 5,000 households, and was designed to be representative at district-level. The part of the 1968 survey that collected these data was an income and expenditure-type survey.

Radios: The share of households in the district owning at least one radio exists in all years.

Durable housing: The share of households that lived in houses with a durable wall, durable roof, or both durable wall and roof was available in 1968 and in 1987.

Agricultural tools: The share of households with at least one panga, at least one hoe, or at least one type of livestock.

Education: The average level of education of individuals in specific age categories.

2 Administrative data

To measure flows of migrant capital, and describe the composition of miners, we collected and digitized data from the National Archives in Malawi and from The Employment Bureau of Africa (TEBA) archives in South Africa, from the Malawian National archives and Rhodes House Library at Oxford University in the U.K.

Migrant capital: Our data record the monthly flows of migrant money from South Africa to specific districts in Malawi, for the period October 1966 to November 1975. These records come from documents entitled "Attestation and Despatch Returns to the Ministry of Labour", found in Malawi's National Archives in Zomba and in the TEBA Archives at the University of Johannesburg, South Africa. To construct a time series of the flows in a consistent currency unit, we converted GBP to the Malawi Kwacha using an exchange rate of 2:1, the official exchange rate at the time the Malawi currency was adopted in 1971. Capital flows were recorded in each of three categories: deferred pay, voluntary remittances, and deposits. Our analysis uses only the deferred pay amounts that were set by contract. These flows make up 89% of the total flows of money over the period.

3 Other Geographic covariates

We also created the following variables:

Area: geographic area for 24 districts was calculated in ArcGIS.

High Malaria Area indicator: we computed altitude for each point on the Malawian grid map using data from the national map seamless server `http://seamless.usgs.gov/index.php` and the Viewshed tool in ArcGIS. We aggregated these measures to district level. Then we defined areas of high, medium or low malaria susceptibility based on standard measures of altitude: high malaria areas (altitude below 650m), medium malaria areas (altitudes between 650m and 1100m) and low malaria areas (altitudes over 1100m)

Estate indicator: We identified which districts contained a large tea or tobacco plantation using information in Christiansen (1984). The FAO's crop suitability index measuring whether a district is highly suitable for tobacco or tea production significantly predicts this estate district indicator

References

Chirwa, Wiseman Chijere. 1992."TEBA is power: Rural labour, migrancy and fishing in Malawi, 1890s – 1985", Ph.D thesis, Queen's University, Kingston, Ontario.

Crush, Jonathan. S., Alan Jeeves and David Yudelman. 1991. South Africa's Labor Empire. Westview Press.

Malawi National Statistical Office. 1970. *National Sample Survey of Agriculture 1968 and 1969: Customary land in rural areas only*. Zomba

Malawi National Statistical Office. 1984. *National Sample Survey of Agriculture 1980 and 1981: Customary land in rural areas only.*, Zomba

Malawi National Statistical Office. 1970. *National Sample Survey of Agriculture 1992 and 1993: Vol. 1, 2, 3, 4.*, Zomba

Malawi National Statistical Office. 1991. *Malawi 1987 Population and Housing Census Summary of Final Reports.*, Volumes I and II, Zomba

Malawi National Statistical Office. 1979. *Malawi 1977 Population Census Final Report*, Volumes I and II, Zomba

Malawi National Statistical Office. 1967. *Malawi 1966 Population Census Final Report*, Government Printer, Zomba

Nyasaland Governor. 1956. "Memorandum on Labor Migrancy in Malawi." 1956 Nyasaland Provincial Office Memo, December 7 1961.

Nyasaland Protectorate. 1931. *Report on the Census of 1931.* Government Printer, Zomba

Nyasaland Protectorate. 1946. *Report on the Census of 1945.* Government Printer, Zomba

Wilson, Francis. 1972. Labour in the South African Gold Mines 1911-1969. Cambridge University Press.