Structural Transformation Under Subsistence Constraints: Evidence from Climate Shocks in Africa

CHUNXIAO JING *

July 2025

Abstract

Structural transformation, the movement of labor from low-productivity agriculture to higher-productivity manufacturing and services, is a key driver of economic growth in developing economies. However, in much of Africa, this process remains slow, with a large share of the population still engaged in subsistence agriculture. This paper develops a micro-founded two-period model incorporating Stone–Geary utility and transition costs to examine how food subsistence constraints and mobility barriers shape individuals' sectoral choices. To empirically test the model, we combine panel data from the LSMS with high-frequency satellite data (NDVI, temperature, and rainfall) to construct a continuous climate shock index using Random Forest in Niger, Nigeria, Malawi, and Ethiopia. Empirical results show that individuals below the food-subsistence threshold are more responsive to positive income shocks, while those above are less likely to move. Simulations reveal that subsidizing transition costs significantly increases structural transformation, but with heterogeneous impacts across food-subsistence groups and countries. Our findings provide new insight into the microeconomic frictions that hinder labor mobility and offer actionable policy tools to accelerate structural change in low-income settings.

*We are grateful to Jeremy Foltz and Vladimir Gimpelson for their insightful comments and suggestions. This research received no external funding. The analysis uses publicly available microdata from the World Bank's LSMS. The author declares no relevant or material financial interests related to this research. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

1 Introduction

Labor moving from low-productivity agriculture to higher-productivity manufacturing or services—a process known as structural transformation—has long been recognized as a key driver of economic growth in developing economies (Lewis et al., 1954; Syrquin, 1988). However, the pace of structural change in Africa has remained low compared with other regions over the past decades (Morsy et al., 2023; Mensah et al., 2018; Fox et al., 2017). As a result, the agricultural sector continues to account for a large share of labor employment in many African countries (Christiaensen and Maertens, 2022; Herrendorf et al., 2022; Barrett et al., 2017; McMillan and Headey, 2014). With a substantial portion of the population still engaged in subsistence agriculture, overall economic growth in Africa remains sluggish (Aristizabal-Ramirez et al., 2022; Diao et al., 2019; McMillan et al., 2017; Herrendorf et al., 2014).

According to Lewis et al. (1954), surplus labor from agriculture can be reallocated to the manufacturing sector, where productivity is higher. Harris and Todaro (1970) further explain that migration from rural to urban areas is driven by expected wage differentials, with higher urban manufacturing wages attracting workers away from agriculture. However, in Africa, while agricultural labor is indeed shifting toward non-agricultural sectors, many agricultural workers remain underemployed (De Janvry et al., 2022; Tsuda, 2022; Beegle and Christiaensen, 2019; McCullough, 2017; AfDB, 2016), and a persistent wage gap exists between agriculture and non-agriculture (Baysan et al., 2024; McCullough, 2017; Gollin et al., 2014; Vollrath, 2014). One key reason for this paradox is that agricultural workers in Africa face both food insecurity (Bambio et al., 2022; Dillon and Barrett, 2017; Karlan et al., 2014; Gollin et al., 2002) and significant transition costs associated with leaving the sector (Baseler, 2023; Morando, 2023; Morten, 2019; Munshi and Rosenzweig, 2016). Therefore, accelerating structural transformation in Africa requires addressing these constraints. In particular, increasing agricultural income could help alleviate food insecurity and offset transition costs, potentially enabling labor to shift more efficiently into higher-productivity non-agricultural sectors.

To address this question, we develop a two-period model that incorporates a Stone–Geary utility function, explicitly accounting for food subsistence requirements and transition costs, building on the framework proposed by Gollin (2002). In the model, agricultural workers are rational decision-makers (Schultz, 1954) who maximize their lifetime utility by choosing whether to remain in the agricultural sector or move to the non-agricultural sector in the second period. If agricultural income rises sufficiently such that those currently above the food-subsistence threshold can afford the transition cost—or those below it are lifted above the threshold and can cover the cost—then individuals will evaluate their options. Ultimately, only those whose utility from moving exceeds that of staying will choose to exit agriculture.

Building on the theoretical model, we develop two empirical strategies to estimate how changes in agricultural income affect individuals' decisions to transition from agriculture to non-agriculture. To conduct the analysis, we construct individual-level panel datasets from four African countries: Ethiopia (2019, 2021), Malawi (2013, 2016, 2019), Niger (2011, 2014), and Nigeria (2010, 2019). Each panel spans a three-year interval, meaning the actual timing of occupational transition may occur at any point within the gap.

In the first empirical strategy, we regress the probability of transition on the agricultural income shock. A key empirical challenge is that agricultural income is not observed annually and is therefore unavailable for the intervening years. To address this, we use climate shocks—observable at high temporal resolution—as a proxy for changes in agricultural income. These shocks affect crop yields and thus influence earnings in agriculture.

The empirical analysis proceeds in three steps. First, we construct a continuous climate shock measure using NDVI, temperature, and rainfall data modeled through a Random Forest algorithm. Second, we define a binary structural transformation variable from the LSMS panel data, coded as one if the individual moves from agriculture to non-agriculture, and zero if they remain in agriculture. Third, we estimate a logit regression to analyze the probability of transitioning. Our results show that individuals below the food-subsistence threshold are significantly more likely to move in response to positive climate shocks, while those above the threshold are not.

This pattern is consistent with the intuition that people below the subsistence level typically lack the resources needed to afford the transition cost. When a positive climate shock raises their agricultural income, a subset of these individuals gains enough surplus to finance the move. However, the fact that individuals above the subsistence threshold are not responsive to the shock does not imply they are immobile. Instead, it likely reflects that those with a strong preference to exit agriculture have already done so, leaving behind individuals with fewer assets, a comparative advantage, or personal preference for farming (e.g., innate agricultural skills). This highlights that the food subsistence level, the transition cost, and the systematic preference shape an individual's mobility decisions.

The impact of climate shocks observed in the empirical analysis is relatively temporary and random, tending to regress back to the mean over time. Moreover, the empirical results suggest that individuals in different food-subsistence groups exhibit distinct preferences when it comes to transitioning out of agriculture. To more precisely examine how changes in agricultural income influence structural transformation—while accounting for these underlying preferences—we simulate the model under a scenario in which individuals receive a subsidy to offset the transition cost.

The simulation results show that subsidizing transition costs increases the overall movement rate in all four African countries. However, the composition of movers differs across countries. In rel-

atively richer countries like Nigeria and Ethiopia, individuals below the food-subsistence level are more likely to move. In contrast, in poorer countries such as Niger and Malawi, it is primarily individuals above the food-subsistence level who take advantage of the subsidy and choose to exit agriculture.

Our paper contribute to the Micro-founded Theory of Structural Change with Subsistence and Transfer Costs. More recent contributions have advanced the theoretical underpinnings by incorporating non-homothetic preferences, particularly Stone–Geary utility functions, to capture the role of subsistence consumption in shaping labor allocation (Morando, 2023; Comin et al., 2021; Herrendorf et al., 2013; Buera and Kaboski, 2009; Matsuyama, 2008; Gollin et al., 2002; Kongsamut et al., 2001). These models explain how minimum consumption needs can delay transitions out of agriculture, particularly in low-income settings. Parallel to this, other studies have emphasized the role of migration or transition costs as frictions impeding labor reallocation, often modeling such costs as barriers to efficient labor mobility (Baseler, 2023; Lagakos et al., 2020; Morten, 2019; Munshi and Rosenzweig, 2016; Kennan and Walker, 2011). Despite these advances, few studies have jointly considered subsistence constraints and transition costs in a unified micro-founded framework. An exception is Pan and Sun (2024), who incorporate Stone–Geary preferences and endogenous migration costs in a spatial general equilibrium model focused on remittances. However, to our knowledge, no study to date embeds both subsistence requirements and transition costs into an individual-level model of sectoral choice within the context of structural transformation. This paper fills that gap by developing a two-period model in which rational agricultural workers weigh their food security status against transition costs to maximize lifetime utility. In doing so, it bridges two strands of the literature and provides a richer framework for understanding constrained labor mobility in developing economies.

Another contribution of this study lies in the construction of a generalizable climate shock measure. Unlike previous empirical studies examining the relationship between agricultural productivity and structural change—which often rely on shocks triggered by specific events—our approach creates a continuous climate shock variable that can be applied broadly without special conditions. For example, many existing studies use weather-driven shocks, such as extreme deviations in rainfall or temperature (Colmer, 2021; Branco and Féres, 2021; Emerick, 2018; Jessoe et al., 2018; Cattaneo and Peri, 2016; Hirvonen, 2016), or technology-induced shocks, such as the introduction of new agricultural technologies (Gollin and Rogerson, 2014; Bustos et al., 2016; Adamopoulos and Restuccia, 2020; Gollin et al., 2021; Rajaonarison, 2014). In contrast, our method employs machine learning techniques to generate a continuous climate shock index based on NDVI, temperature, and rainfall, without depending on rare events. This makes the shock variable more widely usable across settings and better suited for studying long-term changes in agricultural environments.

The remainder of the paper is organized as follows. Section 2 describes the data sources and outlines the construction of the climate shock variable. Section 3 presents the theoretical model that characterizes individuals' sectoral transition decisions. Section 4 examines the empirical relationship between climate shocks and labor movement across sectors. Section 5 estimates the structural model and explores counterfactual scenarios. Section 6 concludes and discusses the broader implications of the findings.

2 Data Sources and Climate Shock Construction

2.1 LSMS Data

To examine the impact of climate shocks on structural change in Africa, we construct individual panel data for four African countries (Ethiopia, Malawi, Niger, and Nigeria) using the World Bank's Living Standards Measurement Study (LSMS). The LSMS-ISA program provides the representative and longitudinal surveys we need for this research. (Bentze and Wollburg, 2024; Grosh and Glewwe, 1998) Besides, the dataset provides better measurement at the micro level across developing countries, which enhances our understanding of structural change in Africa. (Lagakos and Shu, 2023)

To track the individual movement across countries, we need to have panel surveys that interview the same individuals in at least two waves. Additionally, the time intervals between the two surveys vary across countries. We aim to create balanced panel data, where individuals are interviewed at the same time interval, thereby making estimation outcomes more comparable. Finally, we construct a balanced individual-level panel data with a 3-year gap using the LSMS survey data from four countries (Ethiopia (2019, 2021), Malawi (2013, 2016, 2019), Niger (2011, 2014), and Nigeria (2010, 2019)). This data tracks the same households over time and provides detailed household information such as household status, individual's employment, household expenditure, household agricultural inputs and outputs, household food security, etc.

To construct the cross-country 3-year panel data, we start with the household survey that includes the variables needed for estimation, such as age, education level, employment status, sector, food-subsistence level, income, etc. We merge the individuals based on their ID over the years and double-check with their gender and age. For the dataset without individual ID, we first merge the households based on household ID and then match individuals based on their household ID, gender, and age. In this panel data, we are able to track an individual's sectoral movement, which is the dependent variable in our research.

Aside from the household survey, we also use the agricultural survey for agricultural income be-

cause the agricultural sector accounts for about 50% of the labor force in Sub-Saharan Africa. In the agricultural survey, households report their agricultural outputs in different units such as kilograms, plates, bags, pails, etc. To compute the agricultural output values, we need both the crop quantities and the crop prices. For the quantities, we first use the unit conversion provided to transform the self-reported outputs into kilograms based on the crop types. Then we compute the crop price. Following the idea from De Magalhães and Santaeulàlia-Llopis (2018), we create a regional price from the crop sold and its kilogram quantity. We first calculate the price of each type of crop sold by households. Then we use the median of the price of each crop as the selling price at the regional level. Finally, we use the agricultural output quantity times the regional price as the household agricultural income. Since the main panel data is at the individual level, we divide the household agricultural income by the number of household members and get the individual agricultural income.

In the full panel dataset, we have individuals from 4 different countries. To make the data comparable, we deflate the income variables to the base year 2010. We deflate the agricultural income using the food price inflation rate from the FAO and the wage income using the wage inflation from the IMF. Then we apply the PPP conversion factor from the World Bank to transform local currency to International US dollars.

2.2 Climate Data

In our research, we collect three climate variables, and they are annual mean NDVI, rainfall, and temperature. The Normalized Difference Vegetation Index (NDVI) is a widely used remote sensing index that measures vegetation health and density by analyzing reflected light in the red and near-infrared (NIR) regions of the electromagnetic spectrum. The higher the NDVI values, the healthier the vegetation. In our research, we use NDVI as a proxy for household-level agricultural food productivity. Using NDVI, rainfall, and temperature helps address the endogeneity problem in agricultural outputs and labor use.

LSMS data provides the longitude and latitude of the household. We can then collect the NDVI, rainfall, and temperature values for each household, as these households are typically located close to their farmland in Africa. The buffer of the GPS coordinator is 2km for urban areas and 5 km for rural areas. We collect the annual mean NDVI data from MODIS, the annual mean rainfall data from CHIRPS, and the annual mean temperature data from ERA_LAND through Google Earth Engine (GEE) based on the GPS coordinates.

As we have discussed in the LSMS data part, there is a three-year gap for each individual. When we observe an individual move from agriculture to non-agriculture, we don't know which year she moved. We need to have the agricultural productivity in all three-year intervals. The advantage of

using climate data for agricultural productivity is that they are observed continuously and can fill the gap in our dataset. In our data, we collected three years of NDVI, rainfall, and temperature values for each individual.

2.3 Climate Shock Construction

2.3.1 Two Drivers of Agricultural Productivity

In the LSMS data, different households plant different types of crops, and even within the same household, they may choose different crops over time. Additionally, there are measurement errors in agricultural output data. This makes it difficult to estimate agricultural productivity accurately. Using NDVI as a proxy helps overcome this challenge. Several studies have discussed the strong correlation between NDVI data and agricultural outputs. Donaldson and Storeygard (2016) shows that satellite data can be effectively used to measure agricultural output at specific locations. Lobell et al. (2015) introduced a general crop yield model to predict crop yields using satellite data. Farmaha et al. (2016) trained a regression model to simulate maize output using NDVI data. Tang et al. (2022) used NDVI data to capture the poverty dynamics in agricultural communities in four African countries.

As a proxy for agricultural productivity, NDVI still suffers from the endogeneity problem, such as the simultaneous determination of agricultural productivity and labor reallocation, and the omission of unobserved local economic factors that affect both agricultural productivity and labor supply. (Branco and Féres, 2021; Emerick, 2018; Jessoe et al., 2018) Therefore, we cannot directly use NDVI as the explanatory variable for the sectoral movement.

We categorize the drivers of agricultural productivity into 2 parts. One is the climate-driven activities, including rainfall and temperature, which are exogenous. The other one is human-driven activities, such as technology, labor allocation, household decisions, etc. In these two drivers, human activities absorb the endogeneity problems. To study the effect of agricultural productivity on labor reallocation across sectors, we focus on the climate-driven agricultural productivity.

2.3.2 Exogeneous Climate Shock

In agricultural productivity, we have both the exogenous driver (climate activity) and the endogenous driver (human activity). To separate these two drivers, we use the Machine Learning Method. We use the NDVI growth rate as the output variable, and temperature growth rate and rainfall growth rate as the main inputs to run the Random Forest Regression.[1]

In the process, data are bootstrapped and randomly selected with replacement to form different train-

[1] We are using growth rate instead of the level because the changes in productivity, which is called shock in this study, affect people's movement decision in the model.

ing datasets in each decision tree. In our panel dataset, we have 79508 individual-year observations. Since we collect three years of climate data for each individual, the number of data points that can be trained explode to 238524. However, not all data points are used in the training. We randomly separate the data into 80% (190819 data points) as the training set and 20% (47705 data points) as the testing set. In the training set, a random subset of features, which includes rainfall and temperature in our research, is used to build the tree to ensure diversity in the model. After the trees are trained, each of them makes a prediction, and then we average the predictions, which is the final prediction for regression.

We apply the regression model to the testing set to make a prediction. The outcomes show that the mean absolute error (MAE) between the real NDVI value and the predicted NDVI value is 0.04. This implies that the regression model using rainfall changes and temperature changes captures the NDVI changes well. Besides, the R^2 score of the regression model is 0.21. This is because we only account for the exogenous activities, and human activities play an important role in the agricultural output. The R^2 score implies that the regression model has some prediction power, but focuses on the climate part of the agricultural productivity.

Finally, we are able to separate the climate-driven part (climate shock), which is predicted in the regression model, and the human-driven part, which is the residual part in the model. A simplified regression function is:

$$NDVI_{it} = \underbrace{\beta_1 Rainfall_{it} + \beta_2 Temperature_{it}}_{\text{Climate Shock}} + \underbrace{residual}_{\text{Human Activities}} \tag{1}$$

Climate shock constructed here represents the short-term and random part of the agricultural productivity. Short-term means that the impact of climate productivity is temporary. In our context, we only observe the 3-year movement in 4 countries, which is a short-term structural change. Therefore, the panel data matches the characteristics of the climate shock. Random means that people might move back and forth. However, as long as the movement happens, this is the first step and the short-term version of the structural change. This indicates that people could move from one sector to another when a certain condition is satisfied.

2.4 Summary Statistics

In our research, the main dependent variable move is a dummy variable that indicates an individual's movements across sectors. If an individual moves from agriculture to non-agriculture (manufacturing or service), the indicator equals 1. If an individual stays in agriculture, the indicator equals 0. We are only able to observe the move once in Ethiopia, Niger, and Nigeria because the observations are from 2 survey waves. In Malawi, we are able to observe 2 moves. Overall, we observe that fewer

report their sectors in all survey years, leading to low observations. Besides, the movement rates are pretty low across countries.

Another dependent variable we use is the log of income. Climate shock affects people's income. When people are rich enough to reach the food subsistence level and cover the transition cost, they will choose to move. Therefore, climate shock is correlated with income. We find that the average income level is the highest in Nigeria (4.077 USD per month). The average income levels in Ethiopia, Malawi, and Niger are lower.

The explanatory variables are Climate_0lag, Climate_1lag, Climate_2lag. Climate_0lag is the present year's climate shock, Climate_1lag is the 1-year lagged climate shock, and Climate_2lag is the 2-year lagged climate shock.For example, in Niger, the observation years are 2011 and 2014. Then, for each individual, the Climate_0lag variable is the climate-driven productivity growth rate between 2013 and 2014, the Climate_1lag variable is the climate-driven productivity growth rate between 2012 and 2013, and the Climate_2lag variable is the climate-driven productivity growth rate between 2011 and 2012.

Other control variables are: (1) Urban: 1 for urban and 0 for rural; (2) Gender: 1 for male and 0 for female; (3) age; (4) age^2; (5) Years of education: different countries have different education system, we transfer the education level into education years to make the variable comparable; (6) Sector: 1 for agriculture, 2 for manufacturing, and 3 for service; (7) Formal: 1 for working in the formal labor market and 0 for working informal sector; (8) Food Subsistence Level: this is from the question "Have you ever faced with the situation where you don't have enough food to eat in the past year", the variable equals 1 if the answer is yes and 0 if the answer is no.

Table 1: Summary Statistics

	Ethiopia		Malawi		Niger		Nigeria	
	Obs	Mean	Obs	Mean	Obs	Mean	Obs	Mean
Move	1,585	0.0498	2,340	0.1107	4,836	0.0682	6,118	0.0209
$Log(income)$	5,468	2.7180	9,155	1.4972	15,278	2.9557	22,346	4.0777
Climate_0lag	16,171	-0.0058	9,255	-0.0058	22,712	0.0008	30,581	-0.0179
Climate_1lag	16,171	0.0046	9,255	-0.0109	22,712	0.0101	30,581	-0.0039
Climate_2lag	16,171	0.0298	9,255	0.0440	22,712	-0.0111	30,581	0.0075
Urban	16,870	0.7212	9,263	0.2535	22,712	0.4158	30,617	0.2955
Gender	16,870	0.4666	9,269	0.5785	22,751	0.4704	30,617	0.4779
Age	16,870	32.5047	9,269	32.2930	22,751	35.1436	30,617	35.6823
Age^2	16,870	1260.271	9,269	1214.432	22,751	1482.407	30,617	1551.601
Education year	16,870	6.3530	9,053	6.8234	18,328	3.8975	21,306	9.8560
Sector	5,713	1.8848	9,050	1.4480	16,015	1.6106	15,756	1.6673
Formal	5,714	0.3271	9,050	0.1093	14,709	0.0537	11,161	0.0814
Food Subsistence Level	16,857	0.2235	9,269	0.6507	22,715	0.4503	30,498	0.2545

*

3 Model

Poor countries suffer from the food security problem(Schultz, 1953), which explains the low structural change from agriculture to non-agriculture. There are a number of people who are below the food subsistence level and are trapped in the agricultural sector. (McArthur and Sachs, 2019; Gollin et al., 2007; Gollin et al., 2002) Meanwhile, there is a transition cost, including mobility cost, human capital acquisition, etc, if the individuals want to reallocate out of agriculture. (Adamopoulos et al., 2024; Deininger et al., 2022; Porzio et al., 2022; Dekle and Vandenbroucke, 2012) In the model, we introduce both the food-subsistence level and the transition cost into the utility function and discuss individuals' sectoral move decisions.

Based on the food-subsistence model by Gollin et al. (2002), we build a two-period model with climate-driven agricultural productivity (climate shock) and transition cost. When the climate shock is positive, households' agricultural income increases. For the individuals above the food-subsistence level, when the increase in income is large enough to cover the transition cost, they will choose to move if the moving utility is higher than the staying utility. For the individuals below the food-subsistence level, if the climate-driven productivity is large enough, they will have enough income to reach the food-subsistence level and pay the transition cost. They are able to decide whether to move or stay.

3.1 Utility and Budget Constraints

According to Schultz (1964), smaller farmers are rational and efficient in maximizing their utility based on constraints such as poor access to technology, lack of credit, etc. In our two-period model, we allow agricultural individuals to make choices between staying in agriculture and moving to non-agriculture to maximize their overall utility.

An individual i chooses consumption c_1 in period 1 and c_2 in period 2. For simplicity, there is no transfer between the two periods. This is consistent with the finding that smallholder farmers in Africa are faced with financial constraints. (Balana and Oyeyemi, 2022; Mukasa et al., 2017; Karlan et al., 2014) The Stone-Geary intertemporal utility function is

$$U_i = \begin{cases} 0, & \text{if } c_{1i} < \gamma \\ log(c_{1i} - \gamma - t * \mathbb{1}\{ifmove = 1\}) + \beta log(c_{2i}), & \text{if } c_{1i} \geq \gamma \end{cases} \tag{2}$$

where

- c_{1i} and c_{2i} are the consumption of household i in both period 1 and period 2
- γ is the subsistence level.
- t is the transition cost if people choose to move from agriculture to non-agriculture.
- β is the discount value.
- if $c_{1i} < \gamma$, household will stay in agriculture.

The budget constraint is:

$$y_{1i} = c_{1i} \tag{3}$$

$$y_{2i} = c_{2i} \tag{4}$$

- The production in period 1 is $y_{1i} = A_i * (1 + \theta_{1i})$. That is, the agricultural production is determined by the human-driven output (A_i) and the climate shock (θ_{1i}). When there is a positive climate shock ($\theta_{1i} > 0$), the agricultural production increases. Similarly, if there is an increase in human-driven productivity, the agricultural production also increases.
- The production in period 2 is $y_{2i} = A_i * (1 + \theta_{2i}) * (1 - \mathbb{1}\{ifmove = 1\}) + M_i * \mathbb{1}\{ifmove = 1\}$. M is the non-agricultural production. If the individual moves, her output in period 2 equals M_i because $(1 - \mathbb{1}\{ifmove = 1\}) = 0$. If not move, her output in period 2 equals $A_i * (1 + \theta_{2i})$ because $\mathbb{1}\{ifmove = 1\} = 0$.
- We assume the individual makes the decision by the end of period 1. Therefore, we use expected income ($\mathbb{E}(1 + \theta_2) = 1$ and $\mathbb{E}(M_i) = \beta_M$) in period 2

We plug the budget constraints into the utility function and get:

$$U_i = \begin{cases} 0, & \text{if } A_i * \theta_{1i} * H_i < \gamma \\ log(A_i * (1+\theta_{1i}) * H_i - \gamma - t * \mathbb{1}\{move = 1\}) & \\ +\beta log(A_i * H_i * (1 - \mathbb{1}\{move = 1\}) + M_i * \mathbb{1}\{move = 1\}), & \text{if } A * \theta_{1i} * H_i \geq \gamma \end{cases} \tag{5}$$

If the individual's first-period income is below the food subsistence level ($A_i * \theta_{1i} * H_i < \gamma$), their utility equals 0 and they will stay in the agricultural sector, which is the first part of Equation (5). If the individual's first-period income reaches the food subsistence level ($A_i * \theta_{1i} * H_i \geq \gamma$), they can choose whether to move or stay by comparing their aggregate utility. When we discuss the move in the empirical part, we mainly focus on the second part of the utility function, that is,

$$log(A_i * (1+\theta_{1i}) * H_i - \gamma - t * \mathbb{1}\{move = 1\}) + \beta log(A_i * H_i * (1 - \mathbb{1}\{move = 1\}) + M_i * \mathbb{1}\{move = 1\}) \tag{6}$$

3.2 Individual Decision

Individuals make the sectoral change decision based on their utility, constraints, and preferences. In the previous section, we show the utility function and constraints. What is missing here is preferences. In the model, we consider 2 types of preference.

1. Preference shock (ε_{stay} and ε_{move}): The shock follows Type I extreme value (Gumbel). These shocks represent idiosyncratic, unobservable preferences that individuals have toward moving vs. staying. This explains the observations that even if two people have the same observed productivity and constraints, they may choose differently due to these unobserved taste shocks. This is the error term in our empirical part.

2. Systematic preference parameters ($pref_{stay,below}$, $pref_{stay,above}$, $pref_{move,below}$, and $pref_{move,above}$): these parameters are defined to show the movement preference of people above and below the food-subsistence level.

    - From the data, we observe that the agricultural income of individuals above and below the food-subsistence level overlaps. There are some individuals with low agricultural income that have enough food, while some individuals with high agricultural income lack food. This is because we didn't observe the assets, loans, wealth level, etc.
    - Besides, we assume people have different innate agricultural skills (Green Thumb) that make them prefer to stay in agriculture. In the data, we only observe the human-driven agricultural output (A_i), and agricultural skills are observed. The preference parameters indicate individuals' preferences between agriculture and non-agriculture.

- In the reduced form, we absorb the systematic preferences by introducing control variables (whether they are above or below the food subsistence level, whether they are urban or rural, whether they are male or female, whether they work in the formal or informal labor market, etc) and fixed effects (region and year).

Overall, individuals are making the decision by comparing the utilities:

$$U_{i,stay} = log(A_i * (1+\theta_{1i}) * H_i - \gamma) + \beta log(A_i * H_i) + pref_{stay,below/above} + \varepsilon_{stay} \tag{7}$$

$$U_{i,move} = log(A_i * (1+\theta_{1i}) * H_i - \gamma - t) + \beta log(M_i) + pref_{move,below/above} + \varepsilon_{move} \tag{8}$$

If $U_{i,stay} \geq U_{i,move}$, individuals choose to stay in agriculture. On the other hand, if $U_{i,stay} < U_{i,move}$, individuals choose to move to non-agriculture.

4 Empirical Outcomes

To see whether people will move away from agriculture to non-agriculture empirically when they get richer, we choose to regress the probability of movement on agricultural productivity. The human-driven agricultural productivity, such as seed improvement, fertilizer efficiency, the usage of organic inputs, etc, in sub-Saharan Africa is stagnated. (Gollin, 2023; Suri and Udry, 2022; Erenstein et al., 2022; Jayne and Sanchez, 2021) Besides, the human-driven agricultural productivity leads to the endogeneity problem, as we have mentioned in Section 2.3. Therefore, we use the climate-driven agricultural productivity, which is both exogenous and variant.

In our reduced form, we estimate the relationship between the movement rate and the climate shock. The main regression function is the logit regression:

$$Pr(move_{it}) = \alpha + \beta_1 \Delta climate_{it} + \beta_2 \Delta climate_{it}^2 + \beta_3 age + \beta_4 age^2 + \beta_5 educ_year + urban_{it} + gender_{it} + formal_{it} + region_i + year_t + \varepsilon_{it} \tag{9}$$

where i represents individuals and t is years.[2] $move_{it}$ is a dummy variable that indicates an individual's movements across sectors. If an individual moves from agriculture to non-agriculture (manufacturing or service), the indicator equals 1. If an individual stays in agriculture, the indicator equals 0. $\Delta climate_{it}$ is the climate shock variable, which is the growth rate of climate-driven agricultural productivity. The control variables are age, years of education, urban, gender, and formal. We fix the region effect and the year effect.

[2] In Ethiopia, Niger, and Nigeria, we only observe two survey years, which means we capture one move record of each individual in those countries. Therefore, we are not able to fix at the individual level for the main logit regression that checks the movement decision and climate shock. Instead, we use as many control variables as possible.

As shown in the model part, the idea behind the regression is that with the positive climate shock, the agricultural income increases, and people have enough income to cover the transition cost if they choose to move. To show that people's income increases in climate shock, we also run the regression between the agricultural income and the climate shock:

$$ln_inc_{it} = \alpha + \beta_1 \Delta climate_{it} + \beta_2 \Delta climate_{it}^2 + \beta_3 age + \beta_4 age^2 + \beta_5 educ_year + individual_i + year_t + \varepsilon_{it} \tag{10}$$

where i represents individuals and t is years.[3] ln_inc_{it} is the log of agricultural output for those who do not reply the wage income and the log of wage income for to individuals who receive wages in agricultural.

In the Model part, we have discussed the different decision processes of people who are below the food subsistence level and of people who are above. Therefore, we look at two groups seperately. Besides, it is shown that climate changes such as the rainfall change and temperature change has nonlinear impact on migration. (Nawrotzki et al., 2017;Bohra-Mishra et al., 2014) When estimating the relationship between the probability of move from agriculture to non-agriculture and climate shocks, we are using both the climate shock and the climate shock squared. Similarly, we estimate the nonlinear impact of climate change on the agricultural income.

4.1 Agricultural Income and Climate Shock

In the regressions, we break the individuals into two groups. Table 2 shows the regression outcomes between agricultural income and climate shocks for the group that is below the food subsistence level. We find that the agricultural income increases as the climate shock value increases in the first order across all countries. In Niger and Ethiopia, however, the impact of climate shock decreases in the second order. This implies that the marginal growth rate of agricultural income caused by climate shock decrease, which is consistent with the nonlinear climate shock argument. It is worth noting that the coefficients of the climate shock square in Niger and Ethiopia are both pretty large, indicating that the impact of the climate shock on the agricultural income peaks and goes down when the climate shock value is big. In Nigeria and Malawi, the coefficient of climate shock is positive and significant, while the second order coefficients are positive but insignificant. This means that the impact of climate shock is positive and linear.

In Table 3, we estimate the effect of climate shock on agricultural income for the group that is above the food subsistence level. The regression outcomes are mixed. In Niger, we still have a positive coefficients in the first order and a negative second order, implying that the agricultural income

[3] As for the income regression, we have two-year records of each individual, and we fix the regression at the individual level.

first increase, and peak and decrease if the climate shock is big. In Nigeria, both the coefficients of climate shock and climate shock square are positive, showing a positive increase in agricultural income. In Malawi, the agricultural income increases as the climate shock becomes more negative or more positive. As for Ethiopia, the income does not change in climate shock.

Overall, there is a positive correlation between the agricultural income and the climate shock, except for the above group in Ethiopia. [4]

Table 2: Agricultural Income and Climate Shock (Below Food-subsistence Group)

	(1)	(2)	(3)	(4)
	Niger	Nigeria	Malawi	Ethiopia
climate_0lag	2.408**	4.276**	5.701***	9.912***
	(2.52)	(2.13)	(3.42)	(2.99)
$climate_0lag^2$	-47.73***	3.353	19.97	-108.9**
	(-3.03)	(0.14)	(0.67)	(-2.33)
age	0.120***	-0.197**	-0.121*	0.211
	(3.94)	(-2.60)	(-2.07)	(1.63)
age2	-0.000397	0.00170*	0.000257	-0.00188
	(-1.15)	(2.14)	(0.33)	(-1.22)
educ_year	0.0302	0.0325	0.0575	0.00875
	(1.21)	(0.83)	(1.29)	(0.11)
_cons	5.202***	12.62***	8.899***	-0.436
	(8.46)	(7.84)	(8.48)	(-0.16)
individual fixed effect	yes	yes	yes	yes
year fixed effect	yes	yes	yes	yes
N	3771	1404	5068	356

$*$ t statistics in parentheses

$**$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[4] In Table 2 and Table 3, the numbers of observations in all countries are lower than the numbers we observed in the Summary Statistics table. This is because our regression only applies to the individual with at least two years' record in agriculture. There are individuals who either never worked in agriculture or have only one year's record in agriculture. Besides, several individuals either fail to report their income (either wage income or agricultural output) in the first or second survey years. Finally, there are households in some regions losing track in the second survey due to conflicts or resampling.

Table 3: Agricultural Income and Climate Shock (Above Food-subsistence Group)

	(1)	(2)	(3)	(4)
	Niger	Nigeria	Malawi	Ethiopia
climate_0lag	3.486***	5.833***	0.701	-2.623
	(3.29)	(6.11)	(0.21)	(-1.42)
$climate_0lag^2$	-46.01**	37.38***	168.6***	-0.438
	(-2.34)	(3.14)	(2.95)	(-0.01)
age	0.0812*	-0.0901	-0.142	-0.0543
	(2.01)	(-1.94)	(-1.52)	(-0.47)
age2	0.000193	0.000909	0.00142	0.00167
	(0.40)	(1.77)	(1.13)	(1.17)
educ_year	-0.0520*	-0.0419*	-0.00831	-0.0687
	(-2.03)	(-2.05)	(-0.11)	(-1.10)
_cons	6.333***	11.19***	8.794***	4.532*
	(7.94)	(11.82)	(5.21)	(2.10)
individual fixed effect	yes	yes	yes	yes
year fixed effect	yes	yes	yes	yes
N	3106	4139	1749	1094

$*$ t statistics in parentheses

$**$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.2 Sectoral Move and Climate Shock

In the previous section, we have shown the positive correlation between agricultural income and the present climate shock. Now, we study the relationship between sectoral move and climate shock. The challenge in estimating the impact of climate shock on move is that we have a 3-year gap in the data. This means we only observe the original sector in the first year and the destiny sector in the last year. We don't know which year the individual moves from one sector to another. Fortunately, the climate shock variables have a full record that can fill the gaps. We regress the sectoral move on the present climate shock, the 1-year lagged climate shock, and the 2-year lagged climate shock separately.[5]

We find that the 1-year lagged climate shock has a positive and significant impact on sectoral move in the group that is below the food subsistence level in Niger, Nigeria, and Malawi, which is shown

[5]There is a serial correlation between the climate shocks, and we decide not to use all of them in the regression function.

in Table 4. This is consistent with the positive correlation in Table 2. The negative and insignificant coefficient for the 1-year lagged climate shock further enhances the credibility of the regression outcomes. The special land policy that farmers might lose inheritance for land rights when absent from the land keeps them staying in agriculture (Ghebru et al., 2018; Clech et al., 2020; Zewdu and Malek, 2010), even though they are richer in the positive climate shock.[6] For Ethiopian farmers who are not able to leave agriculture, an increase in the climate shock increases their income, and thus they are less likely to leave. This explains the negative coefficient of Climate_1lag.

For the group that is above the food subsistence level, in Table 5, the impact of the 1-year lagged climate shock is insignificant across countries. The flat marginal effect of climate shock indicates that individuals above the food subsistence level are less likely to move in a positive climate shock compared to the below group. According to the model, individuals can make decisions between moving and staying once they reach the food subsistence level. In the above group, those individuals who prefer to move have already leave the agricultural sector, and the rest are the ones who prefer to stay in agriculture.

Overall, we find that the climate shock increases the sectoral move among people below the food subsistence level in Niger, Nigeria, and Malawi. Besides, the impact of climate shock is negative but insignificant among people above the food subsistence level in Nigeria, Malawi, and Ethiopia.[7]

[6] Some studies have shown that the land policy pushes young workers from agricultural to non-agricultural sectors, but the main source of migration has nothing to do with agricultural income. (Bezu and Holden, 2014; Kosec et al., 2018)

[7] In Table 4 and Table 5, the numbers of observations in all countries are lower than the numbers we observed in the Summary Statistics table. The reason is similar to that for Tables 2 and 3.

Table 4: Sectoral Move and Climate Shock (Below Food-subsistence Group)

	(1)	(2)	(3)	(4)
	Niger	Nigeria	Malawi	Ethiopia
climate_1lag	7.40*	29.33**	9.792**	-12.58
	(1.90)	(2.04)	(1.98)	(-1.17)
$climate_1lag^2$	82.51	-335.3	55.85	-302.1
	(1.39)	(-1.63)	(0.55)	(-1.37)
age	0.0398	0.0944	0.0251	0.364
	(0.94)	(0.90)	(0.57)	(1.80)
age2	-0.000335	-0.00115	-0.000204	-0.00448
	(-0.69)	(-0.94)	(-0.37)	(-1.74)
educ_year	0.0761**	0.0907	0.130***	0.112
	(2.08)	(1.14)	(4.03)	(0.96)
_cons	–2.804**	-5.719*	-6.403***	-10.08*
	(-2.98)	(-2.32)	(-4.03)	(-2.43)
urban fixed effect	yes	yes	yes	yes
gender fixed effect	yes	yes	yes	yes
formal fixed effect	yes	yes	yes	yes
region fixed effect	yes	yes	yes	yes
year fixed effect	yes	yes	yes	yes
N	1684	423	1711	216

$*$ t statistics in parentheses

$**$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Sectoral Move and Climate Shock (Above Food-subsistence Group)

	(1)	(2)	(3)	(4)
	Niger	Nigeria	Malawi	Ethiopia
climate_1lag	2.702	-0.989	-4.281	-4.429
	(0.62)	(-0.15)	(-0.49)	(-0.44)
$climate_1lag^2$	39.32	-62.32	-110.1	-126.0
	(0.65)	(-0.57)	(-0.68)	(-0.89)
age	0.0288	0.116	0.0373	0.0504
	(0.60)	(1.35)	(0.54)	(0.45)
age2	-0.000236	-0.00185	-0.000328	-0.000765
	(-0.43)	(-1.74)	(-0.39)	(-0.55)
educ_year	0.126**	0.202***	0.0601	0.0571
	(3.22)	(3.37)	(1.46)	(0.99)
_cons	-2.645*	-9.275***	-5.241**	-3.735
	(-2.50)	(-4.71)	(-2.85)	(-1.61)
urban fixed effect	yes	yes	yes	yes
gender fixed effect	yes	yes	yes	yes
formal fixed effect	yes	yes	yes	yes
region fixed effect	yes	yes	yes	yes
year fixed effect	yes	yes	yes	yes
N	1316	1433	481	654

$*$ t statistics in parentheses

$**$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.3 Robustness Check

4.3.1 Negative vs Positive Climate Shocks

In Section 4, we have shown that the probability of moving out of agriculture increases in the 1-year lagged climate shock for the below group. According to the model, larger climate shock leads to higher agricultural output and thus more people are moving out of the agricultural sector. If the positive climate shock covers poor farmers' transition cost, we should find that the probability of the sectoral move increases in positive climate shock variables. Therefore, we separate the climate shock into positive and negative ones and then run the same logit regression for both the below group and the above group.

In Table 6, we find that the coefficients of climate shock are positive and significant in the below group when the climate shock is larger than 0 in Niger, Nigeria, and Malawi. The coefficient in Ethiopia is still negative and insignificant, which is consistent with the outcome in Table 4. On the other hand, the coefficient is insignificant but positive in Niger and Malawi, and negative but insignificant in Nigeria and Ethiopia. [8]As for the above group (Table 7), we find no significance in climate shock on sectoral movement in most cases. The only significance comes from Niger when the climate shock is positive, where the probability of moving decrease in positive climate shock. Overall, the findings are consistent with the outcomes in Section 4.2.

Table 6: Sectoral Move and Climate Shock (Below Food-subsistence Group)

	$climate_1lag < 0$				$climate_1lag > 0$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Niger	Nigeria	Malawi	Ethiopia	Niger	Nigeria	Malawi	Ethiopia
climate_1lag	34.32	-145.03	5.390	-589.0	11.37*	103.1*	36.02**	-39.73
	(1.41)	(-0.46))	(0.22)	(-0.98)	(1.68)	(1.90)	(2.03)	(-1.47)
c.climate_1lagc.climate_1lag	329.6	-25515.38	-107.6	-7901.4		-1060.4*		
	(1.09)	(-1.37)	(-0.30)	(-1.10)	(2.00)	(-1.87)	(1.35)	(1.07)
age	-0.00199	1.637	0.0128	7.616	0.0925	0.0588	0.145	0.240
	(-0.03)	(0.98))	(0.26)	(1.27)	(1.47)	(0.51)	(0.95)	(0.89)
age2	0.0000554	-0.0213	0.0000527	-0.119	-0.000834	-0.000836	-0.00245	-0.00286
	(0.08)	(-0.98)	(0.09)	(-1.27)	(-1.19)	(-0.61)	(-1.11)	(-0.89)
educ_year	0.0779	-1.186	0.112**	0.0666	0.0643	0.0652	0.195*	0.0218
	(1.47)	(-1.27))	(3.05)	(0.10)	(1.23)	(0.66)	(2.55)	(0.14)
_cons	-1.771	2.618	-3.969**	-137.6	-3.758**	-6.641*	-8.278**	-7.424
	(-1.28)	(0.97))	(-2.58)	(-1.26)	(-2.64)	(-2.16)	(-2.85)	(-1.23)
urban fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
gender fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
formal fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
region fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
year fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
N	872	120	1139	65	708	294	440	122

∗ *t* statistics in parentheses

∗∗ ∗ $p < 0.05$, ∗∗ $p < 0.01$, ∗∗∗ $p < 0.001$

[8]The large, negative, and insignificant outcome in Nigeria and Ethiopia is due to the small negative sample size (120 for Niger and 65 for Ethiopia).

Table 7: Sectoral Move and Climate Shock (Above Food-subsistence Group)

	$climate_1lag < 0$				$climate_1lag > 0$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Niger	Nigeria	Malawi	Ethiopia	Niger	Nigeria	Malawi	Ethiopia
climate_1lag	44.19	-50.70	-19.43	-11.36	-42.24*	-14.92	41.86	-20.85
	(1.38)	(-0.85)	(-0.54)	(-0.12)	(-1.77)	(-0.58)	(0.41)	(-0.52)
c.climate_1lagc.climate_1lag	450.1	-1055.7	-438.6	-628.1	452.5**	128.4	-433.5	45.02
	(1.20)	(-1.14)	(-0.91)	(-0.39)	(2.00)	(0.58)	(-0.25)	(0.14)
age	0.0624	0.0868	0.0257	0.0565	0.0330	0.109	-0.101	0.0927
	(0.78)	(0.54)	(0.32)	(0.35)	(0.52)	(1.01)	(-0.53)	(0.54)
age2	-0.000680	-0.00178	-0.000274	-0.000616	-0.000252	-0.00168	0.00156	-0.00152
	(-0.73)	(-0.85)	(-0.28)	(-0.33)	(-0.35)	(-1.30)	(0.67)	(-0.68)
educ_year	0.139**	0.268***	0.0349	0.0435	0.116**	0.162**	0.0982	0.0890
	(2.30)	(2.68)	(0.70)	(0.50)	(2.02)	(2.03)	(0.77)	(0.98)
_cons	-3.039	-9.740**	-1.224	-5.178	-1.834	-8.804***	-2.054	-3.120
	(-1.70)	(-2.99)	(-0.51)	(-1.28)	(-1.29)	(-3.36)	(-0.53)	(-0.93)
urban fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
gender fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
formal fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
region fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
year fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
N	577	541	326	228	732	892	101	374

$*$ t statistics in parentheses

$**$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.3.2 Sectoral Move: Non-agriculture to Agriculture

With the positive climate shock, it is possible that agriculture becomes more attractive. It seems that farmers could gain more from agriculture, and there is no need for them to move. One explanation is that the climate-shock-led productivity growth is not enough to fill the gap between agricultural productivity and non-agricultural productivity, but is enough to cover the transition cost for the poor agricultural households.

To show this, we further check the movement decision of those who work in the non-agriculture sector. We find that people of both below and above the food-subsistence level group are not moving out of the non-agriculture sector. In Niger, the coefficients are significant and negative, which implies that the probability of people moving out of non-agricultural decreases in the climate shock. These results mean that the productivity growth in agriculture is not attracting those who find jobs in the non-agriculture sector. Besides, the results can serve as a signal that though we are looking at the temporary sectoral movement, it is likely that people will stick in non-agriculture once they move out of agriculture and turn the temporary movement into a permanent one.

Table 8: Sectoral Move and Climate Shock (Non-agriculture to Agriculture)

	Below Group				Above Group			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Niger	Nigeria	Malawi	Ethiopia	Niger	Nigeria	Malawi	Ethiopia
climate_1lag	-7.204***	1.518	-0.205	-0.133	-6.502***	3.131	-2.437	1.702
	(-2.71)	(0.27)	(-0.10)	(-0.04)	(-4.23)	(0.91)	(-1.01)	(0.92)
c.climate_1lagc.climate_1lag	102.0***	-21.72	49.42	-4.742	119.0***	-42.10	-62.96	7.662
	(3.83)	(-0.35)	(1.18)	(-0.11)	(6.82)	(-1.20)	(-1.19)	(0.31)
age	-0.0607**	-0.247***	-0.135***	-0.194***	-0.0911***	-0.250***	-0.171***	-0.183***
	(-2.64)	(-3.62)	(-6.32)	(-3.99)	(-5.60)	(-5.49)	(-6.43)	(-6.60)
age2	0.000457	0.00274**	0.00120***	0.00242***	0.000957***	0.00307***	0.00157***	0.00243***
	(1.62)	(3.19)	(4.47)	(3.90)	(4.77)	(5.27)	(4.72)	(6.66)
educ_year	-0.0773***	-0.153**	-0.196***	-0.115**	-0.120***	-0.0898**	-0.254***	-0.0780***
	(-3.57)	(-3.11)	(-11.64)	(-3.03)	(-9.27)	(-2.64)	(-13.57)	(-5.18)
_cons	1.770***	10.62***	7.967***	5.394***	2.175***	8.771***	7.203***	6.545***
	(3.84)	(6.45)	(9.52)	(4.37)	(6.57)	(9.14)	(10.57)	(11.17)
urban fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
gender fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
formal fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
region fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
year fixed effect	yes	yes	yes	yes	yes	yes	yes	yes
N	2883	470	4279	699	5591	1338	2233	3862

$*$ t statistics in parentheses

$**$ * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5 Structural Estimation and Counterfactual Analysis

In the empirical part, we have shown that the probability of moving from agriculture to non-agriculture increases in climate shock for people who are below the food subsistence level because the agricultural income increase. The impact of climate shock, however, is relatively temporary and random. The mean value of the climate shock is about 0 in the long run. We run the simulation and estimate the parameters in the model. Then we are able to analyze the situation where individuals are subsidized with the transition cost and see if people are more likely to move from agricultural to non-agriculture.

From the structural model, we have 7 parameters to estimate, and they are:

1. β: The discount value between period 1 and period 2 is set to be 0.95.
2. γ: the food subsistence level.
3. t: transition cost.
4. $pref_{stay,above}$: preference for staying when above the subsistence level. They might be better at agriculture than we observed/assumed.

5. $pref_{move,above}$: Preference for moving when above the subsistence level. They might have an extra asset and are not that good at agriculture.

6. $pref_{stay,below}$: Preference for staying when below the subsistence level. They may be in debt and not good at agriculture.

7. $pref_{move,below}$: Preference for move when below the subsistence level.

We apply the Simulated Method of Moments to estimate the parameters. In the four African countries we study, the movement rates are close to 0. In this simulation setting, using group-specific movement rates or the differences across groups as moment conditions often leads to a flat objective function, meaning the loss surface changes very little in response to variations in the parameters. This flatness arises because group-level means are relatively insensitive to small parameter changes —especially when the simulated probabilities saturate near 0 or 1. As a result, the estimation procedure may stall at local minima or fail to find meaningful gradients for optimization.

To address this, we use regression coefficients as the moments instead. These coefficients capture continuous relationships between movement decisions and covariates across all individuals. Unlike grouped means, regression-based moments are more sensitive to parameter shifts, especially when the model affects the slope or level of the decision rule. This improves the curvature of the loss surface and facilitates smoother convergence in the Simulated Method of Moments (SMM) estimation.

In the simulation-based estimation, we use a set of moment conditions that include both regression coefficients and group-specific migration rates. Specifically, we estimate an OLS regression of the binary migration outcome on key covariates—such as education, age, and food subsistence level—and use the resulting regression coefficients as simulated moments. These coefficients capture the empirical relationships between individual characteristics and the likelihood of migration. Additionally, we include aggregate and subgroup migration rates (e.g., overall movement rate and movement rates by food-subsistence level) as moments. Together, these moments guide the calibration of model parameters to ensure that the simulated data closely replicate patterns observed in the actual data. Overall, the 7 moments we target are:

1. Constant term

2. Coefficient on years of education

3. Coefficient on age

4. Coefficients on the food subsistence level

5. Overall movement rate

6. Movement rate for the individuals who are below the food subsistence level

7. Movement rate for the individuals who are above the food subsistence level

We also test the robustness of our model and use other moments, including the movement rate in four quartiles of climate shock and four quartiles of age. After we get the coefficients, we set the subsidy value equals the transition cost parameter value in the counterfactual case. In this way, we can analyze the changes in the movement rates with extra income.

5.1 Estimation Results

Table 9 shows the SMM estimation outcomes in four countries. The food-subsistence level is 3 USD in Malawi, 7 USD in Niger, 8 USD in Ethiopia, and 61 USD in Nigeria. Since we are using deflated monthly income in the estimation, γ in each country represents the monthly agriculture income that an individual need to have enough food to consume. The food-subsistence level should be correlated to the GDP per capita in each country. If we look at the GDP per capita in those countries in 2019, the monthly GDP per Capita is 48 USD in Malawi, 47 USD in Niger, 69 USD in Ethiopia, and 189 USD in Nigeria. GDP per Capita is a lot larger in Nigeria compared to the other three countries. At the same time, the food-subsistence level is also a lot larger in Nigeria than in the other countries.

As for the transition cost t, the parameter varies across countries. If we compute the share of the transition cost on the food-subsistence level, we find that the transition cost is the largest in Niger (100%), followed by Ethiopia (50%). This means that it is harder for individuals to transfer from agriculture to non-agriculture. This is consistent with the regression outcomes that the coefficient of the climate shock is relatively smaller and only at 10% significance in Niger, and the coefficient of climate shock is insignificant in Ethiopia.[9] In these two countries, the income increase from climate shock is not large enough to cover the moving cost. The transition cost in Nigeria and Malawi is about one-thirds of the food subsistence level, which explains the higher probability of movement in the empirical part.

Finally, we look at the preference parameters. Among those who are below the food-subsistence level, the preference parameter of the moving individual ($pref_{move,below}$) is always larger than that of the staying individuals. This implies that the below and move people either have more unobserved assets or a strong preference towards non-agriculture than the below and stay people. Similar for the above and move people.

We list the fit of our simulation for the 7 moments we chose in Table 10. The simulated values fits the empirical values well for the overall movement and the movement by the food-subsistence level in all countries. As for the coefficients, most of the coefficients from the simulated data have the same sign as the actual data. Besides, the values of simulated coefficients are relatively close

[9] As we have mentioned, there is a special land policy that keeps farmers on the land. The transition cost does not take the policy into account. Therefore, the transition cost is not as large as we expect to see.

to the actual data regression outcomes, considering the sensitivity of the regression outcomes in the simulated data.

Table 9: Model Fit

Parameter	Meaning	Niger	Nigeria	Malawi	Ethiopia
γ	Food-subsistence Level	7	61	3	8
		(0.0013)	(0.0013)	(0.0018)	(0.0010)
t	transition cost	7	21	1	4
		(0.0010)	(0.0035)	(0.0004)	(0.0022)
β	Discount value	0.95	0.95	0.95	0.95
$pref_{stay,above}$	Preference of stay among the above	0.5	0	-1	0
		(0.0005)	(0.0011)	(0.0028)	(0.0046)
$pref_{move,above}$	Preference of move among the above	0.5	1	0	1
		(0.0018)	(0.0037)	(0.0026)	(0.0008)
$pref_{stay,below}$	Preference of stay among the below	0.5	-0.5	-0.5	0.5
		(0.0009)	(0.0006)	(0.0016)	(0.0008)
$pref_{move,below}$	Preference of move among the below	0.5	0.5	1	1
		(0.0017)	(0.0009)	(0.0026)	(0.0044)

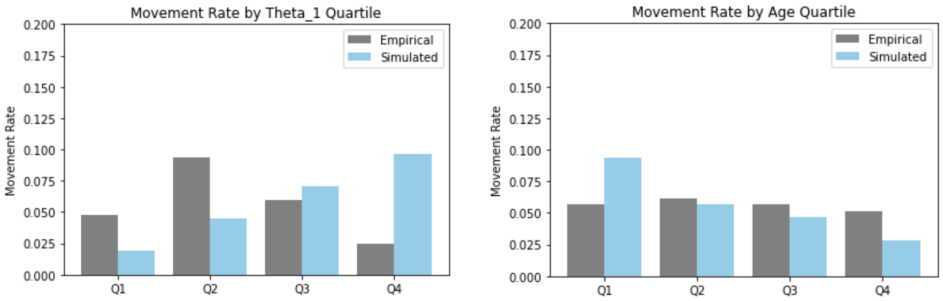
Table 10: Model Fit

	Niger		Nigeria		Malawi		Ethiopia	
Moments	Empirical	Simulated	Empirical	Simulated	Empirical	Simulated	Empirical	Simulated
Constant term	0.0203	0.0612	-0.0919	-0.0891	-0.0288	0.0156	-0.1156	-0.0984
Coefficient on years of education	0.0113	0.0152	0.0084	0.0019	0.0215	0.0276	0.0210	0.0307
Coefficient on age	0.0006	-0.0005	0.0013	0.0029	0.0021	0.0004	0.0030	0.0017
Coefficients on the food subsistence level	-0.0064	-0.0101	0.0115	0.0053	-0.0700	-0.0979	0.0499	0.0246
Overall Movement	0.0565	0.0576	0.0312	0.0258	0.1110	0.1110	0.0718	0.0745
Movement of the above food-subsistence level	0.0610	0.0650	0.0285	0.0239	0.1953	0.2256	0.0581	0.0666
Movement of the below food-subsistence level	0.0530	0.0517	0.0402	0.0324	0.0817	0.0712	0.1150	0.0996

5.2 Robustness Check

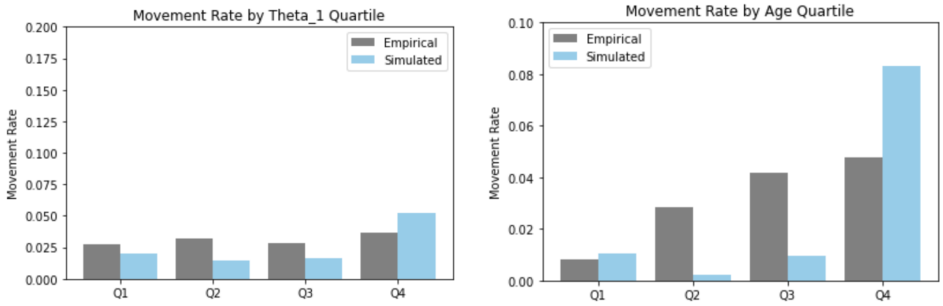
We use both the quartiles of the climate shock and the quartiles of the age. We find that the simulated movement rates by groups fit the actual data either in both climate shock and age quartiles (Malawi and Ethiopia), or in at least one of them (Niger and Nigeria).

Figure 1: Robustness (Niger)



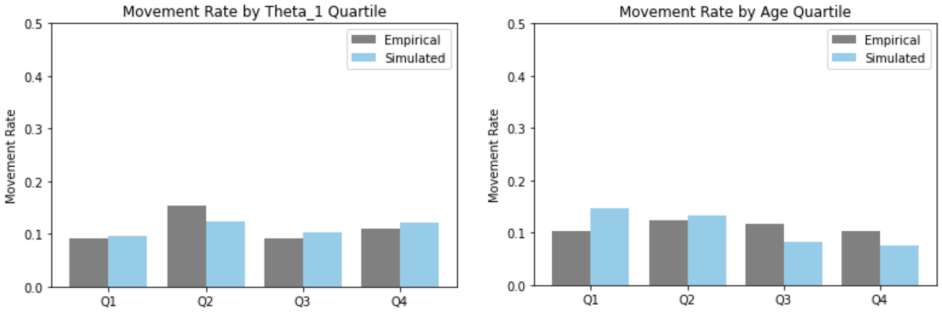
(a) Movement Rate by Theta Quartile **(b)** Movement Rate by Age Quartile

Figure 2: Robustness (Nigeria)



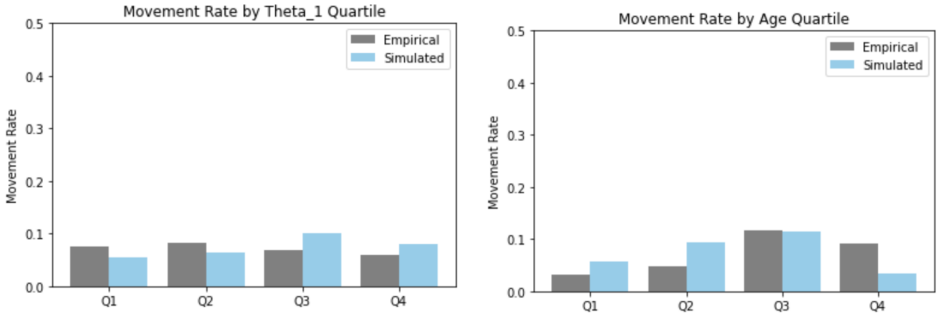
(a) Movement Rate by Theta Quartile **(b)** Movement Rate by Age Quartile

Figure 3: Robustness (Malawi)



(a) Movement Rate by Theta Quartile **(b)** Movement Rate by Age Quartile

Figure 4: Robustness (Ethiopia)



(a) Movement Rate by Theta Quartile **(b)** Movement Rate by Age Quartile

Note:

5.3 Counterfactual Analysis: Subsidize the transition cost

In this counterfactual case, people recevie a subsidy that is equal to the average transition cost, that is 7 USD in Niger, 21 USD in Nigeria, 1 USD in Malawi, and 4 USD in Ethiopia. This means that the transition cost is covered by the individuals. We find that the overall movement rates increase a lot in all countries. Besides, different food-subsistence group benefits from the subsidy differently. In Niger and Malawi, individuals below the food-subsistence level are more likely to move to non-agriculture. The GDP per Capita level is the lowest in these two countries. People are poor, and people in the below group are even poor. When there is a subsidy, it is easier for people in the above group to move.

In Nigeria and Ethiopia, on the other hand, individuals above the food-subsistence level are more likely to move to non-agriculture. This is because people in these two countries are relatively richer. It is easier for the below group individuals to reach the food subsistence level. When the subsidy comes, the below group individuals have higher income that helps them to reach the food-subsistence level and cover a part of the transition cost. With extra assets and a higher preference towards the non-agricultural sector ($pref_{move,below}$), they choose to move.

Table 11: Counterfactual Case 1: t=0

	Niger		Nigeria		Malawi		Ethiopia	
Moments	Empirical	Simulated	Empirical	Simulated	Empirical	Simulated	Empirical	Simulated
Overall	0.0565	0.4180	0.0312	0.3816	0.1110	0.3381	0.0718	0.4517
Movement of the above food-subsistence level	0.0610	0.4305	0.0285	0.3790	0.1953	0.4461	0.0581	0.4370
Movement of the below food-subsistence level	0.0530	0.4084	0.0402	0.3904	0.0817	0.3006	0.1150	0.4981

*

6 Conclusion

This paper develops a micro-founded framework to understand how agricultural income, food subsistence constraints, and transition costs jointly shape labor reallocation decisions in the context of structural transformation. We build a two-period model with Stone–Geary preferences and endogenous migration costs, and empirically examine individual-level transition patterns using panel data from four African countries. Our key innovation lies in the construction of a generalizable climate shock variable—derived from NDVI, temperature, and rainfall using machine learning techniques—that serves as a proxy for agricultural income shocks. Empirically, we show that individuals below the food-subsistence threshold are more likely to move following positive climate shocks, while those above are less responsive, likely due to self-selection or preference for farming. We then simulate counterfactual policies and find that subsidizing transition costs substantially increases sectoral mobility, particularly among different subsistence groups depending on the country context.

Our findings have several implications. First, the climate shock variable we develop can be applied more broadly in future research without requiring extreme weather or technological events, making it a powerful tool for analyzing productivity shocks in various agricultural settings. Second, our structural model offers a policy-relevant framework for estimating the level of subsidy required to help households overcome food insecurity and enable transitions out of agriculture. Simulations suggest that such targeted subsidies can significantly accelerate structural change. However, the same policy may have heterogeneous effects across food-subsistence groups and countries, highlighting the importance of context-specific design in labor mobility interventions.

A limitation of our approach is that both the model and the data are short-term in nature; individuals may migrate temporarily and return to agriculture over time. Nevertheless, initiating the first movement out of low-productivity sectors is a critical step toward long-term structural transformation. Our study thus provides new micro-level evidence and tools to support policy efforts that foster inclusive economic growth in low-income agrarian economies.

References

Adamopoulos, T., Brandt, L., Chen, C., Restuccia, D., & Wei, X. (2024). Land security and mobility frictions. *The Quarterly Journal of Economics*, *139*(3), 1941–1987.

Adamopoulos, T., & Restuccia, D. (2020). Land reform and productivity: A quantitative analysis with micro data. *American Economic Journal: Macroeconomics*, *12*(3), 1–39.

AfDB, A. (2016). African development report 2015-growth, poverty and inequality nexus: Overcoming barriers to sustainable development. *African Development Report*.

Aristizabal-Ramirez, M., Leahy, J. V., & Tesar, L. (2022). *A north-south model of structural change and growth* (tech. rep.). National Bureau of Economic Research.

Balana, B. B., & Oyeyemi, M. A. (2022). Agricultural credit constraints in smallholder farming in developing countries: Evidence from nigeria. *World Development Sustainability*, *1*, 100012.

Bambio, Y., Deb, A., & Kazianga, H. (2022). Exposure to agricultural technologies and adoption: The west africa agricultural productivity program in ghana, senegal and mali. *Food Policy*, *113*, 102288.

Barrett, C. B., Christian, P., & Shiferaw, B. A. (2017). The structural transformation of african agriculture and rural spaces: Introduction to a special section. *Agricultural economics*, *48*(S1), 5–10.

Baseler, T. (2023). Hidden income and the perceived returns to migration. *American Economic Journal: Applied Economics*, *15*(4), 321–352.

Baysan, C., Dar, M. H., Emerick, K., Li, Z., & Sadoulet, E. (2024). The agricultural wage gap within rural villages. *Journal of Development Economics*, *168*, 103270.

Beegle, K., & Christiaensen, L. (2019). *Accelerating poverty reduction in africa*. World Bank Publications.

Bentze, T., & Wollburg, P. (2024). *A longitudinal cross-country dataset on agricultural productivity and welfare in sub-saharan africa*. World Bank.

Bezu, S., & Holden, S. (2014). Are rural youth in ethiopia abandoning agriculture? *World Development*, *64*, 259–272.

Bohra-Mishra, P., Oppenheimer, M., & Hsiang, S. M. (2014). Nonlinear permanent migration response to climatic variations but minimal response to disasters. *Proceedings of the national academy of sciences*, *111*(27), 9780–9785.

Branco, D., & Féres, J. (2021). Weather shocks and labor allocation: Evidence from rural brazil. *American Journal of Agricultural Economics*, *103*(4), 1359–1377.

Buera, F. J., & Kaboski, J. P. (2009). Can traditional theories of structural change fit the data? *Journal of the European Economic Association*, *7*(2-3), 469–477.

Bustos, P., Caprettini, B., & Ponticelli, J. (2016). Agricultural productivity and structural transformation: Evidence from brazil. *American Economic Review*, *106*(6), 1320–1365.

Cattaneo, C., & Peri, G. (2016). The migration response to increasing temperatures. *Journal of development economics*, *122*, 127–146.

Christiaensen, L., & Maertens, M. (2022). Rural employment in africa: Trends and challenges. *Annual Review of Resource Economics*, *14*(1), 267–289.

Clech, L., Jones, J. H., & Gibson, M. (2020). Inequality in the household and rural–urban migration in ethiopian farmers. *Evolutionary Human Sciences*, *2*, e9.

Colmer, J. (2021). Temperature, labor reallocation, and industrial production: Evidence from india. *American Economic Journal: Applied Economics*, *13*(4), 101–124.

Comin, D., Lashkari, D., & Mestieri, M. (2021). Structural change with long-run income and price effects. *Econometrica*, *89*(1), 311–374.

De Janvry, A., Duquennois, C., & Sadoulet, E. (2022). Labor calendars and rural poverty: A case study for malawi. *Food Policy*, *109*, 102255.

De Magalhães, L., & Santaeulàlia-Llopis, R. (2018). The consumption, income, and wealth of the poorest: An empirical analysis of economic inequality in rural and urban sub-saharan africa for macroeconomists. *Journal of Development Economics*, *134*, 350–371.

Deininger, K., Jin, S., & Ma, M. (2022). Structural transformation of the agricultural sector in low- and middle-income economies. *Annual Review of Resource Economics*, *14*(1), 221–241.

Dekle, R., & Vandenbroucke, G. (2012). A quantitative analysis of china's structural transformation. *Journal of Economic Dynamics and Control*, *36*(1), 119–135.

Diao, X., McMillan, M., & Rodrik, D. (2019). The recent growth boom in developing economies: A structural-change perspective. In *The palgrave handbook of development economics: Critical reflections on globalisation and development* (pp. 281–334). Springer.

Dillon, B., & Barrett, C. B. (2017). Agricultural factor markets in sub-saharan africa: An updated view with formal tests for market failure. *Food policy*, *67*, 64–77.

Donaldson, D., & Storeygard, A. (2016). The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, *30*(4), 171–198.

Emerick, K. (2018). Agricultural productivity and the sectoral reallocation of labor in rural india. *Journal of Development Economics*, *135*, 488–503.

Erenstein, O., Jaleta, M., Sonder, K., Mottaleb, K., & Prasanna, B. M. (2022). Global maize production, consumption and trade: Trends and r&d implications. *Food security*, *14*(5), 1295–1319.

Farmaha, B. S., Lobell, D. B., Boone, K. E., Cassman, K. G., Yang, H. S., & Grassini, P. (2016). Contribution of persistent factors to yield gaps in high-yield irrigated maize. *Field crops research*, *186*, 124–132.

Fox, M. L., Thomas, M. A. H., & Haines, C. (2017). *Structural transformation in employment and productivity: What can africa hope for?* International Monetary Fund.

Ghebru, H., et al. (2018). Land policy and the youth bulge'in ethiopia: How social and economic transformations are scrutinizing the status quo.

Gollin, D. (2023). Agricultural productivity and structural transformation: Evidence and questions for african development. *Oxford Development Studies*, *51*(4), 375–396.

Gollin, D., Hansen, C. W., & Wingender, A. M. (2021). Two blades of grass: The impact of the green revolution. *Journal of Political Economy*, *129*(8), 2344–2384.

Gollin, D., Lagakos, D., & Waugh, M. E. (2014). The agricultural productivity gap. *The Quarterly Journal of Economics*, *129*(2), 939–993.

Gollin, D., Parente, S., & Rogerson, R. (2002). The role of agriculture in development. *American economic review*, *92*(2), 160–164.

Gollin, D., Parente, S. L., & Rogerson, R. (2007). The food problem and the evolution of international income levels. *Journal of Monetary Economics*, *54*(4), 1230–1255.

Gollin, D., & Rogerson, R. (2014). Productivity, transport costs and subsistence agriculture. *Journal of Development Economics*, *107*, 38–48.

Grosh, M. E., & Glewwe, P. (1998). Data watch: The world bank's living standards measurement study household surveys margaret e. grosh and paul glewwe. *Journal of Economic Perspectives*, *12*(1), 187–196.

Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment and development: A two-sector analysis. *The American economic review*, *60*(1), 126–142.

Herrendorf, B., Rogerson, R., & Valentinyi, A. (2013). Two perspectives on preferences and structural transformation. *American Economic Review*, *103*(7), 2752–2789.

Herrendorf, B., Rogerson, R., & Valentinyi, A. (2014). Growth and structural transformation. *Handbook of economic growth*, *2*, 855–941.

Herrendorf, B., Rogerson, R., & Valentinyi, A. (2022). *New evidence on sectoral labor productivity: Implications for industrialization and development* (tech. rep.). National Bureau of Economic Research.

Hirvonen, K. (2016). Temperature changes, household consumption, and internal migration: Evidence from tanzania. *American Journal of Agricultural Economics*, *98*(4), 1230–1249.

Jayne, T. S., & Sanchez, P. A. (2021). Agricultural productivity must improve in sub-saharan africa. *Science*, *372*(6546), 1045–1047.

Jessoe, K., Manning, D. T., & Taylor, J. E. (2018). Climate change and labour allocation in rural mexico: Evidence from annual fluctuations in weather. *The Economic Journal*, *128*(608), 230–261.

Karlan, D., Osei, R., Osei-Akoto, I., & Udry, C. (2014). Agricultural decisions after relaxing credit and risk constraints. *The Quarterly Journal of Economics*, *129*(2), 597–652.

Kennan, J., & Walker, J. R. (2011). The effect of expected income on individual migration decisions. *Econometrica*, *79*(1), 211–251.

Kongsamut, P., Rebelo, S., & Xie, D. (2001). Beyond balanced growth. *The Review of Economic Studies*, *68*(4), 869–882.

Kosec, K., Ghebru, H., Holtemeyer, B., Mueller, V., & Schmidt, E. (2018). The effect of land access on youth employment and migration decisions: Evidence from rural ethiopia. *American Journal of Agricultural Economics*, *100*(3), 931–954.

Lagakos, D., Marshall, S., Mobarak, A. M., Vernot, C., & Waugh, M. E. (2020). Migration costs and observational returns to migration in the developing world. *Journal of Monetary Economics*, *113*, 138–154.

Lagakos, D., & Shu, M. (2023). The role of micro data in understanding structural transformation. *Oxford Development Studies*, *51*(4), 436–454.

Lewis, W. A., et al. (1954). Economic development with unlimited supplies of labour.

Lobell, D. B., Thau, D., Seifert, C., Engle, E., & Little, B. (2015). A scalable satellite-based crop yield mapper. *Remote Sensing of Environment*, *164*, 324–333.

Matsuyama, K. (2008). Structural change. In *The new palgrave dictionary of economics* (pp. 1–6). Springer.

McArthur, J. W., & Sachs, J. D. (2019). Agriculture, aid, and economic growth in africa. *The World Bank Economic Review*, *33*(1), 1–20.

McCullough, E. B. (2017). Labor productivity and employment gaps in sub-saharan africa. *Food policy*, *67*, 133–152.

McMillan, M., & Headey, D. (2014). Introduction–understanding structural transformation in africa.

McMillan, M., Rodrik, D., & Sepulveda, C. (2017). *Structural change, fundamentals and growth: A framework and case studies* (tech. rep.). National Bureau of Economic Research.

Mensah, E. B., Owusu, S., Foster-McGregor, N., & Szirmai, A. (2018). Structural change, productivity growth and labor market turbulence in africa. *UNU-MERIT Working Paper Series*, *25*, 1–150.

Morando, B. (2023). Subsistence farming and factor misallocation: Evidence from ugandan agriculture. *The World Bank Economic Review*, *37*(4), 570–598.

Morsy, H., Shimeles, A., & Nabassaga, T. (2023). Structural change and inequality in africa. *Journal of African Economies*, *32*(Supplement_2), ii228–ii245.

Morten, M. (2019). Temporary migration and endogenous risk sharing in village india. *Journal of Political Economy*, *127*(1), 1–46.

Mukasa, A. N., Simpasa, A. M., Salami, A. O., et al. (2017). Credit constraints and farm productivity: Micro-level evidence from smallholder farmers in ethiopia. *African Development Bank*, *247*, 1–40.

Munshi, K., & Rosenzweig, M. (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review*, *106*(01), 46–98.

Nawrotzki, R. J., DeWaard, J., Bakhtsiyarava, M., & Ha, J. T. (2017). Climate shocks and rural-urban migration in mexico: Exploring nonlinearities and thresholds. *Climatic change*, *140*(2), 243–258.

Pan, X., & Sun, C. (2024). Internal migration, remittances and economic development. *Journal of International Economics*, *147*, 103845.

Porzio, T., Rossi, F., & Santangelo, G. (2022). The human side of structural transformation. *American Economic Review*, *112*(8), 2774–2814.

Rajaonarison, H. M. (2014). Food and human security in sub-saharan africa. *Procedia Environmental Sciences*, *20*, 377–385.

Schultz, T. W. (1953). The economic organization of agriculture.

Schultz, T. W. (1964). Transforming traditional agriculture.

Suri, T., & Udry, C. (2022). Agricultural technology in africa. *Journal of Economic Perspectives*, *36*(1), 33–56.

Syrquin, M. (1988). Patterns of structural change. *Handbook of development economics*, *1*, 203–273.

Tang, B., Liu, Y., & Matteson, D. S. (2022). Predicting poverty with vegetation index. *Applied Economic Perspectives and Policy*, *44*(2), 930–945.

Tsuda, S. (2022). Refugee inflows, surplus farm labor, and crop marketization in rural africa. *Journal of Development Economics*, *155*, 102805.

Vollrath, D. (2014). The efficiency of human capital allocations in developing countries. *Journal of Development Economics*, *108*, 106–118.

Zewdu, G. A., & Malek, M. (2010). Implications of land policies for rural-urban linkages and rural transformation in ethiopia.