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

Taryn Dinkelman † Grace Kumchulesi ‡ Martine Mariotti §

May 29, 2020

Abstract

Can migrant capital contribute to long-run structural change in sending community labor markets? We study how rural labor markets in Malawi changed after exogenous shocks to international labor migration opportunities. Merging archival data on sub-national remittance flows with census data on employment, we track how work shifted across sectors in places receiving different amounts of migrant earnings as a result of the shocks. In labor markets receiving more migrant capital, workers -– particularly women – moved out of farming and into more capital-intensive non-farm service sectors over the next thirty years. High migrant capital areas accumulated more non-farm physical capital and human capital, and were wealthier fifteen years after the migration episode. Our results demonstrate that temporary international migration and the associated remittances have the potential to change rural labor markets in the long run. [135 words]

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

*This is an output from a project funded by the UK Department for International Development (DFID) and the Institute for the Study of Labor (IZA) for the benefit of developing countries (GA-C2-RA4-181). 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: gracek78@yahoo.co.uk.

§Mariotti: Australian National 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). Yet, economists continue to puzzle over persistent productivity gaps between agricultural and non-agricultural sectors in developing countries, the lack of structural transformation in many places, and what options for structural change are feasible in low-income countries.[1]

Recent empirical evidence points to technology, trade, and environmental shocks being important triggers for structural transformation in the labor market. In Brazil, the introduction of new labor-saving agricultural technologies 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 the US, labor reallocated across space and sectors in response to environmental shocks that changed relative supplies (and hence prices) of factors of production across sectors (Hornbeck and Naidu, 2014).

Our paper provides new empirical evidence on an alternative trigger for structural change as well as its potential long-run effects. We analyze a natural experiment in Malawi that exogenously and temporarily increased the local supply of capital from international migrants. We show that this temporary capital surge led to changes in the structure of local labor markets that lasted at least 40 years. In this African rural setting, work shifted from farming into services and continued to shift for decades after the surge dissipated. These results provide some of the first evidence on the importance of capital accumulation outside of a country in promoting long-run changes in local labor markets.[2]

Specifically, we exploit spatial variation in migrant money injected into local labor markets in Malawi between 1967 and 1975 in the wake of unanticipated shocks to international migration opportunities. These years demarcate historical events that first expanded and then shut down the ability of Malawians to take up lucrative jobs on South African gold mines. The first event in 1967 removed all quotas that had previously limited contract mine migration to low levels. The second was the banning of migration and repatriation of all

[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]Summarizing the cross-country literature on the effects of remittances on home country outcomes, Clemens and McKenzie (2018) conclude that economists remain "surprisingly unsure of its broad development effects".

Malawian migrants in response to a mineworker plane crash that killed traveling miners in 1974. In the eight intervening years, legal temporary migration to South Africa rose by 200% and then fell to zero. This sharp shock to 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.[3]

To study how the labor markets of migrant-sending communities changed in response to these capital injections, we build a panel dataset of money flows and district outcomes by digitizing archival material on sub-national remittances and match it to six waves of Malawian population census data. We estimate differential trajectories of economic change in local labor markets in the decades following the international migration ban, controlling for district and decade fixed effects, and allowing for time-varying effects of baseline covariates.

The thought experiment motivating our empirical strategy is the following: compare changes in labor market outcomes across two districts, each of which sends the same total number of migrants to South Africa during the migration surge, but each of which receives different total amounts of migrant money as a result. Conditional on total migrants from a district, variation in migrant capital across districts is driven by when these migrants leave, because mining wages were rising exogenously over time.[4] Our strategy amounts to comparing outcomes across districts that sent more of their migrants late, when mining wages were high (late-sending areas), and districts that sent more of their migrants early (early-sending areas) when mining wages were lower.

Our identification assumption is that conditional on total migrants sent and other covariates, late-sending districts would have evolved similarly to early-sending districts in the absence of the injection of capital. While there are no employment data to directly test this parallel trends assumption, we show that prior to the migration shock these early- and late-sending districts are comparable on baseline economic variables. They appear to differ only in their access to mining jobs through historically-placed recruiting stations. In our prior and related paper (Dinkelman and Mariotti, 2016), we established that these recruiting stations from the early 1900s were "as if" randomly located in districts, and uncorrelated with characteristics that drive local economic outcomes.

Our main finding is that after this temporary international migration episode ended, districts with larger capital injections saw a greater share of workers shift out of agriculture and into capital-intensive, non-farm service activities. These shifts began in the first ten years following the end of migration, once all migrants had returned, and continued into

[3]This amount is equivalent to 185 million USD in 2015 (Fagernas and Shurich, 2004).

[4]Changes in mining wages were driven by changes in the gold mining industry in South Africa and the world, and unrelated to events in Malawi. Workers signed up for two year contracts at a fixed rate.

the second and third decades post-shock. Men shifted out of farming first, with effects for women lagged by about a decade. Both women and men shifted into the service sector, specifically into construction and retail sector jobs. Our results are larger and most robust for women. We show that this reallocation of especially female employment triggered by the migrant capital injection was accompanied by a spatial shift in population towards market towns and urban areas within a district. Such spatial shifts are additional indications of structural change in the rural economy.

In terms of magnitudes, we find that for each additional one million USD received in the district through migration, 975 workers shifted into service sector jobs. Each additional non-farm job cost about USD1,025 in migrant capital. A back-of-the-envelope calculation suggests that the additional inflow of migrant money accounted for between 5% (for men) and 19% (for women) of the total structural shift out of agriculture between 1977 and 2008.

Our findings are robust to a range of specification checks. We find similar employment results when we hone in on the variation in migrant capital generated by migrants returning immediately after the mine-worker plane crash. This variation depends on even smaller differences in the time profile of migration across districts, and is even less likely to be plagued by concerns about migrant selection across early- and late-sending migrant districts. Our results are also unlikely to be driven by differentially positive economic growth in high migrant capital districts. Far from booming in advance of the capital injection, high migrant capital districts experienced slower population growth decades before the migration shock.

Although our analysis controls for total number of migrants from a district, we cannot separate the effect of capital injections from the migrants that carried this money back home. Nonetheless, 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. Moreover, our results are also strongest in subsequent decades for women, not for men who would have been the migrants themselves.

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. To understand more about mechanisms, we test whether migrant capital inflows directly changed patterns of investment across sectors in ways that could account for the persistence of the temporary capital shock. To do this, we assemble data on investment goods from population and agricultural censuses and household income and expenditure surveys, before and after the migration shock.

In areas receiving more migrant capital, people do not appear to invest more in agricultural investment goods in decades after the shock, but instead invest in the non-farm sector and in the human capital of the next generation of workers.[5] These investment patterns suggest that at least some of the shock to migrant capital was channeled towards higher return non-farm sectors, and contributed to the shift in workers towards the service sector.

In a final exercise documenting persistence of the original migrant capital shock, we use demographic and health survey data to show that fifteen years after the end of this temporary migration, households in districts that received more migrant capital were significantly wealthier and held more private assets.

The primary contribution of our paper is to empirically demonstrate that temporary international migration offers a route to capital accumulation that in turn contributes to long-run structural change in the labor market. 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, and the impact here is over a long-run period of forty years.[6] 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.[7]

By connecting migrant capital and labor market outcomes over time we also build on prior research in the migration literature showing that unexpected positive shocks to migrant earnings increase savings and investment in migrant households (e.g. Yang (2008)). We go beyond short- and medium-run effects on individual migrant households and estimate long-run market-level effects of migrant capital flows driven by specific migration episodes.[8]

A second key contribution of our paper is that we highlight the role that temporary international migration and associated migrant capital could play in Africa, the region with perhaps the greatest potential for structural change. Not only is migration an important feature of labor markets in low-income countries, but remittance flows to developing countries

[5]This 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.

[6]In Bustos, Garber, and Ponticelli (2020), a new agricultural technology provides the impetus for capital accumulation in the farming sector in rural areas. The authors cleverly show that these savings move through the Brazilian financial system to expand credit to the non-farm sector in urban areas, expanding employment and output in that sector over ten years. Both their paper, and ours, provide evidence that an increased supply of capital plays an important role in facilitating structural change. Interestingly, each paper finds that capital accumulation accounts for around 20% of the shift out of farm employment.

[7]Mesnard (2004) suggests that international labor migration can help individuals overcome liquidity constraints at home. Foster and Rosenzweig (2008) model how rural to urban labor migration and associated remittances facilitate structural change in the rural labor market through local demand shocks.

[8]There is a long tradition of estimating the impact of migrant money coming back into migrant households themselves, e.g. Dustmann and Kirchkamp (2002), Woodruff and Zeneto (2007), Yang (2008), Kinnan, Wang and Wang (2016). Recent work has turned to estimating immediate effects of the end of migration flows on market-level outcomes (Theoharides, 2018), and medium-run effects of positive remittance shocks on local labor markets and human capital (Khanna, Theoharides, and Yang, 2020).

are now the largest source of external funds in low-income countries, exceeding total foreign aid flows and foreign direct investment (The World Bank, 2019). Our paper complements the handful of papers estimating the short-run developmental impacts of seasonal and guest worker programs on origin countries, e.g. Gibson, McKenzie and Rohorua (2014); Gibson and McKenzie (2014); Kosack (2019). The experience of Malawi indicates 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.

The paper begins with a simple conceptual framework to fix ideas about the channels through which an injection of migrant capital might matter in rural economies. In section 3, we describe labor markets in Malawi and their evolution over time. Section 4 discusses the natural experiment and describes the variation in migration opportunities and spatial variation in receipt of migrant capital that is central to our identification strategy. Section 5 outlines our main empirical strategy and analyses datasets, and section 6 presents main results. In section 7, we marshal direct evidence on how capital injections affected accumulation behaviors across communities over time. We conclude by examining wealth profiles of communities fifteen years after the end of migration.

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

How might a temporary capital injection affect a rural economy? First, and most directly, the return of migrant capital increases local incomes of migrants. Local demand shocks generated by returning migrants increase 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). Returns to work change across sectors, bidding up the price of non-farm relative to farm work, encouraging workers to move out of farming. However, for temporary capital injections to have more than a transitory effect on incomes and relative returns to working across sectors, there should be a mechanism for capital accumulation. This is where other channels of impact become important.

Second, migrant capital provides liquidity for investment in all sectors. In the farm sector, farmers may choose to invest in farm capital such as seeds, fertilizer, or farming equipment. As long as the economy is closed (as we expect is the case for districts in Malawi), more farm capital lets farmers meet minimum food production requirements with less labor, and excess labor can be released to the non-farm sector.[9] In the non-farm sector, migrant capital can enable individuals to overcome fixed start-up costs of entering the non-farm sector, 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). If migrants are not starting businesses, 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.[10]

Third, migrant earnings could be invested in human capital (either education or health). This may directly improve worker productivity of the next generation. If literate or healthier workers have higher returns in non-farm relative to farm sectors, these types of human capital investments that boost the share of skilled workers could then shift the sectoral allocation of labor over time (e.g. as in Porzio and Santangelo (2019)).

Underlying each of these channels is the idea that there is insufficient capital in the local economy to break out of a poverty trap. Incomes are so low that savings cannot be sustained, and people cannot afford to buy anything except 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 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. As savings and investment accumulate, the impact of the initial income shock persists. In the second half of the paper we provide some evidence on these channels of persistence.

3 Evolution of rural labor markets: The case of Malawi

In the 1960s, agriculture accounted for half of Malawian GDP; by 2015, this contribution had shrunk to 30%. Manufacturing was around 13% of GDP in the 1960s, and 10% in 2015. Over

[9]For example, Robinson (2014) shows wide variation in maize prices across markets in Malawi, suggesting it is reasonable to assume that districts in Malawi function like closed economies. In addition, researchers have documented large frictions in markets for land and other agricultural inputs (seeds, fertilizers) in Malawi (Beegle, Galasso and Goldberg, 2017; Restuccia and Santaeulalia-Llopis, 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.

[10]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.

the same period, the share of services in GDP increased from around 40% to almost 55% (The World Bank, 2016). 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 shows 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.[11] We construct a Herfindahl index to describe how concentrated employment is in any one sector, for each district. Lower values of this index reflect lower concentrations, or more diversity of employment within the district. Over time, this index falls from 0.89 to 0.54 for women's employment, and from 0.62 to 0.36 for men's employment.

It is useful to fix ideas about the nature of non-farm work in rural Malawi. Figures 1a and 1b graph the percentages of workers employed in non-farm industries and occupations, for each of the top five most prevalent non-farm categories. 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 hotels and restaurants for women, and manufacturing and construction 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 follow industry patterns. Almost 30% of women working in the non-farm sector report themselves as working proprietors (self-employed). 16% are sales and shop assistants, 8% are teachers, and 7% are food and beverage producers. About 17% of men report being working proprietors and 11% are shop assistants. 11% work as brick-makers or carpenters in construction, 7% in protective services (security guards), and 5% are teachers. These charts show that the non-farm sector is dominated by work in retail trade and small-scale personal and general services, often run by the self-employed.

Several facts about how services are produced in Malawi are also worth noting. We illustrate these facts using data on rural households from a national income and expenditure survey in Table 2. First, small businesses are common in rural areas. All rural households farm, and one in five of these households also owns a non-farm enterprise. Second, although

[11] Labor force participation rates for both men and women are high in Malawi. Between 84% and 96% of working age people were working, or looking for work in different decades (results not shown). More details on the data are in the Data Appendix.

many of the service sector jobs documented in Table 1 and Figure 1 are in small-scale household enterprises, these activities use more capital on average than farming does. Table 2 shows that non-farm businesses report using between 8 and 10 times more working and physical capital (excluding land) relative to farm businesses. Third, labor tends to be more productive (on average) in the non-farm sector than in farming. In our data, average labor productivity in the non-farm sector (value added per worker) is about about three times higher than in the farm sector (Table 2, last column). 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 better off in such a job relative to the farm sector. These descriptive statistics for the rural economy form the backdrop for our analysis of what happens to work in decades following the shocks to international labor migration.

4 The natural experiment: Shocks to international labor migration

For a time during the 20th century, mine work in South Africa provided another feasible sector of work for Malawian men who were willing to migrate. In this section, we briefly describe the nature of this international migration and exogenous shocks to the opportunity to migrate that are relevant for our empirical analysis. We show how the number of migrants and migrant capital flows vary over time and space during these years. Then, we discuss why some districts received different amounts of capital as a result of the migration shocks. 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.

The demand for international migrant mine workers during the twentieth century was centrally administered by the mines' labor recruiting agency, the Witwatersrand Native Labour Association (*Wenela*).[12] Workers were employed on fixed two-year contracts. 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 (own calculations, South African Census data 1970 and 1980).

Mine work was significantly more lucrative than work at home. Miners could earn at least 2.5 times more on the mines per year, relative to the next best wage-earning job at

[12]Members of the South African Chamber of Mines gave this agency authority for recruiting mineworkers from across the southern African region. *Wenela* merged with the South African recruiting agency in the mid-1970s and became *Teba*, The Employment Bureau of Africa. Much of the archival material we use in our analysis are original documents from *Wenela/Teba* that we collected and digitized.

home: farm labor on an agricultural estate (tobacco or sugarcane). Even working on tobacco and sugar cane farms was not an option open to everyone; the majority of farmers were subsistence farmers. Because miners were contractually obliged to save up to 60% of their earnings and receive this as deferred pay upon repatriation, most of the migrant earnings were returned to Malawi. This formed the basis of their ability to accumulate capital for use in rural sending regions.

Opportunities for taking up mine work in South Africa varied over time. Figure 2a, reproduced from Dinkelman and Mariotti (2016)), shows the number of Malawian migrants employed on two-year contracts on South African mines in each year from 1950 to 1990.[13] Prior to 1967, legal migration was limited by recruiting quotas that were held at less than 2% of the working age male population. In 1967, a new labor treaty (Treaty, 1967) removed this quota and migration expanded from 40,000 to over 120,000 men in five years. The 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. The number of Malawians working on South African mines fell to zero in the two years following 1974. 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.

Flows of money paralleled the ramp up of migration in the late 1960s and the equally dramatic decline of migration in the mid-1970s. Figure 2b plots the log of total money flows documented by *Wenela* in 1975 USD and the log of total deferred pay flows back to Malawi in each year.[14] Between 1966 and 1975, total migrant capital flows rose by about 20%. After the removal of the labor quota in 1967, deferred pay flows increased slowly at first and then more rapidly, as the number of migrants surged. Money flows spiked after the plane crash (indicated by the second vertical line), when all miners were repatriated. This later part of the migration period (1974-1975) represents the period of largest, coordinated capital flows back to rural districts in Malawi.

Total deferred pay inflows over the entire period were 53 million USD. Depending on how one discounts, between one fifth and one quarter of the total amount of migrant capital returned to Malawi in the 1 to 2 years after the plane crash. At peak migration in early 1973, Malawi received 2.75 million USD from miner earnings in a single month, or almost 115,000USD on average per district. Each migrant returning from a two-year contract would

[13] Data sources are described in more detail in the next section and in the data appendix.

[14] Amounts are converted (from Great British Pounds, GBP, or Malawian Kwacha, MWK) to USD at the fixed 1975 exchange rate of 1.2 MWK to 1 USD. Total money sent back was the sum of deferred pay, voluntary remittances, and deposits.

have received between 130USD and 295USD, depending on when he left. As a benchmark, average per capita GDP in 1970 in Malawi was only 63USD.[15]

Within Malawi, these shocks to international labor migration affected different districts to different degrees. Some districts had more total migrants between 1967 and 1975 than others, and were therefore exposed to migration and migrant capital shocks of different sizes. Because migration is a choice, comparisons of outcomes across high migrant-sending and low migrant-sending areas are likely to be confounded by factors that drive different levels of migration across districts.

Instead of comparing high and low migrant-sending regions with each other, we compare districts that sent the same total number of migrants, but at different times during the 8 year period. A key observation for our work is that *different time profiles of migration* generated spatial variation in the amount of capital received by each district. Specifically: the year and month in which a migrant left a district affected total deferred pay eventually received by the district because mine wages were rising over time. Between 1966 and 1974, nominal mine wages had increased by over 200% due to changes in the gold price.[16] Districts sending more migrants later in the period would therefore have received more money per miner. Furthermore, districts with more migrants closer to the end of their work contracts in 1974, at the time of the labor ban, would have received more deferred pay after the ban.[17]

An obvious question is: why would some districts send more migrants earlier in the 1967-1974 period? What is different about early-sending districts and late-senders? Insights from a simple migration model are useful here. Workers typically choose to migrate when the return to migration is higher than the costs of migrating. The higher the migration cost, the higher the foreign wage needs to be to induce migration. Late-sending districts are likely those where workers faced the highest costs of migrating. In these areas, the mining wage had to rise high enough to induce these districts to send more workers.

In Dinkelman and Mariotti (2016), we discuss the various costs associated with getting a mining job and explain that having a *Wenela* recruiting station in one's district significantly lowered the cost of getting a mine job. In that paper, we explain how obtaining a mine job

[15] We take the total deferred pay flowing back to Malawi in each year from 1967 to 1975, and conservatively divide this by the number of migrants employed in South African mines two years prior (to account for the two year contract), to come up with this range. This is almost surely an underestimate of the total capital per returning miner, because it assumes that each migrant had only one migration episode.

[16] In the mid-1970s and early 1980s, the mining industry chose to raise mining wages for political economy reasons. Such increases were made possible by large increases in the global gold price (Crush, Jeeves and Yudelman, 1991).

[17] Districts with a larger share of migrants on repeat contracts would also have received more capital since miners earned a small raise each time they re-contracted. We do not have information on the share of repeat migrants, however, given that most Malawian miners were under age 30, it is unlikely that variation in shares of career miners over districts drives the bulk of the variation in total deferred pay flowing back to Malawi.

entailed significant travel and administrative costs, implying that those living closest to a recruiting station had the largest incentive to take such jobs conditional on the wage. We show that the placement of these stations in the early 1900s was unrelated to key historical economic variables of the district, including changing educational profiles of each district prior to the migration shock.

This fact suggests the following test of our hypothesis that migrants in late-sending districts faced higher costs of migrating to South Africa: regress the share of migrants leaving from each district in each month and year on whether the district had an historical recruiting station, interacted with time. We do this in Table 3, also controlling for baseline district characteristics, and district fixed effects.[18] The results in this table show that the presence of a recruiting station, which lowers the cost of migration, affects the timing of migrant flows. In the first column, we see that in all districts, a larger share of migrants are sent in later periods. The second and third columns show that this relationship is robust to controlling for district-level observables measured before the international migration shock, and for time-invariant district-specific unobservables when adding district fixed effects. The fourth column directly tests our hypothesis. The share of migrants coming from each district is increasing over time and this increase is faster (slower) for districts without (with) a pre-existing recruiting station. Districts facing higher costs of getting a migrant job were sending relatively larger shares of migrants later in the period when wages were highest.

The main variation we leverage in our empirical work is this spatial variation in the size of the migrant capital shock. It is important to note that conditional on the number of migrants leaving a district, baseline district-level characteristics cannot predict differences in money received by districts. These baseline district covariates include literacy rates and population density in 1931, the share of men and women not earning any cash income in 1966 as a proxy for the local economy prior to the migration shocks, the share of men and women married in 1966, an indicator for whether the district is a malaria area, and two region indicators. Each of these variables is plausibly relevant for local labor market performance and structure over time. Table 4, column (1) shows the correlation between total deferred pay received by a district and these baseline district-level variables. In Table 4, column (2), we take out common month and year fixed effects from total deferred pay going to a district, then aggregate the residuals to district-level and correlate this residualized migrant capital variable with total migrants and baseline district-level variables. Results are

[18] We do not have data on number of migrants returning each month to each district. We impute the number of migrants leaving each district in each month by dividing the district-year-month-specific migrant capital inflows by 60% of the annual wage rate (deferred pay shares were fixed at 60%). Because the specifics of mining contracts did not vary by district, there is no reason to expect measurement error to invalidate our imputation.

similar across residualized and non-residualized specifications. Results in this table reassure us that districts receiving different levels of migrant capital, conditional on the total number of migrants, are not different on observables in the pre-period.

5 Empirical strategy and Data

5.1 Estimation

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

$$Y_{dt} = \sum_t \alpha_t K_d * Decade_t + \sum_t \beta_t L_d * Decade_t + \kappa_t + \delta_d + W_d Trend_t \lambda + \epsilon_{dt} \qquad (1)$$

where Y_{dt} is, for instance, the share of workers in agriculture, manufacturing or services, d is the district, t is the decade (1977, 1987, 1998 or 2008), K_d is the amount of deferred pay in millions of USD received by district d between 1966 and 1975 (the total capital shock), L_d is the total number of migrants in thousands returning to each district between 1966 and 1977 (the total migrant shock), $Decade_t$ 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.

This specification allows us to estimate the effect of the capital shocks 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 earning cash wages to evolve differently over time. We isolate the impact of K_d across districts sending the same number of migrants by controlling for L_d.

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 an additional one million USD was received per district, while α_{1998} and α_{2008} provide the same parameter for the 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 also 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 (1) using a difference-in-differences design.

It is important to note that we do not pursue an instrumental variables strategy here. As we show in Table 3, having a recruiting station predicts the time profile of migrant flows, and hence migrant capital inflows. However, we cannot use recruiting stations to instrument for the capital shock because this variable predicts both the total flow of migrants (positively) out of a district as well as the time profile of migration (whether districts are early-senders, and hence receiving less migrant capital conditional on migrants). Instead, we control for high and low-migration districts directly, and compare remaining variation in migrant capital generated by differences in the time profile of migration *within* high- and within low-migrant sending districts.

5.2 Identification assumptions and threats to validity

For all outcomes, our identification assumption is that in the absence of the capital shock, districts sending the same number of migrants but receiving different amounts of migrant capital would have evolved in the same way after 1977, conditional on the other covariates. Because the spatial variation in migrant capital derives from differences between early- and late-sending districts, we need to assume that the time profile of migration out of a district is not confounded with systematic differences in types of migrants leaving and returning to Malawi at different times. The biggest threats to validity are that: (i) districts sending more of their migrants early are systematically different from districts sending more of their migrants late; and (ii) different types of migrants are leaving from early- versus late-sending districts.

We have already shown in Table 4 that districts receiving more migrant capital are not statistically different from each other on a range of baseline characteristics that could be important for evolving local labor market outcomes. We implement two additional tests to support our identification assumptions. We re-estimate (1) replacing K_d with $K_{d,post-crash}$, the migrant capital returning to Malawi immediately after the plane crash. Focusing on this component of migrant capital allows us to exploit an even more subtle difference between districts: districts with a greater share of migrants closer to the end of their contracts in April 1974 would have received more capital than districts with a greater share of migrants closer to the start of their contracts *in the month and year* of the crash. Arguably, cross-district differences in the timing of migration just around the April 1974 threshold are much less likely to be driven by differences in local labor market conditions in Malawi or by differential migrant selection across districts. Finding similar results for both measures of capital flows, K_d and $K_{d,post-crash}$, goes a long way towards assuaging concerns that the structural change results are driven by differential migrant selection over time.

Second, we check for differences in population growth across areas with more versus less migrant capital, before the migration episodes occur. We do this because population growth is an interesting outcome to track economic development, and because it helps us rule out that high migrant capital areas were already growing faster before the shock. These two additional results help to build 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.

5.3 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 digitized historical Census data available at the district-gender level from 1945, 1966 and 1977 and matched this with Census data from 1987 (the 10% sample), 1998 (100% sample) and 2008 (10% sample). Details of variable construction are in the data appendix.

Key labor market variables are defined for men and women using labor market questions that remain largely 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.

In addition to the Census data, we collected and digitized material from *Wenela* administrative records that capture 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 converted to USD, aggregated over time to district-level, and scaled so that one unit of K_d represents a one million dollar transfer over this period. The average transfer to a district over eight years was 2.25 million USD.

We do not have data on district-specific numbers of migrants leaving or returning each year between 1967 and 1977. We estimate the district-specific number of migrants in this decade by multiplying the national share of ever migrants who report returning to the country between 1967 and 1977 by the district-specific number of men who report ever migrating at all.

Means and standard deviations for the main covariates in our analysis dataset are in Online Appendix Table 1. The important takeaway from this set of summary statistics is that international migration was common in Malawi. By 1977, on average 36% of all adult men in a district had worked abroad, earning on average 130USD (or two times Malawi's GDP per capita) in the eight years following 1967. There is substantial variation in deferred payment amounts across districts: the standard deviation in this variable is 178USD. Per person in an average district, migration generated an additional 24USD or about 3USD per year. With per capital GDP in Malawi in 1970 at 63USD, this means that the average person (regardless of migrant status) in an average district received a 5% increase in money available to them as a result of deferred pay.

6 Main results

6.1 Structural change in high versus low migrant capital districts

Figures 3 (for women) and 4 (for men) use the raw Census data to illustrate broad shifts out of farming in high and low migrant capital districts. In each figure, we plot the average share of workers in the agricultural (top panel) or services (bottom panel) sector in each decade in districts receiving above median (solid line) or below median (broken line) levels of deferred pay. We do not control for any variables in these figures.

Shares of men and women working in agriculture start out roughly the same in high and low capital shock districts in 1977, two years after all migrants have returned home (Figures 3a and 4a). By 1987, these districts begin to diverge, with the shift out of agriculture occurring faster in the high capital shock places. Patterns in the service sector are reversed: while high and low capital shock districts start out at similar rates in 1977, by 2008, there are much larger shares of workers in services in the high capital shock districts (Figures 3b and 4b). Labor shifts into services to a greater extent in districts with the larger capital inflows. The effects of the capital shock persist, and grow larger, over time. Men shift out of farming first, with women lagging this shift by about a decade.

6.2 Impacts on broad sector of work

Table 5 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 district fixed effects, decade fixed effects, and interactions of all baseline district-level controls with a trend term. The first column for each outcome presents estimates excluding controls for the number of migrants in each district. The second column includes these controls. Robust standard errors are clustered at district level. We report statistical significance using p-values from the small sample t distribution to account for the small number of districts (24). We also show p-values generated using randomization inference in square brackets.[19]

Districts with larger capital flows 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. Effects are still present, although attenuated, when we control for the number of migrants carrying back this capital. For all sectors of work in Panel A, point estimates are larger and more significant for women.

For each additional one million dollars that flowed back to a district by 1977, the share of women working in agriculture fell by 0.39 percentage points in the first decade following the shock, by 1.14 percentage points in the next decade, and by 1.4 percentage points by the third decade after the shock. These effects fall by 20-50% when we control for the number of male migrants. In column (2), the share of women in agriculture fell by 0.2 percentage points in the first decade post-shock and 0.9 percentage points in the second and third decades after the shock. Column (4) shows smaller shifts of women into manufacturing (0.1 percentage points or lower) and column (6) shows larger shifts into services (0.5-1 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

[19] Young (2019) argues that p-values generated through randomization inference are more reliable than inference based on robust clustered standard errors in situations with few clusters.

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 0.2 percentage points in the first and second decades after the shock. Shifts into the service sector were largely positive. Male employment in services increased by a significant 0.3 in 1987, although the effects in the first and third decades were more muted.[20]

Overall, the employment shifts in response to the capital inflows are positive, and persistent. They suggest some measure of structural change, especially for women's work, facilitated by exposure to labor migration opportunities. We can 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). If we sum up our estimates of the α's in Table 5 column (2) (-0.002 $-$ 0.00928$-$0.00915 =-0.02053) and multiply by the average amount of migrant capital received by a district (2.25 million USD), we estimate that migrant earnings in an average district accounted for about 19% of the structural reallocation of female labor out of farming and into non-farm work. The contribution of migrant capital to reallocation of male labor across sectors is smaller, at around 5%. Noting that an average district had 58,000 women and 57,000 men, these percentages translate into an additional 907 female jobs and an additional 68 male jobs in the service sector (975 total new service sector jobs). The average cost of one service sector job was around 1,025USD.

6.3 Isolating variation in migrant capital following the plane crash

In Table 6, we show results for sector of work for women (Panel A) and men (Panel B), using $K_{d,post-crash}$ as the main treatment variable. All other controls are included in each specification. Coefficients in this table should be scaled to (multiplied by) the mean of capital inflows during this time, 0.6 million USD per district. We see very similar patterns of structural change as in Table 5: women shift out of agriculture and into services, as do men. The shifts for women start later, and continue for longer, than the shifts for men. Results for women are again stronger than for men. Overall, our results using the full amount of

[20] In the Robustness Appendix, we show that our main results – especially for women – are stronger when we omit the district fixed effects and baseline controls interacted with trend terms (Appendix Table 3). The point estimates are similar (and larger) when we exclude Lilongwe, the capital, although estimates become more noisy (Appendix Table 4). We show that the functional form of the control for number of migrants in equation 1 does not drive our results: results are similar when we include decade dummies interacted with migrants and migrants squared, or when we interact these decade dummies with the log of migrants (Appendix Table 5).

migrant capital are quantitatively similar to results using only the post-plane crash amount of migrant capital.

6.4 Impacts 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 7 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).[21] 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 (1).

These estimates indicate that men and women experience similar reallocations across sub-sectors of service employment. For each additional million USD received before 1977, the share of women working in construction rose by between 0.04 and 0.1 percentage points, in general services by 0.2 percentage points, in retail by 0.4-0.8 percentage points, and in transport by 0.01 percentage points. Relative to mean levels of employment in each non-farm sub-sector, the largest increases for women were in construction and retail services. For men, more capital in the district shifted work out of manufacturing (0.2-0.5 percentage points) and transport (0.2 percentage points) and into construction (0.2-0.5 percentage points) and retail (0.08-0.5 percentage points). The largest relative shifts for men were towards construction and retail. For women, the impacts of the capital shock on movement into construction, services and retail persist and grow larger over time.

6.5 Impacts on type of worker

Table 8 further investigates whether the capital infusion affects the types of workers found in local labor markets. We estimate (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 working for a wage. Panel A shows how access to migrant capital affected the employment situation of women, Panel B for men.[22]

[21] 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.

[22] 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.

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 have a smaller share of (male and female) wage workers, and at the same time, a larger share of women and men 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.

6.6 Impacts 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. We control for number of total migrants between 1966 and 1975 in all regressions. 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 are included, and standard errors are estimated as before.

Figure 5 plots the estimates of α_t, the relationship between the amount of migrant capital received by each district between 1966 and 1977, and district-level population outcomes before and after the migration surge. Standard errors bars are included, and the omitted category is 1945. Each point on the line represents the marginal impact of receiving one million USD of migrant capital between 1966 and 1977, on the log of the age-group specific population counts in the district in each Census year (Figure 5a) or on the share of the population living in urban areas in the district (Figure 5b). The regression tables underlying these figures are in Online Appendix Table 2.

The figure shows that districts that were going to receive an extra one million USD in migrant capital between 1966 and 1975 did not look significantly larger (population-wise) compared with districts that were about to receive less migrant capital, prior to 1966. Instead of being on a growth trajectory before the migration episodes, districts about to receive more deferred pay were shrinking prior to 1966. Between 1966 and 1977, this pattern reverses. Districts receiving more deferred pay by 1977 experience significant population growth by 1977. This growth is sustained over the next decades, and is just picking up male migrants returning home. The largest significant effects are seen for kids under age 5, perhaps suggestive of a small baby boom post-migration shock. Although confidence intervals are wide, the impact of the capital shock on population is persistent and positive. By the end of the labor migration period, total population in the high capital inflow districts had increased by just over 3 percentage points.

In Figure 5b, we see there was also a spatial reorganization of population within districts in the wake of the shock. By 1977, districts that received one million USD more than other districts had 0.005% more of their population living in an urban area, a 7% increase off the baseline urbanization rate. 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 The persistence of accumulation

A key question is how the capital shock could have led to shifts out of agricultural work and into the service sector over the long run, especially for women? In this section, we explore three channels. First, we look for evidence of investments in farm capital. Second, we examine what happened to physical investments in the non-farm sector post-shock. Third, we examine impacts on human capital investments of the next generation of workers. Farm and non-farm investment outcomes are measured at district-level across different Census years, with most of these outcomes measured before and after the migration episode. Human capital outcomes are measured at district-cohort level, where cohorts come of age for primary school in years either before, during, or after the migration shock.

We estimate regressions that take the form of equation (1), but where t now includes observations from before the migration shock. These regressions are therefore difference-in-differences specifications.

7.1 Long-run investments in farm and non-farm physical capital

In Table 9, Panel A, we look at how ownership of productive farm assets changes after the migrant capital shock. We measure ownership of hoes, pangas (machetes), any livestock, and oxcarts. Outcomes are taken from the National Sample Survey of Agriculture in 1968 and the National Household Income and Expenditure Survey data in 1998 and weighted up to district level using sample weights (see Data Appendix for details). For hoe, panga and livestock outcomes, we measure the share of households owning any of these items before, and twenty years after, the migration shock. Oxcart ownership is only measured in 1987, 1998 and 2008.

We find no evidence that districts receiving more migrant capital were differentially investing in farm-specific capital over time. This lack of impact on farm investments lines up with the low levels of physical capital in farm businesses observed in Table 2.[23]

In Table 9, Panel B, we examine changes in ownership of assets that may be used in non-farm work. We measure the share of households in the district in a given year that own a radio, have durable walls, have a durable roof, and have both durable walls and a roof. Radio ownership is measured in all Census waves, while the other outcomes are measured in a subset of years. Years of data included in the regressions are noted in the table. All outcomes in this panel are measured prior to the migration shock in 1968/9 and in some years after the shock.

The table shows that ten years after the capital shock, districts receiving the largest amounts of deferred pay saw increases in the share of households with a radio in 1977 (1.4 percentage points). The radio ownership gap between districts fades over time. Districts receiving more capital also see increases in the share of households with durable walls and roofs, and a durable roof alone. The share of houses with better quality housing increased by 1.3 percentage points, or by 10%, in the decade following migrant capital inflows. This result ties in nicely with the result from Table 7 that in districts receiving more migrant capital, larger shares of men and women were employed in the building and construction sector. Note that for many types of self-employed work in the service sector, a more durable home may also represent an important investment in protecting inventory (for retail trade) and/or for offering services (e.g. as a restaurant, bar, hairdresser etc).

7.2 Long-run impacts on human capital investments

In Dinkelman and Mariotti (2016), we showed that districts more exposed to the total migration shock through the location of the recruiting stations invested more in human capital of those who were children at the time of the migration shocks.

In Table 10, we explore a parallel analysis in the current setting, relating educational attainment to the inflow of money from migration. We focus on total accumulated education of adults aged 20 to 65 in 1998, some of whom would have been eligible for schooling during parts of the migration surge. We ask whether human capital is higher for cohorts living in districts that received the largest capital inflows during their years of primary school eligibility. We estimate the following regression for education outcomes of cohort c in district d ($Educ_{cd}$):

[23] One caveat to our results is that data on inputs like fertilizer, hybrid seeds, or type of crop planted are not available prior to the early 2000s.

$$Educ_{cd} = \gamma_1 K_d * EarlyTreated_c + \gamma_2 K_d * LateTreated_c + \gamma_3 K_d * PostTreated_c + \sum_c \rho_c L_d * C_c + \phi_c + \mu_d + W_d Trend_c \sigma + \omega_{cd} \quad (2)$$

We exploit an additional piece of time variation (related to timing of birth) to check whether cohorts eligible for primary schooling between 1966 and 1973 (the $EarlyTreated_c$ cohorts) had more education in the long run, if they were in high capital inflow districts. We perform the same check for differences in education levels across cohorts eligible for primary schooling in the 1974-1977 period ($LateTreated_c$ cohorts) and for those eligible at the end of the shock, 1977-1980 ($PostTreated_c$ cohorts), across high and low capital shock districts. Our control cohorts (omitted cohort category) are those eligible for schooling before 1966. All other controls in W_d are the same as in equation (1), including district fixed effects, μ_d.[24] We add cohort dummies C_c and interactions of cohort C_c with the number of returning migrants L_d. The idea of this specification is to compare individuals who were and were not age-eligible to be in primary schooling at the time of the shock, across districts that received more versus less migrant money, controlling for number of migrants. This regression is estimated separately by gender.[25]

Results in Table 10 show that for each additional million USD received by the district, female education rose by 0.1 years in early treated cohorts, 0.14 years in late treated cohorts, and 0.12 years in post- treated cohorts. Male education rose by 0.1 years, 0.13 years, and 0.11 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 migration surge. Although the estimates are not statistically different from each other by gender, females start at a much lower base level of education: the mean educational attainment for females is about half that of males. This means the average treatment effect size is larger for females. The average effect of migrant capital on female level of education is a 6.4-7.4% increase, while for males the increase is 3-4%. Exposure to more migrant capital also raises female enrollment in primary school by 2.8-3.7%, with smaller (1%) increases in enrollment for males. The most robustly significant estimates are those for the youngest females and males, eligible for primary schooling just after the end of the migration shock. Because human capital is a long-lasting asset, we interpret these results as evidence that migrant money was not simply consumed, but also invested.

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

[25] In Dinkelman and Mariotti (2016), we use the Census 1998 data for cohorts ages 20 to 44 and Census 1977 data for cohorts ages 40 to 55 because there is evidence of differential mortality of men at older ages in the 1998 Census.

Micro-economic evidence suggests that people with more schooling are more likely to work off-farm relative to on-farm (for example: Huffman (1980) for the historical US, Yang (1997) for China and Fafchamps and Quisumbing (1999) for Pakistan). In the Malawian setting, 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. Since average effects for females are larger, the effect of migrant capital on education could be a key channel for the structural change in employment we estimate for women. And, because human capital takes time to produce, this investment would only be able to contribute to structural reallocation in the later parts of our study period.

To get a sense of how much the additional human capital might have contributed to the shift in sector of work out of agriculture and into services, we focus on the changes between 1998 and 2008. We divide our estimate of the impact of an additional one million USD in migrant capital on the percentage decline in workers employed in agriculture (for women, the share of workers in farming falls by 0.92 percentage points in Table 5 column (2) row (1)) by our estimate of the impact of an additional one million USD in migrant money on the percent gain in years of education of the most exposed adult cohorts (the *Late Treated* cohorts gain an additional 0.14 years of education in Table 10 column (1); relative to average levels of education this is a 7.4 percent gain). The human capital gained as a result of the migrant capital inflows is therefore estimated to have contributed about 19% (0.92/0.074) towards the total shift out of agriculture among women in the later part of our period. This contribution is smaller for men, who experience a smaller shift into the service sector in the same period. Based on this back of the envelope calculation, human capital investments triggered by migrant capital inflows accounted for at most one fifth of the shift out of farming for women.

7.3 Did migrant capital injections improve well-being?

We have shown the extent of structural change in rural households resulting from migrant capital triggered by the migration shock. We have also shown the impact on household physical capital and human capital investments. Whether this structural transformation translated into improvements in well-being depends on whether the movement of labor across sectors allowed people to earn more for their labor. We turn to the 1992 Malawi Demographic Health Survey (DHS) to ask whether districts that received more capital after the migration shock were eventually wealthier than districts that received less capital. This allows us to test whether there was ultimately any benefit to this structural transformation for Malawians.

We estimate the cross-sectional correlation between measures of household wealth in 1992 and the amount of migrant capital received by each district, controlling for an increasing number of covariates. A limitation of this analysis is that we can only measure wealth outcomes in the cross-section. We therefore need to assume that baseline district-level controls (historical population density, historical literacy rates, a malaria risk indicator, the share of men and women married in 1966 and working for no wage in 1966) and region fixed effects adequately account for any differences in initial wealth conditions across districts. To more sharply isolate the impact of unexpected injections of migrant capital, we focus on the migrant capital returning immediately after the plane crash and estimate the following:

$$HHAssets_{hd} = \gamma_1 K_{d,post-crash} + \rho L_d + W_d \sigma + \mu_r + \omega_{hd} \tag{3}$$

where W_d and L_d are as before, and $K_{d,post-crash}$ is the amount of capital received after the plane crash. μ_r is a region fixed effect and 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.[26]

Table 11 shows that larger injections of migrant capital after April 1974 are strongly and significantly correlated with households being wealthier fifteen years later. Each cell in the table presents the estimated coefficient on the migrant capital variable (measured in millions of USD received post-plane crash) and its associated standard error, along with the number of observations in each regression in brackets. In the first column, we show the estimated coefficient from regressions of each outcome on the migrant capital variable with no other controls. The remaining columns show how these estimates change when we add additional controls. In column (2), we add the number of migrants; and in columns (3) and (4) we add region fixed effects and baseline district-level controls. We do not control for household-level characteristics (e.g. age, education and gender of household head, and household size), as these are potentially endogenous variables.

Communities that received a larger injection of migrant capital appear wealthier in the long run. An additional one million USD received by the district increases a household's wealth index by almost 0.2 standard deviations (column (1)). This effect is robust to including controls for number of migrants. With all controls included (column (4), the estimate falls only little, to 0.16 standard deviations.

The same patterns are observed for most of the other outcomes. Households in districts that received more migrant capital have about 0.25 more total assets (on a mean of 3.5

[26]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`

total assets in 1992). Homes in these areas are of higher quality (more houses have durable roofs and durable floors), results that reflect our earlier findings on the investment channels. Households in these districts are also 2 percentage points more likely to have electricity, 4 percentage points more likely to have a radio, and .8 percentage points more likely to have a car. Across the board, the results are robust to including all additional controls.

Interestingly, households in areas receiving more migrant capital are not differentially likely to have access to utilities most often provided by the government, like improved toilets and improved water sources (Table 11, column (4), last two rows). Although the government was spending on infrastructure projects in the post-migration period, these projects were apparently not targeted towards areas that received the largest amounts of migrant capital fifteen years earlier. Instead, we see that additional migrant capital returning to districts translated into higher privately-held household wealth. Taken together, these results suggest that the initial migrant capital injection did lead to lasting improvements in economic well-being in Malawi.

8 Conclusion

We marshal historical data from Malawi and exploit a natural experiment that shocked access to international labor migration 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, conditional on total migrants, 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 from migration. 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 unknown 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. Empirically measuring economic impacts of these capital flows requires outcomes to be measured over a long enough period of time. 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

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.

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

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.

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

Fagernas, Sonja and Cedrik Schurich. 2004. "The Fiscal Effects of Aid in Malawi", *Economic and Statistics Analysis Unit Working Paper*, Overseas Development Institute/Department for International Development, September.

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

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

Hornbeck, Richard and Suresh Naidu. 2014. "When the Levee Breaks: Black Migration and Economic Development in the American South", *American Economic Review*, 104(3): 963-990

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, Caroline Theoharides, and Dean Yang 2020. "Abundance from abroad: Migrant earnings and economic development in the Philippines", Working paper.

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. 2019. "Guest Worker Programs and Human Capital Investment: The Bracero Program in Mexico, 1942-1964" *Journal of Human Resources*, forthcoming

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

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.

Mesnard, Alice. 2004. "Temporary migration and capital market imperfections", *Oxford Economic Papers*, 56: 242-262.

Porzio Tommaso and Gabriella Santangelo. 2019. "Does schooling cause structural transformation?", December. `https://www.dropbox.com/s/vecp5pfxcifnw35/PorzioSantangelo_2019.pdf?dl=0` Accessed May 2019.

Restuccia, Diego and Raul Santaeulalia-Llopis. 2017. "Land Misallocation and Productivity", Working Paper.

Robinson, Amanda Lea. 2014. "Internal Borders: Ethnic Diversity and Market Segmentation in Malawi", Working Paper. Accessed December 2015.

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

Theoharides, Caroline. 2018. "Banned from the Band: 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.

Figure 1a: Most prevalent industrial classifications, non-farm sector 2008

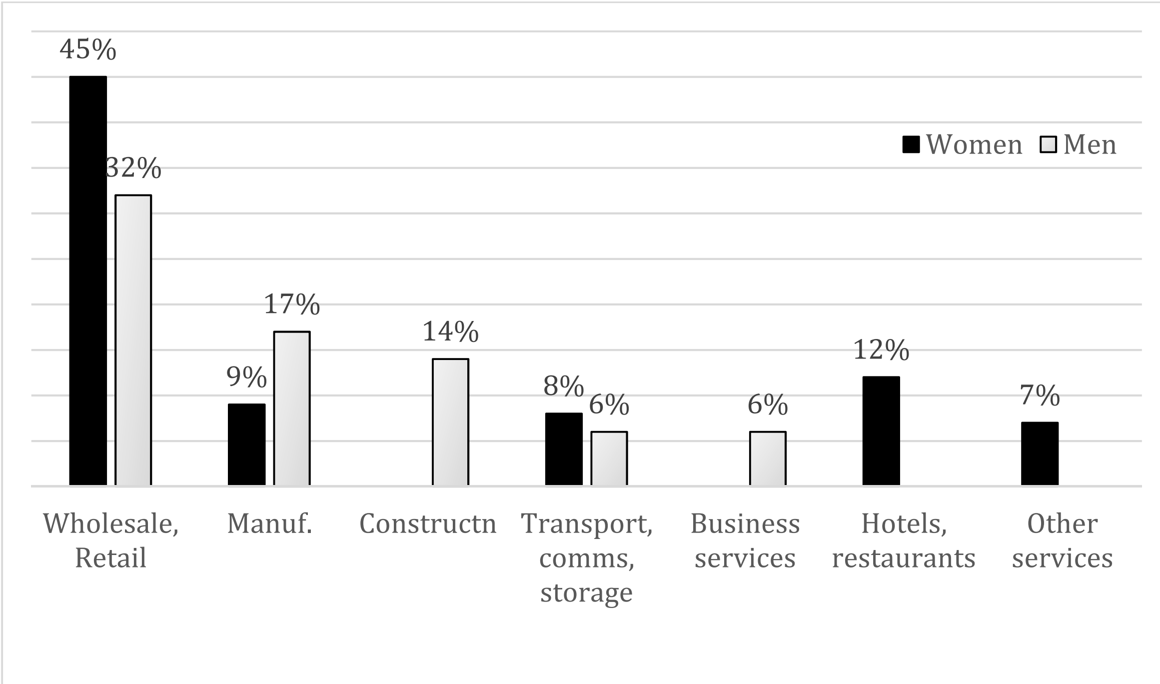
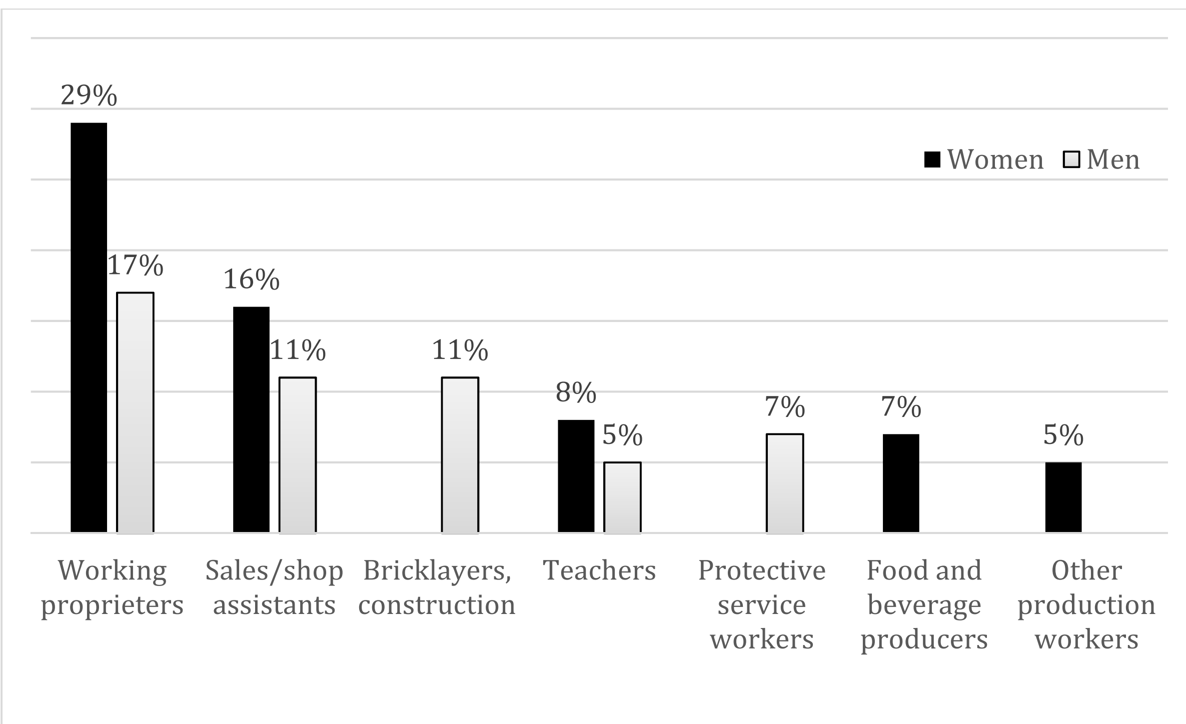
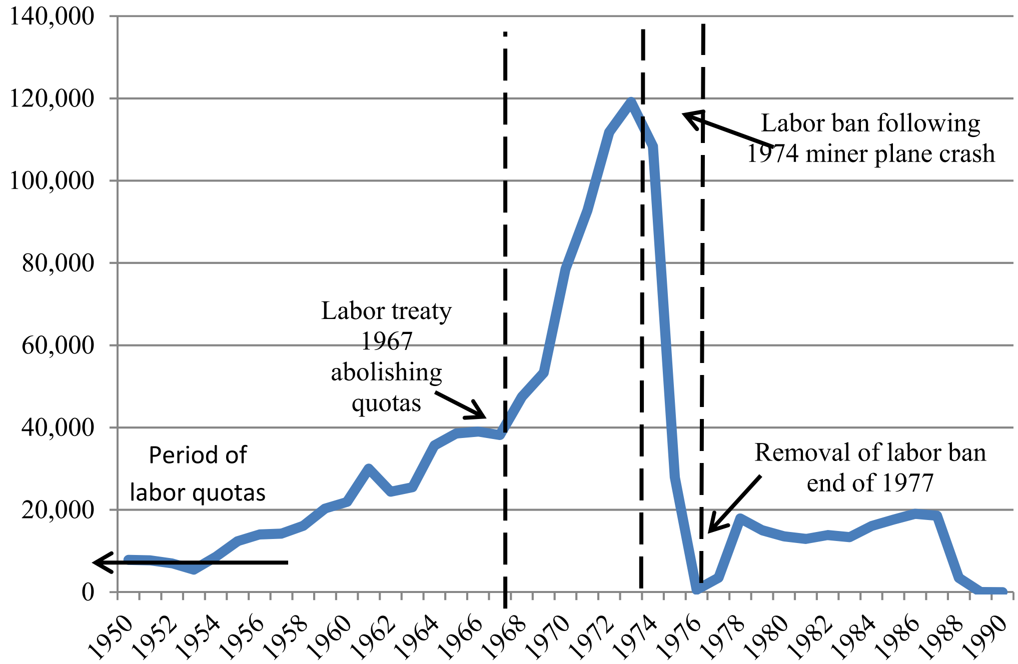


Figure 1b: Most prevalent occupation classifications, non-farm employment, 2008



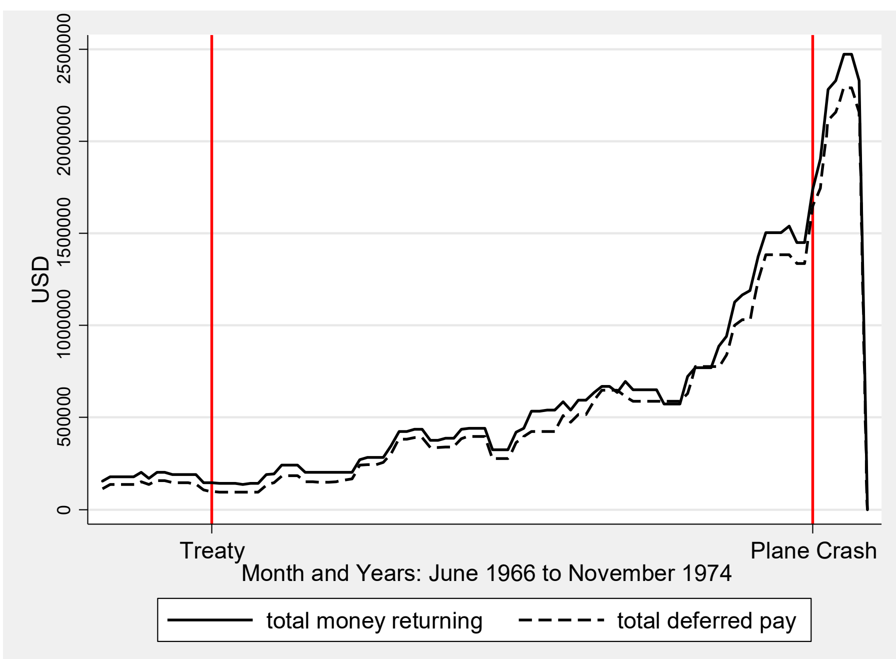
Notes: Figures indicate the share of men and women employed in non-farm industries (top figure) or non-farm occupations (bottom figure) using two digit industry and occupation classifications in the 2008 Census.

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



Source: Dinkelman and Mariotti (2016). Figure 2 shows number of workers contracted by Wenela to work on South African mines in each year. The three dotted lines represent (from left to right) the abolition of labor quotas in August 1967, the moratorium on migration after the April 1974 Malawian plane crash and the legal resumption of mine migration in 1978.

Figure 2b: Migrant capital flows over time, 1966-1975



Source: Archival material collected by the authors

Figure 3a: Sectoral shifts in the labor market: Women in agriculture

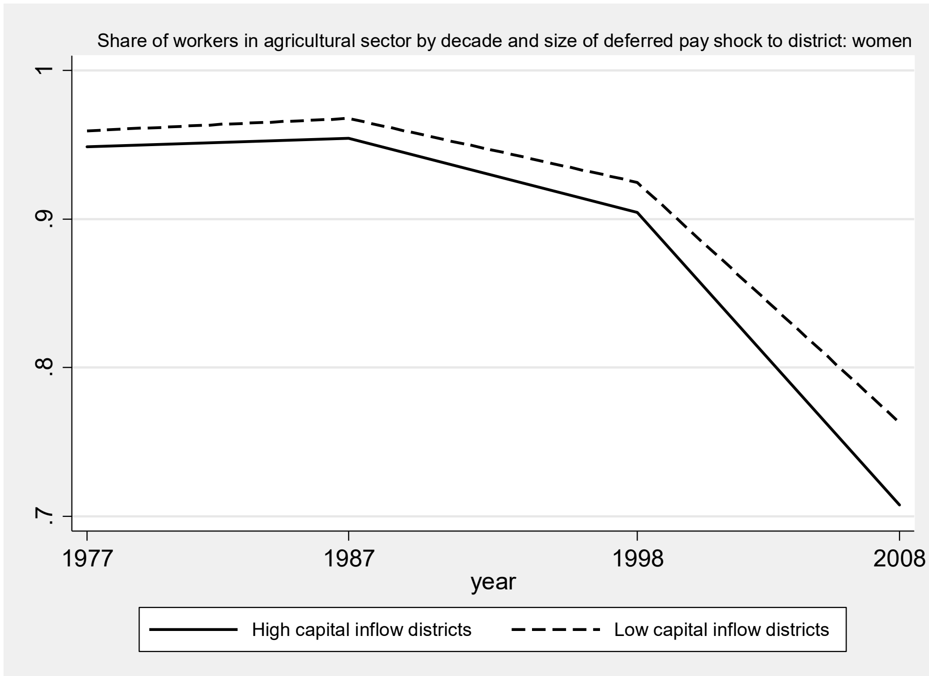
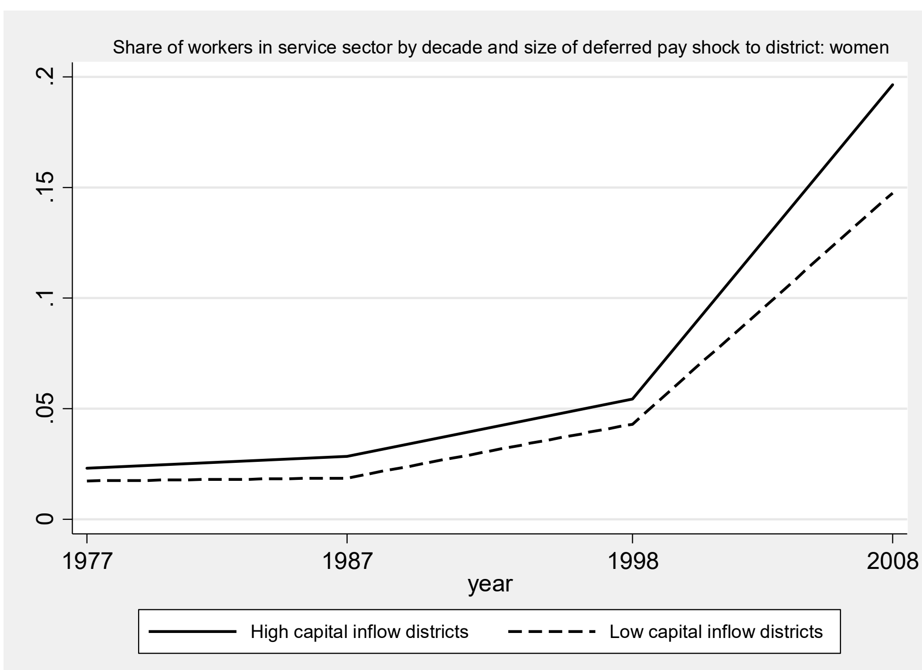


Figure 3b: Sectoral shifts in the labor market: Women in services



Notes: Share of employed women working in agricultural (top) or services (bottom) sectors over time and by type of district. High capital inflow districts are the districts receiving above median levels of migrant deferred pay before 1977.Low capital inflow districts are those receiving below median levels of deferred pay. Means are weighted using Census weights.

Figure 4a: Sectoral shifts in the labor market: Men in agriculture

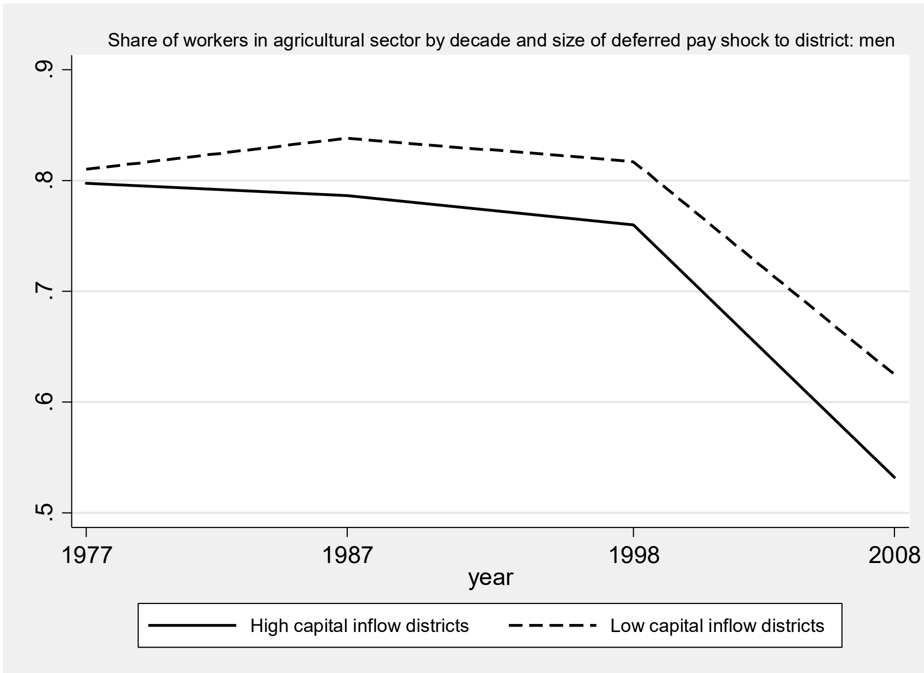
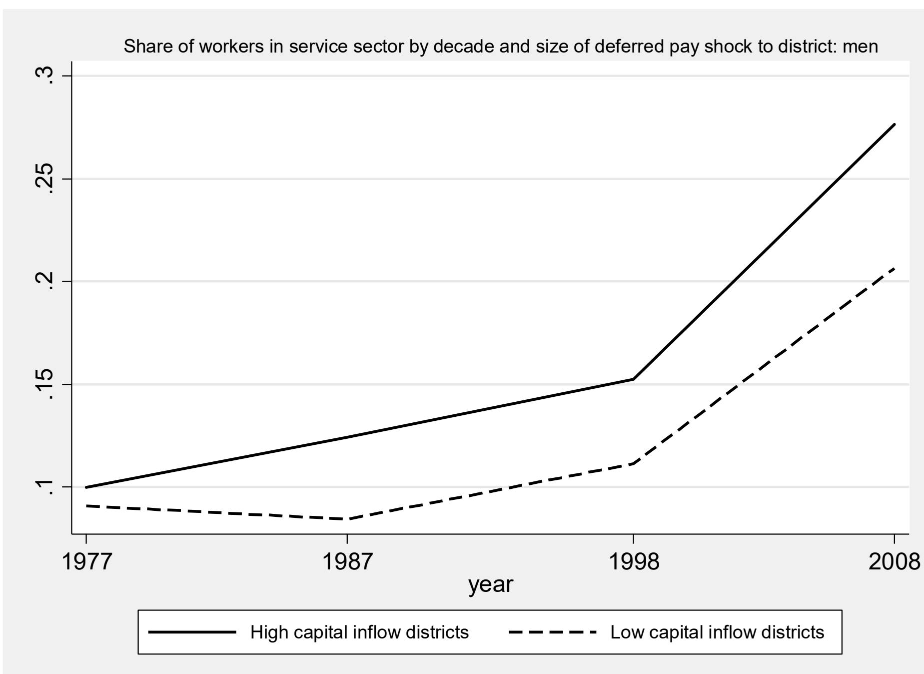


Figure 4b: Sectoral shifts in the labor market: Men in services



Notes: Share of male workers in agricultural (top) or services (bottom) sectors over time and by type of district. High capital inflow districts are the districts receiving above median levels of migrant deferred pay before 1977. Low capital inflow districts are those receiving below median levels of deferred pay. Means are weighted using Census weights.

Figure 5a: Effects of migrant capital inflows on population growth

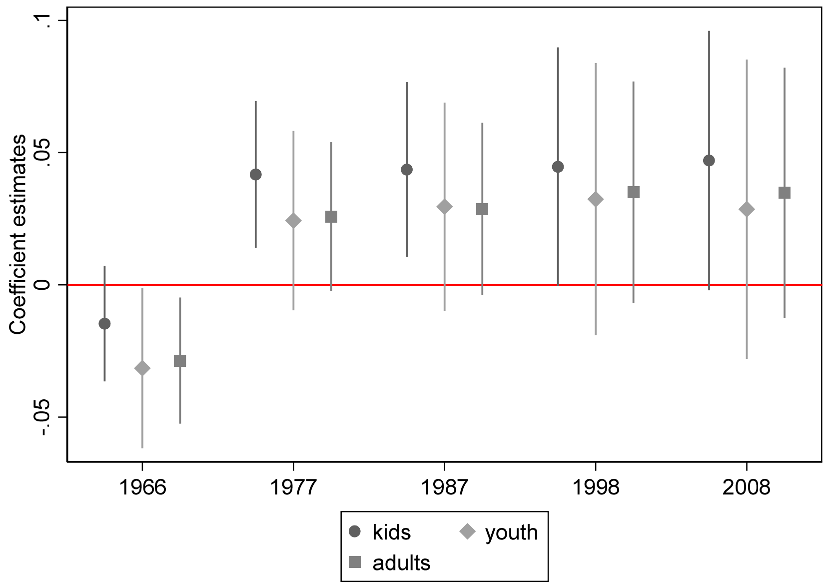


Figure 5a plots coefficients from equation (1) estimated using the log of the age-specific population totals in the district as outcome. Age groups are: adults (19+), youth (5-18) and children (under age 5). Points shown are coefficients on the interaction of Census year dummies with the district-level migrant capital shock. Base year is 1945. Mean value of migrant capital in a given year is 2.25 (million USD).

Figure 5b: Effects of migrant capital inflows on urbanization

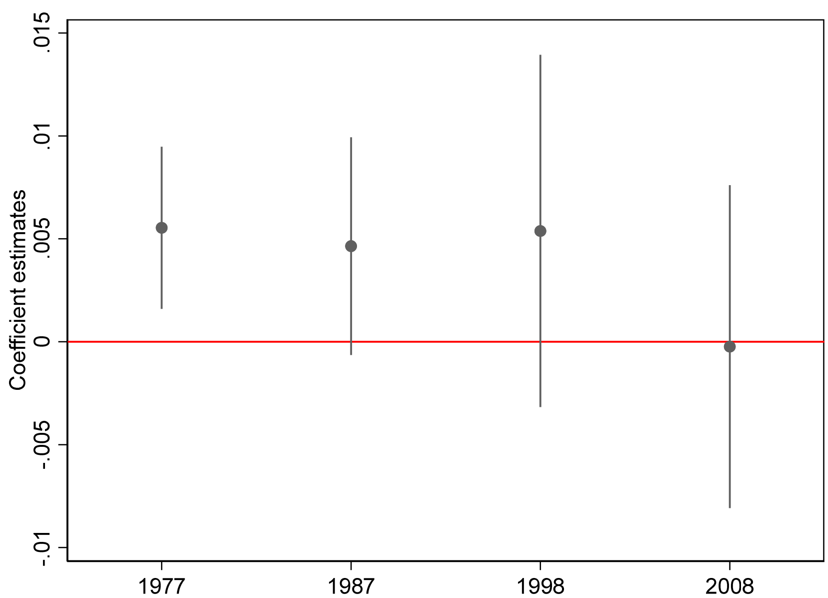


Figure 5b plots coefficients from equation (1) estimated using the share of urban population in the district as outcome. Points shown are coefficients on the interaction of Census year dummies with the district-level migrant capital shock. Base year is 1945. Mean value of migrant capital in a given year is 2.25 (million USD).

Table 1: Sector of work in Malawi over time

	Employment shares by decade			
	1977	**1987**	**1998**	**2008**
Sector of work (Industry): Women				
Agriculture	0.943	0.941	0.888	0.695
Manufacturing	0.016	0.013	0.012	0.038
Services	0.028	0.037	0.067	0.210
Industrial concentration index	0.893	0.893	0.805	0.538
Sector of work (Industry): Men				
Agriculture	0.760	0.761	0.731	0.532
Manufacturing	0.093	0.077	0.074	0.133
Services	0.120	0.135	0.171	0.278
Industrial concentration index	0.618	0.622	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: Farm and non-farm production in Malawi -- Capital and labor inputs

Annual means, 1998 USD	Working capital	Physical capital (equipment)	Land	Total capital = working + physical + land	Effective labor (N.*share of year working)	Revenue = sales + value of home production	Net Value Added = Revenues - Input costs*	Value Added of Labor = Net value added /effective labor
Farm Households N=9,280	20	13	125	158	0.44	97	71	162
Households w/ non-farm business N=1,964	172	139	na/a	311	0.77	540	368	476
Ratio: Non-farm/farms	8.6	10.4		2.0	1.8	5.6	5.2	2.9

Data are from the 1997/1998 Malawi Household Integrated Income and Expenditure Survey (HIES). Unit of observation is the household, means are weighted, values (except for effective labor units) are annual USD. Statistics for farm households include all rural households; statistics in second row are restricted to any rural household running at least one household business with only one worker (the majority of household businesses are single-operator). Working capital includes (e.g.) seeds and fertilizers, or stock for household retail businesses. Physical capital equipment includes (for example): hoes, sickles, pangas and axes for farming activities; bicycles and pounding mills for services. Land is only valued for farm operations. Annual revenues include the market value of home produced goods and services. *Net value added measure excludes the value of land. See data appendix for further discussion of how measures of value added were created.

Table 3: Predicting the time profile of migration at district-level

	Share of all migrants leaving each period			
Time (periods since Jan. 1967)	0.000267***	0.000272***	0.000277***	0.000316***
	(0.00002)	(0.00001)	(0.00001)	(0.00002)
Time*Any recruiting station				-0.0000689***
				(0.00002)
Any recruiting station		-0.001		
		(0.001)		
Indicator for high malaria area		-0.0003		
		(0.001)		
Log population density 1945		-0.001		
		(0.001)		
Share of young adults literate in 1945		-0.019		
		(0.014)		
Indicator for estate area		-0.0016		
		(0.0011)		
Share of working men with no cash wage 1966		0.004		
		(0.004)		
Share of working women with no cash wage 1966		0.002		
		(0.003)		
Central region indicator		-0.004		
		(0.003)		
Southern region indicator		-0.005		
		(0.004)		
District FE?	N	N	Y	Y
N	1,637	1,637	1,637	1,637
R2	0.499	0.524	0.552	0.56

Robust standard errors clustered at district-level. Significance levels ***p<0.01, **p<0.05, *p<0.1 where critical values are taken from the small sample t-distribution. Unit of observation is the district-year-month. Time is linearly measured from January in 1967 (time=1) to December in 1973 (time=84). We exclude the year of the plane crash since many workers would have come back without completing their contracts, hence our estimated value of migrants will be measured with larger errors. Any recruiting station is a dummy variable indicating whether a district has an historically-placed recruiting station. All other controls are measured at district level prior to 1967. Outcome is the (imputed) share of total migrants sent from each district d in each month and year between 1967 and 1973. These imputations take the total deferred pay returning to a district d in a specific month and year and divide it by 60% of the annual mining wage rate to get an estimated number of migrants. Details of variable construction are in text.

Table 4: Predicting spatial variation in deferred pay flows at district-level

	Raw deferred pay flow, 1966-1975	Residualized deferred pay flow, 1966-1975
Number of migrants, 1966-1975	0.0180**	0.0171*
	(0.008)	(0.008)
Indicator for high malaria area	-283.50	-258.20
	(187.50)	(204.80)
Log population density 1945	-197.80	-181.30
	(170.50)	(172.50)
Share of young adults literate in 1945	221.90	303.00
	(1559.00)	(1841.00)
Indicator for estate area	-232.20	-237.10
	(147.60)	(154.00)
Share of working men with no cash wage 1966	-3,892.00	-3,929.00
	(2899.00)	(2954.00)
Share of working women with no cash wage 1966	3,163.00	3,056.00
	(2323.00)	(2347.00)
Central region indicator	286.70	88.34
	(244.70)	(249.10)
Southern region indicator	184.20	-11.42
	(245.40)	(245.80)
Partialling out month-year variation	N	Y
Observations	24	24
R-squared	0.626	0.56

Robust standard errors clustered at district-level. Significance levels ***p<0.01, **p<0.05, *p<0.1 where critical values are taken from the small sample t-distribution. Unit of observation is the district. Outcome in column (2) is the residual from a regression of district-month-year flows of deferred pay on year and month dummies, added up over months and years to district-level. Units are millions of USD.

Table 5: Long-run impacts of migrant capital inflows on sectoral employment shares

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Share of women in:	**Agriculture (Mean = 0.88)**		**Manufacturing (Mean = 0.02)**		**Services (Mean = 0.076)**		**Industry concentration index (Mean = 0.79)**	
USD (Mill.)*Three decades post	-0.0144***	-0.00915***	0.00155*	0.001	0.0168***	0.0106***	-0.005	0.003
	(0.002)	(0.002)	(0.001)	(0.001)	(0.003)	(0.003)	(0.004)	(0.003)
	[0.001]***	[0.008]***	[0.034]**	[0.22]	[0]***	[0.009]***	[0.087]*	[0.195]
USD (Mill.)*Two decades post	-0.0114***	-0.00928***	0.000	0.001	0.00953***	0.00505**	-0.0109**	-0.00910***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.004)	(0.003)
	[0]***	[0.005]***	[0.155]	[0.315]	[0.001]***	[0.038]**	[0.013]**	[0.02]**
USD (Mill.)*One decade post	-0.00398*	-0.002	0.001	0.001	0.003	0.000	-0.00573**	-0.00371*
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)
	[0.044]*	[0.237]	[0.112]	[0.124]	[0.062]*	[0.52]	[0.016]**	[0.081]*
Controls for migrants*Post	-	Y	-	Y	-	Y	-	Y
Joint test: USD (Mill.)*Decades post								
p value	0.00	0.00	0.23	0.06	0.00	0.00	0.01	0.00
Randomization inference p value	0.00	0.13	0.45	0.46	0.01	0.25	0.11	0.21
B: Share of men in	**Agriculture (Mean = 0.73)**		**Manufacturing (Mean = 0.09)**		**Services (Mean = 0.15)**		**Industry concentration index (Mean = 0.58)**	
USD (Mill.)*Three decades post	0.003	0.007	-0.001	-0.002	0.003	-0.002	0.014	0.015
	(0.007)	(0.008)	(0.004)	(0.004)	(0.003)	(0.002)	(0.011)	(0.012)
	[0.171]	[0.222]	[0.222]	[0.398]	[0.1]	[0.208]	[0.067]*	[0.185]
USD (Mill.)*Two decades post	-0.004	-0.002	0.001	0.000	0.00618***	0.00320**	0.000	0.001
	(0.005)	(0.005)	(0.003)	(0.003)	(0.002)	(0.001)	(0.007)	(0.007)
	[0.116]	[0.358]	[0.188]	[0.48]	[0.015]**	[0.079]*	[0.253]	[0.49]
USD (Mill.)*One decade post	-0.004	-0.002	0.000	0.000	0.003	0.000	-0.005	-0.004
	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.004)	(0.004)
	[0.073]*	[0.256]	[0.227]	[0.472]	[0.061]*	[0.465]	[0.084]*	[0.241]
Controls for migrants*Post?	-	Y	-	Y	-	Y	-	Y
Joint test: USD (Mill.)*Decades post								
p value	0.01	0.00	0.50	0.67	0.02	0.05	0.00	0.00
Randomization inference p value	0.10	0.19	0.62	0.86	0.21	0.60	0.03	0.06

N=96 in all regressions. 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. USD (Mill.) is the total deferred pay returning to each district between August 1967 and November 1975 in millions of USD. Data are from Census 1977, 1987 (one decade post), 1998 (two decades post) and 2008 (three decades post). Mean is 2.25. Unit of observation is the district-year cell. Total districts=24. All regressions include year and district fixed effects and interactions of a time trend with baseline district controls (adult literacy in 1945, population density in 1945, a malaria dummy, share of men and women married in 1966, share of men and women not earning any cash income in 1966, two region dummies). Industrial concentration is a Herfindahl index measuring how concentrated work is in any one sector.

Table 6: Long-run impacts of migrant capital inflows on sectoral employment shares
Capital shocks triggered by plane crash

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Share of women in:	**Agriculture (Mean = 0.88)**		**Manufacturing (Mean = 0.02)**		**Services (Mean = 0.076)**		**Industry concentration index (Mean = 0.79)**	
USD (mill.) post crash*Three decades post	-0.0551***	-0.0338***	0.003	0.003	0.0592***	0.0362***	-0.0212*	0.013
	(0.007)	(0.007)	(0.002)	(0.004)	(0.008)	(0.009)	(0.012)	(0.010)
	[0.167]	[0.309]	[0]***	[0.311]	[0]***	[0.301]	[0.5]	[0.046]**
USD (mill.) post crash*Two decades post	-0.0438***	-0.0338***	0.000	0.001	0.0329***	0.0165**	-0.0433***	-0.0316***
	(0.007)	(0.007)	(0.002)	(0.003)	(0.007)	(0.007)	(0.013)	(0.010)
	[0]***	[0.267]	[0.5]	[0.558]	[0.333]	[0.343]	[0]***	[0.043]**
USD (mill.) post crash*One decade post	-0.0145*	-0.006	0.001	0.002	0.011	-0.001	-0.0209**	-0.0111*
	(0.007)	(0.005)	(0.002)	(0.002)	(0.007)	(0.005)	(0.009)	(0.006)
	[0.333]	[0.58]	[0.333]	[0.211]	[0.5]	[0.73]	[0.333]	[0.344]
Controls for migrants*Post	-	Y	-	Y	-	Y	-	Y
Joint test: USD (Mill.) post crash*Decade								
p value	0.00	0.00	0.25	0.02	0.00	0.00	0.00	0.00
Randomization inference p value	1.00	0.88	0.00	0.01	1.00	0.93	0.00	0.02

B: Share of men in	**Agriculture (Mean = 0.73)**		**Manufacturing (Mean = 0.09)**		**Services (Mean = 0.15)**		**Industry concentration index (Mean = 0.58)**	
USD (mill.) post crash*thirty years after	-0.011	0.014	-0.005	-0.008	0.0235**	-0.002	0.019	0.033
	(0.018)	(0.023)	(0.008)	(0.012)	(0.009)	(0.005)	(0.027)	(0.037)
	[0.167]	[0.085]*	[0]***	[0.217]	[0.5]	[0.071]*	[0.167]	[0.181]
USD (mill.) post crash*twenty years after	-0.0306**	-0.016	0.002	-0.001	0.0304***	0.0145***	-0.022	-0.012
	(0.014)	(0.014)	(0.007)	(0.008)	(0.007)	(0.004)	(0.020)	(0.024)
	[0.167]	[0.526]	[0]***	[0.495]	[0.167]	[0.182]	[0.333]	[0.423]
USD (mill.) post crash*ten years after	-0.0209**	-0.011	-0.001	-0.001	0.0127**	0.001	-0.0253**	-0.017
	(0.009)	(0.008)	(0.005)	(0.005)	(0.006)	(0.002)	(0.012)	(0.013)
	[0.167]	[0.472]	[0.333]	[0.402]	[0.167]	[0.194]	[0]***	[0.312]
Controls for migrants*Post?	-	Y	-	Y	-	Y	-	Y
Joint test: USD (Mill.) post crash*Decade								
p value	0.00	0.00	0.46	0.72	0.01	0.02	0.00	0.00
Randomization inference p value	1.00	0.01	1.00	0.94	1.00	0.78	1.00	0.00

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. USD (Mill.) post-plane crash is the total deferred pay returning to each district between April 1974 and November 1975 (mean = 0.6), in millions of USD. Data are from Census 1977, 1987 (one decade post), 1998 (two decades post) and 2008 (three decades post). Unit of observation is the district-year cell. Total districts=24. All regressions include year and district fixed effects and interactions of a time trend with baseline district controls (adult literacy in 1945, population density in 1945, a malaria dummy, share of men and women married in 1966, share of men and women not earning any cash income in 1966, two region dummies). Industrial concentration is a Herfindahl index measuring how concentrated work is in any one sector.

Table 7: Long-run impacts of migrant capital inflows on employment shares in non-farm sectors

	(1)	(2)	(3)	(4)	(5)
A: Share of women in:	**General Manufacturing (Mean = 0.03)**	**Construction (Mean = 0.02)**	**General Services (Mean = 0.05)**	**Retail (Mean = 0.06)**	**Transport and communications (Mean = 0.007)**
USD (Mill.)*Three decades post	0.0002	0.00111***	0.00214**	0.00841***	0.0001
	(0.001)	(0.000)	(0.001)	(0.002)	(0.000)
	[0.481]	[0.018]**	[0.052]*	[0.007]***	[0.138]
USD (Mill.)*Two decades post	0.0000	0.000672*	0.0007	0.00430***	6.20e-05*
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
	[0.458]	[0.073]*	[0.282]	[0.017]**	[0.087]*
USD (Mill.)*One decade post	0.0004	0.000470**	0.0001	-0.0002	0.000145***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
	[0.286]	[0.056]*	[0.497]	[0.419]	[0.003]***
Joint test: USD (Mill.)*Decade					
p value	0.310	0.015	0.000	0.000	0.000
Randomization inference p value	0.506	0.342	0.622	0.140	0.060
B: Share of men in	**General Manufacturing (Mean = 0.05)**	**Construction (Mean = 0.04)**	**General Services (Mean = 0.07)**	**Retail (Mean = 0.07)**	**Transport and communications (Mean = 0.01)**
USD (Mill.)*Three decades post	-0.004	0.003	-0.00374***	0.00368**	-0.00203***
	(0.003)	(0.002)	(0.001)	(0.001)	(0.000)
	[0.138]	[0.13]	[0.024]**	[0.037]**	[0.015]**
USD (Mill.)*Two decades post	-0.00457**	0.00480***	0.000	0.00536***	-0.00207***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)
	[0.066]*	[0.004]***	[0.451]	[0.008]***	[0.009]***
USD (Mill.)*One decade post	-0.00209*	0.00202**	-0.001	0.000847**	0.000
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
	[0.113]	[0.056]*	[0.21]	[0.053]*	[0.167]
Joint test: USD (Mill.)*Decade					
p value	0.030	0.000	0.005	0.000	0.000
Randomization inference p value	0.362	0.092	0.258	0.092	0.004

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. USD (Mill.) is the total deferred pay returning to each district between August 1967 and November 1975 (mean = 2.25), in millions of USD. Data are from Census 1977, 1987 (one decade post), 1998 (two decades post) and 2008 (three decades post). Unit of observation is the district-gender cell. Total districts=24. All regressions include district and year fixed effects, interactions of total migrants and decade dummies, and interactions of a linear trend term with baseline district-level variables (adult literacy in 1945, population density in 1945, a malaria dummy, share of men and women married in 1966, share of men and women not earning any cash income in 1966, two region dummies).

Table 8: Long-run impacts of migrant capital on type of employer

	Sample = Workers in all sectors			Sample = All service sector workers		
	(1)	(2)	(3)	(4)	(5)	(6)
A: Share of women in:	**Self-employed**	**Family business/ family farm**	**Wage work**	**Self-employed**	**Family business**	**Wage work**
	Mean: 0.88	*Mean: 0.03*	*Mean: 0.05*	*Mean: 0.07*	*Mean: 0.04*	*Mean: 0.87*
USD (Mill.)*Three decades post	-0.004	0.00544***	-0.0003	-0.004	0.00666**	-0.003
	(0.003)	(0.001)	(0.001)	(0.005)	(0.003)	(0.004)
	[0.171]	[0.004]	[0.405]	[0.244]	[0.05]	[0.31]
USD (Mill.)*Two decades post	-0.0108***	0.00601***	0.00137*	0.000	0.002	-0.005
	(0.003)	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)
	[0.011]	[0.016]	[0.071]	[0.452]	[0.192]	[0.114]
Joint test: USD (Mill.)*Decade						
p value	0.002	0.000	0.000	0.293	0.029	0.329
Randomization inference p value	0.114	0.018	0.184	0.536	0.230	0.556
B: Share of men in	**Self-employed**	**Family business/ family farm**	**Wage work**	**Self-employment**	**Family business**	**Wage work**
	Mean: 0.68	*Mean: 0.03*	*Mean: 0.22*	*Mean: 0.1*	*Mean: 0.03*	*Mean: 0.86*
USD (Mill.)*Three decades post	0.009	0.00257***	-0.007	0.000	0.00422**	-0.003
	(0.006)	(0.001)	(0.006)	(0.002)	(0.002)	(0.003)
	[0.142]	[0.028]**	[0.223]	[0.449]	[0.034]	[0.219]
USD (Mill.)*Two decades post	0.000	0.00227***	-0.001	0.000	0.000	-0.001
	(0.003)	(0.001)	(0.003)	(0.002)	(0.001)	(0.003)
	[0.446]	[0.02]**	[0.441]	[0.459]	[0.421]	[0.424]
Controls for migrants*Post?	Y	Y	Y	Y	Y	Y
Joint test: USD (Mill.)*Decade						
p value	0.044	0.009	0.258	0.986	0.017	0.394
Randomization inference p value	0.336	0.180	0.708	0.994	0.240	0.634

N=72 in all regressions. Robust standard errors clustered at district level and shown in parentheses. Randomization inference p-values in square brackets. Significance levels ***p<0.01, **p<0.05, *p<0.1. USD (Mill.) is the total deferred pay returning to each district between April 1974 and November 1975 (mean = 2.25), in millions of USD. Data are from Census 1977, 1987 (one decade post), 1998 (two decades post) and 2008 (three decades post). Unit of observation is the district-gender cell. Total districts=24. All regressions include district and year fixed effects, total migrants interacted with year fixed effects, and interactions of a linear trend term with baseline district-level variables (adult literacy in 1945, population density in 1945, a malaria dummy, share of men and women married in 1966, share of men and women not earning any cash income in 1966, two region dummies).

Table 9: Long-term impacts of migrant capital inflows on investments in physical capital

	Farm capital			
A: Share of households with (number of*)	(1)	(2)	(3)	(4)
	Hoes(*)	**Pangas(*)**	**Any Cattle**	**Oxcart**
	Mean: 1.82	*Mean: 0.48*	*Mean: 0.1*	*Mean: 0.02*
USD (Mill.)*Three decades post (Year=2008)				-0.008
				(0.009)
USD (Mill.)*Two decades post (Year=1998)	-0.010	0.006	0.007	-0.002
	(0.021)	(0.006)	(0.007)	(0.006)
Base year	1968	1968	1968	1987
Years of data in sample	1968, 1997	1968, 1997	1968, 1997	1987, 1998, 2008
N	46	46	46	69
Joint test: USD (Mill.)*Decade				
p value	n/a	n/a	n/a	0.08
	Non-farm capital			
B: Share of households with	**Radio**	**Durable walls**	**Durable roof**	**Durable roof and walls**
	Mean: 0.28	*Mean: 0.39*	*Mean: 0.13*	*Mean: 0.12*
USD (Mill.)*Three decades post (Year=2008)	-0.00447***	-0.010	0.006	0.007
	(0.002)	(0.021)	(0.006)	(0.007)
USD (Mill.)*Two decades post (Year=1998)	0.000			
	(0.002)			
USD (Mill.)*One decade post (Year=1987)	0.001	0.011	0.0136***	0.0134***
	(0.001)	(0.012)	(0.003)	(0.004)
Millions of USD*Post (Year=1977)	0.00145*			
	(0.001)			
Base year	1968	1968	1968	1968
Years of data in sample	1969, 1977, 1987, 1998, 2008	1969, 1987, 2008	1969, 1987, 2008	1969, 1987, 2008
N	115	69	69	69
Joint test: USD (Mill.)*Decade				
p value	0.00	0.01	0.00	0.00

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 different Census years for each outcome. Outcomes in Panel A: *Share of households with a bike, oxcart, or any cattle, and mean number of hoes and pangas per household in the district. Unit of observation is the district-year cell. Total districts with data in all Census years including 1966: 23. Other controls include interactions of survey year dummies with total migrants between 1966 and 1977 and baseline district variables: adult literacy in 1945, population density in 1945, a malaria dummy, the share of men and women married in 1966, the share of men and women not earning any cash income in 1966, and region dummies. All regressions contain district fixed effects and year dummies. Regressions are not weighted.

Table 10: Long-term impacts of migrant capital inflows on investment in education of the next generation, by gender

	Years of completed education		Any primary schooling	
	(1)	(2)	(3)	(4)
	Females	**Males**	**Females**	**Males**
USD (Mill.)*Early treated cohorts	0.103***	0.100***	0.010***	0.006**
	(0.029)	(0.034)	(0.003)	(0.003)
USD (Mill.)*Late treated cohorts	0.140***	0.131**	0.013**	0.009**
	(0.047)	(0.048)	(0.005)	(0.004)
USD (Mill.)*Post treated cohorts	0.122**	0.112***	0.010**	0.007**
	(0.047)	(0.039)	(0.004)	(0.003)
N	240	240	240	240
Mean of outcome:	1.88	3.23	0.35	0.47
Joint tests:				
p value	0.000	0.000	0.002	0.033
randomization inference p value	0.016	0.05	0.056	0.114

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. Total districts in dataset: 24. *Early treated cohorts* are those age-eligible for primary school during 1967-1973; *Late treated cohorts* are those age-eligible for primary school during 1974-1977; *Posttreated cohorts* are those age-eligible for primary school 1977-1980. Other controls include interactions of cohort dummies with total migrants between 1966 and 1977, and baseline district variables interacted with trend terms: adult literacy in 1945, population density in 1945, a malaria dummy, and region dummies. All regressions contain district fixed effects and cohort and gender dummies. Regressions are unweighted.

Table 11: Does a temporary capital injection lead to greater wealth 15 years later? Evidence from Malawi's 1992 Demograhic Health Survey

	Estimates of β (s.e.) [N] on Millions of USD received post-plane crash			
	Additional controls			
Asset outcome	No other controls	+Num. Migrants	+Region FE	+Baseline district controls and region FE
DHS Wealth index	0.196***	0.211***	0.212***	0.157***
	(0.024)	(0.013)	(0.012)	(0.030)
Count of assets	0.249***	0.261***	0.275***	0.247**
	(0.034)	(0.026)	(0.031)	(0.095)
Durable roof	0.0880***	0.0946***	0.0906***	0.0801***
	(0.010)	(0.005)	(0.006)	(0.013)
Durable floor	0.0704***	0.0763***	0.0780***	0.0669***
	(0.010)	(0.005)	(0.005)	(0.011)
Electricity	0.0232***	0.0241***	0.0239***	0.0121***
	(0.003)	(0.002)	(0.002)	(0.002)
Radio	0.0409***	0.0446***	0.0537***	0.0406*
	(0.012)	(0.010)	(0.009)	(0.021)
Car	0.00873***	0.00807***	0.00799***	0.00871***
	(0.001)	(0.001)	(0.001)	(0.002)
Bike	-0.0230***	-0.0301***	-0.0276***	-0.027
	(0.008)	(0.005)	(0.006)	(0.019)
Improved toilet	0.0530***	0.0658***	0.0691***	0.023
	(0.015)	(0.008)	(0.013)	(0.029)
Improved water source	-0.010	-0.0136**	-0.0153*	0.009
	(0.009)	(0.006)	(0.009)	(0.029)

Sample size = 5,323 in all regressions. Robust standard errors clustered at the district level. Each block of coefficients is estimated in a separate regression of the outcome variable on the amount of migrant capital returning to the district immediately after the plane crash in April 1974, and a series of controls variables: number of migrants 1966-1975, and region fixed effects and baseline district level controls (population density in 1945, literacy rates, an indicator for estates in the district, a malaria risk indicator, the share of men and women married in 1966 and the share of men and women working for no cash wage in 1966). Household level regressions are weighted using DHS sample weights. All outcomes are indicators 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.

Online Publication Only: Data Appendix

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 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.

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?

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

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 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.

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, homeworker, 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: Homeworker, 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)

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

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

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

4 Household Income and Expenditure Surveys

We used micro-level data from the Malawi Integrated Household Income and Expenditure Survey 1997/1998 to characterize the capital intensity of farm and non-farm activities in Malawi and to create a measure of agricultural productivity (value-added of labor in agriculture; see Gollin, Lagakos and Waugh (2014) for a detailed description of this measure).

To measure value added in agriculture, we computed the total value of self-employment output in agriculture, the value of labor income from agricultural work outside of the household and any interest on land rented out. For self-employment output on farms, we valued all crops produced at home, whether for market or home consumption. We valued home produced goods at local or national market prices, whichever was available for the specific crop and unit harvested. We do the same for livestock sold. From this agricultural income total at the household level, we subtracted out the value of non-labor inputs (rented land, fertilizer, seeds) and hired-in labor inputs into agricultural production.

To measure value added in non-farm activities at the household level, we added revenues from non-agricultural household businesses, wage and salary income from non-farm work outside of the home, and subtracted out the costs of intermediate inputs used in self-employment. Almost no households report renting out capital equipment for non-farm use.

Measuring labor used in each sector is tricky, mainly because workers do not have full-time jobs in either sector. We computed effective units of labor used in agriculture, and in non-farm work at household level, by counting up the number of workers reporting their primary occupation is in agriculture as a farmer, or not, and weighting these workers by the average number of weeks worked in the last year. We also included women reporting home production as farm workers in the household. Following the macro literature, we included measures of workers who are currently unemployed but who report sector of work and weeks of work per year.

To create value added of labor measures, we divided household value added in each sector by the total number of workers in the household in each sector.

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.

Online Publication Only: Additional Results Appendix

Appendix Table 1: Summary statistics, district-level data

	Mean	s.d.	min	max	N
Components of migration shock					
Number of adult men (aged 15-64) in the district in 1977	54,809	39,418	13,057	180,466	24
Number of adult men ever been abroad by 1977~	19,557	15,421	4,232	75,324	24
Δ number of migrants, 1966-1977	13,642	10,667	2,816	50,121	24
Total deferred pay per district 1966-1975, Millions of USD	2.25	3.53	0	16.29	24
Total deferred miner pay per person in district 1966-1977, USD	24.04	55.40	0	275.68	24
Total deferred miner pay per migrant from district 1966-1977, USD	129.41	177.76	0	908.46	24
District-level descriptives at baseline					
Northern Region	0.21	0.41	0	1	24
Central Region	0.38	0.49	0	1	24
Southern Region	0.42	0.50	0	1	24
Population, 1945	71,262	60,353	5,919	230,891	24
Population density, 1945	30.61	26.61	5.10	109.05	24
Share of youth literate in English and vernacular, 1945	0.08	0.04	0.03	0.14	24
Altitude: high malaria area=1	0.28	0.35	0	1	24
Share of districts with any agricultural estate	0.46	0.51	0	1	24
Share of men earning no cash income in 1966	0.37	0.10	0.22	0.59	24
Share of women earning no cash income in 1966	0.48	0.14	0.28	0.72	24
Share of districts with at least one *Wenela* recruiting station	0.63	0.50	0.00	1.00	24

Data for the first set of outcomes are district-level data collected from administrative records and from Census 1977. Data for the second set of outcomes comes from 1945 Census data 1and from geographic files for Malawi. Agricultural estate is a dummy variable indicating whether a district contains any cash crop estates (e.g. for tobacco or sugar). Raw means (unweighted).

Appendix Table 2: Impacts on population growth and urbanization

	Ln population	Ln population (females only)	Ln population under age 5	Ln population age 5 to 18	Ln population age 18 +	Share of urban population
USD (Mill.)*Three decades after	0.036	0.0289	0.0470*	0.0286	0.0348	-0.0002
	(0.024)	(0.023)	(0.024)	(0.027)	(0.023)	(0.004)
	[0.226]	[0.27]	[0.159]	[0.329]	[0.21]	[0.501]
USD (Mill.)*Two decades after	0.0372*	0.0301	0.0447*	0.0324	0.0350*	0.00465*
	(0.022)	(0.020)	(0.022)	(0.025)	(0.020)	(0.003)
	[0.194]	[0.227]	[0.151]	[0.289]	[0.177]	[0.159]
USD (Mill.)*One decade after	0.0329*	0.0259*	0.0436**	0.0296	0.0287*	0.00465*
	(0.016)	(0.015)	(0.016)	(0.019)	(0.016)	(0.003)
	[0.155]	[0.19]	[0.095]*	[0.247]	[0.168]	[0.112]
USD (Mill.)*End of migration	0.0291**	0.0211*	0.0417***	0.0243	0.0258*	0.00554***
	(0.014)	(0.012)	(0.013)	(0.016)	(0.014)	(0.002)
	[0.152]	[0.194]	[0.074]*	[0.257]	[0.158]	[0.073]*
USD (Mill.)*One decade before	-0.0261**	-0.0353***	-0.0146	-0.0315**	-0.0286**	
	(0.012)	(0.012)	(0.011)	(0.015)	(0.012)	
	[0.117]	[0.077]*	[0.243]	[0.124]	[0.102]	
N	144	144	144	144	144	119
Joint test: USD (Mill.)*Decade						
p value	0.30	0.33	0.03	0.26	0.47	0.00
Randomization inference p value	0.88	0.89	0.64	0.91	0.93	0.15

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 Census data are from 1945, 1966, 1977, 1987, 1998 and 2008. Omitted category is Millions of USD*(Year=1945) for the first three outcomes; Millions of USD*(Year=1966) for the last outcome. Unit of observation is the district-year cell. Total districts=24. All regressions control for district and year fixed effects, interactions of baseline controls with year dummies, and interactions of number of migrants with year dummies. Regressions are not population-weighted. Urban share of the district not available in 1945.

Appendix Table 3: Long-run impacts of migrant capital inflows on sectoral employment shares
Robustness to omitting district fixed effects, baseline controls interacted with decade dummies, and migrant controls

	Share in Agriculture						Share in Services					
Covariates included:	No others	Only district FE	Only trend controls	Only migrant controls	Only district FE and trend interactions	All controls	No others	Only district FE	Only trend controls	Only migrant controls	Only district FE and trend interactions	All controls
Panel A: Women			*Mean: 0.88*						*Mean: 0.08*			
Millions of USD*Three decades post	-0.0207***	-0.0108***	-0.0241***	-0.0253***	-0.0144***	-0.00915***	0.0191***	0.0118**	0.0221***	0.0242***	0.0168***	0.0106***
	(0.006)	(0.003)	(0.004)	(0.004)	(0.002)	(0.002)	(0.007)	(0.005)	(0.003)	(0.004)	(0.003)	(0.003)
	[0.014]**	[0.008]***	[0.001]***	[0.023]**	[0.001]***	[0.008]***	[0.024]**	[0.027]**	[0.001]***	[0.027]**	[0.0]***	[0.009]***
Millions of USD*Two decades post	-0.0188***	-0.00899**	-0.0211***	-0.0245***	-0.0114***	-0.00928***	0.0135***	0.00621**	0.0155***	0.0173***	0.00953***	0.00505**
	(0.006)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
	[0.022]**	[0.013]**	[0.001]***	[0.021]**	[0.000]***	[0.005]***	[0.02]**	[0.019]**	[0.001]***	[0.021]**	[0.001]***	[0.038]**
Millions of USD*One decade post	-0.0126***	-0.00278**	-0.0138***	-0.0160***	-0.00398*	-0.002	0.00899***	0.00173***	0.00998***	0.0111***	0.003	0.000
	(0.003)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.000)	(0.001)	(0.001)	(0.002)	(0.001)
	[0.007]***	[0.02]**	[0.001]***	[0.015]**	[0.044]*	[0.237]	[0.006]***	[0.008]***	[0.001]***	[0.019]**	[0.062]*	[0.52]
P value for Joint F on Migrant capital	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
Randomization inference p value: joint test	0.026	0.030	0.002	0.026	0.002	0.126	0.006	0.010	0.002	0.028	0.012	0.250
Panel B: Men			*Mean: 0.73*						*Mean: 0.15*			
Millions of USD*Three decades post	-0.0219***	-0.001	-0.0278***	-0.0245***	0.003	0.007	0.0173***	0.00533**	0.0170***	0.0190***	0.003	-0.002
	(0.005)	(0.006)	(0.006)	(0.004)	(0.007)	(0.008)	(0.004)	(0.002)	(0.005)	(0.004)	(0.003)	(0.002)
	[0.008]***	[0.213]	[0.005]***	[0.025]**	[0.171]	[0.222]	[0.006]***	[0.033]**	[0.014]***	[0.03]**	[0.1]	[0.208]
Millions of USD*Two decades post	-0.0277***	-0.00706**	-0.0316***	-0.0330***	-0.004	-0.002	0.0195***	0.00755***	0.0193***	0.0228***	0.00618***	0.00320**
	(0.008)	(0.003)	(0.005)	(0.006)	(0.005)	(0.005)	(0.006)	(0.002)	(0.003)	(0.005)	(0.002)	(0.001)
	[0.012]**	[0.027]**	[0.001]***	[0.027]**	[0.116]	[0.358]	[0.011]**	[0.002]***	[0.001]***	[0.03]**	[0.015]**	[0.079]*
Millions of USD*One decade post	-0.0262***	-0.006	-0.0282***	-0.0328***	-0.004	-0.002	0.0153***	0.00327**	0.0152***	0.0181***	0.003	0.000
	(0.007)	(0.003)	(0.005)	(0.004)	(0.003)	(0.003)	(0.004)	(0.001)	(0.003)	(0.003)	(0.002)	(0.001)
	[0.011]**	[0.047]**	[0.002]***	[0.016]**	[0.073]*	[0.256]	[0.01]**	[0.034]**	[0.001]***	[0.025]**	[0.061]*	[0.465]
P value for Joint F on Migrant capital	0.000	0.003	0.000	0.000	0.006	0.000	0.004	0.001	0.000	0.000	0.022	0.054
Randomization inference p value: joint test	0.080	0.110	0.020	0.030	0.100	0.190	0.170	0.060	0.030	0.190	0.210	0.60

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 where critical values are taken from the small sample *t*-distribution. USD (Mill.) is the total deferred pay returning to each district by 1975 (mean=2.25), in millions of USD. Data are from Census 1977, 1987 (one decade post), 1998 (two decades post) and 2008 (three decades post). Unit of observation is the district-gender cell. Total districts=24. Other controls include: interactions of a linear trend with baseline district variables (adult literacy in 1945, population density in 1945, a malaria dummy, share of men and women married in 1966, share of men and women not earning any cash income in 1966, two region dummies). Controls for number of migrants are interactions of total returning migrants and decade dummies. All regressions include year fixed effects. The final column for each outcome presents the same results as appear in Table 5.

Appendix Table 4: Long-run impacts of migrant capital inflows on employment shares in different sectors Omitting Lilongwe

	(1)	(2)	(3)	(4)
A: Share of women in:	**Agriculture (Mean = 0.88)**		**Services (Mean = 0.02)**	
Millions of USD post crash*Three decades pos	-0.0139***	-0.0101***	0.0162***	0.0104**
	(0.002)	(0.003)	(0.003)	(0.004)
	[0]***	[0.015]**	[0]***	[0.03]**
Millions of USD post crash*Two decades post	-0.0108***	-0.0110***	0.00891***	0.005
	(0.002)	(0.002)	(0.002)	(0.003)
	[0]***	[0.007]***	[0.001]***	[0.102]
Millions of USD post crash*One decade post	-0.00371*	-0.002	0.003	-0.0004
	(0.002)	(0.002)	(0.002)	(0.002)
	[0.05]**	[0.26]	[0.07]*	[0.407]
Controls for migrants*Post?	-	Y	-	Y
Joint test: Migrant capital*Decade, Regular p value	0.00	0.00	0.00	0.00
Randomization inference p value	0.00	0.08	0.00	0.11
B: Share of men in	**Agriculture (Mean = 0.73)**		**Services (Mean = 0.15)**	
Millions of USD post crash*Three decades pos	0.003	0.010	0.003	-0.004
	(0.007)	(0.009)	(0.003)	(0.003)
	[0.169]	[0.172]	[0.107]	[0.146]
Millions of USD post crash*Two decades post	-0.003	-0.002	0.00550***	0.00315*
	(0.005)	(0.005)	(0.002)	(0.002)
	[0.144]	[0.359]	[0.018]**	[0.112]
Millions of USD post crash*One decade post	-0.004	-0.002	0.00208*	0.000
	(0.003)	(0.003)	(0.001)	(0.001)
	[0.077]*	[0.325]	[0.05]**	[0.365]
Controls for migrants*Post?	-	Y	-	Y
Joint test: Migrant capital*Decade, Regular p value	0.01	0.00	0.02	0.00
Randomization inference p value	0.11	0.13	0.21	0.18

N=92 in all regressions. 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. USD (Mill.) is the total deferred pay returning to each district between August 1967 and November 1975 in millions of USD. Data are from Census 1977, 1987 (one decade post), 1998 (two decades post) and 2008 (three decades post). Mean is 2.25. Unit of observation is the district-year cell. Total districts=24. All regressions include year and district fixed effects and interactions of a time trend with baseline district controls (adult literacy in 1945, population density in 1945, a malaria dummy, share of men and women married in 1966, share of men and women not earning any cash income in 1966, two region dummies) and with total number of migrants.

Appendix Table 5: Long-run impacts of migrant capital inflows on employment shares in different sectors
Alternative specifications for migrant controls

	Using log of migrants* Decade dummies		Using migrants*decade dummies and migrants squared*decade dummies	
	(1)	(2)	(3)	(4)
A: Share of women in:	**Agriculture (Mean = 0.88)**	**Services (Mean = 0.02)**	**Agriculture (Mean = 0.88)**	**Services (Mean = 0.02)**
Millions of USD post crash*Three decades pos	-0.00943***	0.00943***	-0.00873***	0.00914***
	(0.003)	(0.002)	(0.002)	(0.003)
	[0.012]*	[0.002]**	[0.012]*	[0.012]*
Millions of USD post crash*Two decades post	-0.00939***	0.004	-0.00994***	0.004
	(0.002)	(0.002)	(0.002)	(0.003)
	[0.008]***	[0.088]*	[0.007]***	[0.104]
Millions of USD post crash*One decade post	-0.001	-0.001	-0.001	-0.001
	(0.002)	(0.001)	(0.002)	(0.002)
	[0.322]	[0.244]	[0.311]	[0.317]
Migrant*Decade controls included?	Y	Y	Y	Y
Joint test: Migrant capital*Decade, Regular p value	0.000	0.000	0.380	0.059
Randomization inference p value	0.098	0.068	0.094	0.130
B: Share of men in	**Agriculture (Mean = 0.73)**	**Services (Mean = 0.15)**	**Agriculture (Mean = 0.73)**	**Services (Mean = 0.15)**
Millions of USD post crash*Three decades pos	0.007	-0.002	0.009	-0.003
	(0.007)	(0.003)	(0.008)	(0.002)
	[0.216]	[0.257]	[0.177]	[0.143]
Millions of USD post crash*Two decades post	-0.003	0.003	-0.003	0.00325*
	(0.004)	(0.002)	(0.004)	(0.002)
	[0.261]	[0.129]	[0.29]	[0.094]*
Millions of USD post crash*One decade post	-0.002	0.000	-0.002	0.000
	(0.003)	(0.001)	(0.003)	(0.001)
	[0.225]	[0.396]	[0.27]	[0.352]
Migrant*Decade controls included?	Y	Y	Y	Y
Joint test: Migrant capital*Decade, Regular p value	0.001	0.087	0.000	0.040
Randomization inference p value	0.140	0.484	0.124	0.334

N=96 in all regressions. 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. USD (Mill.) is the total deferred pay returning to each district between August 1967 and November 1975 in millions of USD. Data are from Census 1977, 1987 (one decade post), 1998 (two decades post) and 2008 (three decades post). Mean is 2.25. Unit of observation is the district-year cell. Total districts=24. All regressions include year and district fixed effects and interactions of a time trend with baseline district controls (adult literacy in 1945, population density in 1945, a malaria dummy, share of men and women married in 1966, share of men and women not earning any cash income in 1966, two region dummies). Controls for log number of migrants interacted with decade dummies are included in columns (1) and (2); controls for interactions of decade dummies with number of migrants and number of migrants squared are included in columns (3) and (4).