Deforestation and Structural Transformation in Sub-Saharan Africa

Kenneth Houngbedji Clément Imbert Samuel M. Marshall
Julien Wolfersberger Liam Wren-Lewis *

Thursday 10th September, 2026

Abstract

By pulling workers out of subsistence agriculture, can structural transformation reduce deforestation in Africa? To answer this question, we combine population censuses with satellite data on deforestation from 1,324 administrative units in 28 Sub-Saharan countries. A shift-share design suggests that non-agricultural employment growth in one location decreases deforestation in locations from which it hires migrants ("migration-close") and increases it in locations from which it likely buys food ("trade-close"). To quantify these effects, we build a three-sector spatial equilibrium model with costly trade and migration and estimate it in 2000 and 2010. We find that manufacturing and mining productivity growth was the main driver of structural change over the period. Overall, it increased deforestation by 4%, but with wide heterogeneity across countries (from –18% to +33%). Model counterfactuals highlight the importance of where this productivity growth happens: growth in the locations most trade-close to forests brings the same amount of structural change as growth in the most migration-close, but the effect on deforestation goes from positive to negative.

JEL Codes: Q23, Q56, O13.

*Houngbedji: IRD Université Paris-Dauphine. Imbert: Sciences Po clement.imbert@sciencespo.fr. Marshall: University of Oxford. Wolfersberger: AgroParisTech and Paris-Saclay University. Wren-Lewis: Paris School of Economics. We would like to thank numerous conference and seminar participants, as well as Kirill Borusyak, Bruno Conte, Nicolas Coeurdacier, Farid Farrokhi, Antoine Leblois, Thierry Mayer, Frederik Noack, Florian Oswald, and Anant Sudarshan for helpful comments. Edoardo Badii, Robin Benabid Jégaden, Iddo Glass, Flavia Grasso, and Alonso Venegas provided excellent research assistance. We gratefully acknowledge funding from the Agence Nationale de la Recherche (ANR-24-CE26-2254), INRAE EXPLOR'AE (ANR-24-RRII-0003), and the STEG research initiative.

1 Introduction

Is economic development always bad for the environment? The idea of a trade-off between growth and environmental preservation is strengthened by large-scale deforestation in countries such as Brazil and Indonesia, which has been driven by export-led agriculture (Balboni et al., 2023). But in Africa, the continent responsible for a third of global forest loss between 2000 and 2020,[1] subsistence agriculture is the dominant driver of deforestation (Curtis et al., 2018) and penetration of international trade is more limited (Atkin and Donaldson, 2015). This opens the possibility that economic development, by pulling workers out of subsistence agriculture, may slow down deforestation.[2]

In this paper, we study how structural transformation has affected deforestation in Sub-Saharan Africa since 2000. Our key contribution is to highlight the role of spatial linkages *within* African countries. Based on a combination of census and satellite data, we show that structural change in one region is not correlated with deforestation in the same region. However, it slows down deforestation in regions from which it usually draws migrants ("migration-close"), but increases deforestation in regions from which it would likely buy food ("trade-close"). We build and estimate a multi-sector spatial general equilibrium model where regions are linked through costly migration and agricultural trade. Model counterfactuals suggest that manufacturing and mining productivity growth is the main driver of structural change; its average effect on deforestation is small (4%), but it ranges from –18% to +33% across countries. We link this heterogeneity to the internal geography of structural change relative to forests. In a model counterfactual that shifts manufacturing and mining productivity growth across space, we show that if it doubles in the top 5% regions that are the most "trade-close", deforestation increases by 1%, whereas if it doubles in the top 5% regions that are the most "migration-close", deforestation declines by 2%.

Specifically, we begin by combining satellite data on deforestation (Hansen et al., 2013) with census data on employment and historical migration across sub-national administrative units from 28 Sub-Saharan African countries in the 2000s and 2010s. Our sample represents 44% of forest cover in Sub-Saharan Africa in 2000, and 60% of deforestation between 2001 and 2020. In the first part of the paper, we present motivating evidence using a reduced sample of 18 countries for which we have two consecutive censuses. In these countries, the share of non-agricultural employment tends to be lower close to forests, and there is no correlation between deforestation and structural change in the same location. We investigate spatial linkages between deforestation and structural change using two shift shares: a "migration shift-share" that combines non-agricultural employment growth in all other locations as a shift with historical migration probabilities between locations as shares, and a "trade shift-share" that combines the same shift with market access (population at destination divided by

[1] As a comparison, South America accounted for 38% of global deforestation between 2000 and 2020, and Asia 22%. Forest loss in Africa amounts to about 4.4 million hectares – the area of Switzerland – per year (FAO, 2020).

[2] Many African countries are undergoing fast structural change, e.g. the agricultural share in the labor force fell from 55% to 30% in Ghana, and from 83% to 65% in Tanzania between 2000 and 2020 (ILO, 2021).

distance) as a share. We find that the migration shift-share is negatively correlated with deforestation, while the trade shift-share is positively correlated with deforestation. These findings suggest that growth in the non-agricultural sector in one location might not affect deforestation locally, but that it might decrease deforestation in other locations connected to it through migration ("migration-close") and increase deforestation in other locations connected to it via trade ("trade-close").

We further provide six additional pieces of evidence to support our interpretation of the shift-share results. First, we show that the estimates do not change when we control for local growth in non-agricultural employment, climatic shocks, or conflict, which could be correlated with non-agricultural growth over space. Second, we implement a placebo check and show that there is no correlation between the shift-shares and deforestation prior to the first census. Third, we find similar patterns if we use other measures of deforestation (Vancutsem et al., 2021) or if we consider changes in agricultural land use instead (Potapov et al., 2022): they both increase with the trade shift-share and decrease with the migration shift-share. Fourth, we test the sensitivity of our results to the choice of trade share (duration instead of distance) and to the inclusion of international migration and trade links between sample countries. Fifth, we check that our results hold in the shock-level specification advocated by Borusyak et al. (2022). Finally, we split the sample between countries with a higher vs lower ratio of agricultural trade (exports plus imports) over agricultural GDP. If countries import their food, or mostly produce for exports, then structural change may move consumption of agricultural goods across space without affecting land use, and hence deforestation. Indeed, we find that the trade shift-share has smaller effects on deforestation in countries more open to international trade.

These shift-share results establish the existence of opposing forces in the relationship between structural change and deforestation, but cannot quantify them in equilibrium. We develop a spatial general equilibrium model with three sectors (agriculture, manufacturing and mining, and services). Locations are connected through costly migration and costly trade in agricultural goods. The model incorporates non-homothetic demand, heterogeneous agricultural productivity across crops and locations, and the endogenous conversion of forest plots to farmland. It thus provides a quantitative framework in which structural change leads to the joint reallocation of labor and demand for agricultural goods across space, with corresponding effects on land use and deforestation.

We then proceed to quantify the model for our 28 sample economies in 2000 and 2010 and recover productivity changes by location and sector. We show that the estimated productivity changes account for most of the structural transformation observed over this period, with manufacturing and mining productivity changes playing the largest role. Overall, we find that productivity growth substantially reduced predicted deforestation: absent these changes, deforestation would have been 28% higher than the observed rate. We find that agricultural productivity growth had a large reducing effect, consistent with a land-saving effect in the spirit of the Borlaug Hypothesis: holding agricultural productivity at its 2000 level raises predicted deforestation by 25%. By comparison, the average effect of manufacturing and mining productivity growth is small: holding it at its 2000 level decreases predicted deforestation

by 4%. This modest average number, however, hides large positive and negative effects across space.

The contrast between two countries, Angola and Ghana, illustrates this heterogeneity. According to the model's decomposition, manufacturing and mining productivity changes generated almost the same amount of structural transformation in the two countries: the non-agricultural employment share rose by 10 p.p. in Angola, and 9.6 p.p. in Ghana. Yet these changes increased deforestation by 1.7 p.p. in Angola and reduced it by 0.65 p.p. in Ghana. Similar amounts of structural transformation can therefore produce forest outcomes of opposite signs. The geography of productivity growth provides a first clue: in Angola, it was concentrated in the densely forested northwest, whereas in Ghana it was strongest in the northern half of the country, where forest cover is low.

The previous shift-share analysis suggested that geographical proximity between structural change and forests can be decomposed into two, via migration- and trade-closeness. We apply the same shift-share design using the same shares, and using as shifts changes in non-agricultural employment between the observed equilibrium and the counterfactual equilibrium without any sectoral productivity growth. We show that in our model counterfactuals as in the data there is less deforestation in locations that are migration-close to a growing non-agricultural center and more deforestation in locations that are trade-close.

To pin down the mechanisms behind these patterns, we next conduct location-specific counterfactuals. In each counterfactual, we set manufacturing and mining productivity in one spatial unit back to its 2000 value while keeping all remaining parameters at their 2010 values, thereby isolating the contribution of productivity growth in that location to national structural transformation and deforestation. These exercises show that, even within the same country, one single productivity change can increase or reduce national deforestation depending on where it occurs. The position of the growing location within migration and trade networks determines whether forested regions are affected primarily through the workers they send or through the agricultural goods they supply. Accordingly, productivity growth generates less forest loss in migration-close regions, and more in trade-close ones. The impact of structural transformation on deforestation therefore depends not only on the magnitude of productivity growth or the number of workers leaving agriculture, but also on the position of the growing location within the country's migration and trade networks.

Finally, we demonstrate the importance of internal geography by comparing two counterfactuals that double manufacturing and mining productivity in two different sets of locations. In each country, we rank locations by the difference between forest cover in the regions to which they are connected through migration and through trade, and double non-agricultural productivity in either the top or the bottom 5% of this ranking. While both counterfactuals increase the non-agricultural employment share by approximately 1.3%, their effects on forests are markedly different. Increasing productivity in locations connected to forests relatively more through migration *reduces* deforestation by almost 2%, whereas increasing it in locations connected relatively more through trade *increases* deforestation by 1%. A given amount of structural transformation thus does not have a fixed environmental cost.

Similar development outcomes can place very different pressures on natural land because the location of growth determines how it is transmitted to agricultural regions. This has direct implications for policies that seek to promote productivity growth in low-income countries: where growth occurs can be as important for forest loss as the structural transformation it generates.

Contribution Our work contributes to the literature on deforestation in tropical countries by examining the role of structural transformation through migration and trade. Early literature on the economics of tropical deforestation uses cross-country regressions to investigate economic and institutional factors behind deforestation (see Barbier and Burgess, 2001, for a review). The recent literature exploits advances in the measurement of deforestation using satellite data and new empirical strategies to get at the causal effect of determinants of deforestation (see Balboni et al., 2023, for a recent review).[3] The literature has highlighted the importance of international trade (Ferreira, 2004; Leblois et al., 2017; Abman and Lundberg, 2020; Hoang and Kanemoto, 2021) and market access through changes in road infrastructures (Chomitz and Gray, 1996; Pfaff, 1999; Andersen, 2002; Pfaff et al., 2007; Asher et al., 2020; Kaczan, 2020). Another strand of the literature has addressed the effects of population forces (Angelsen, 1999) and demand for fuelwood and wood products induced by rising incomes (Foster and Rosenzweig, 2003) on deforestation. Fewer papers look at the deforestation impacts of migration to urban centers on the sending rural locations. Two exceptions are Angelsen et al. (2020), who find mixed impacts of international migration on deforestation in sending communities in case studies in Guatemala, and Brewer et al. (2022), who find that labor loss and the inflow of remittances from rural urban migration translates into reduced tree cover loss in Uganda. Our motivational evidence uses a similar methodological framework to Brewer et al. (2022) and spans several countries, including Uganda: we replicate their finding that diversification through rural-urban migration has a net positive effect on tree cover.[4] Our contribution to this literature is to investigate migration and trade together as determinants of deforestation in the process of structural transformation, and to embed them in a spatial equilibrium model which allows us to quantify their net effect in equilibrium.

Our paper also contributes to the literature on internal trade and migration in developing countries, which increasingly employs new quantitative spatial models (Redding and Rossi-Hansberg, 2017). While many papers have developed spatial equilibrium frameworks to study agriculture (e.g., Sotelo, 2020; Pellegrina, 2022), or aggregate productivity (Tombe, 2015; Bryan and Morten, 2019), their focus is on economic rather than environmental outcomes. Our paper is part of a literature which uses

[3] A burgeoning literature investigates institutional and political factors, such as property rights (Fetzer and Marden, 2017; Wren-Lewis et al., 2020), political incentives (Burgess et al., 2012; Balboni et al., 2021), or anti-poverty programs (Alix-Garcia et al., 2013; Ferraro and Simorangkir, 2020), and evaluates the effectiveness of payments for ecosystem services (Jack and Jayachandran, 2019; Souza-Rodrigues, 2019).

[4] In related work, Madhok et al. (2022) analyze the impact of rural-urban migration on agriculture in rural areas in India. They use a shift-share instrument which is similar to the one we use to identify trade effects, but they do not use data on migration patterns and do not consider deforestation as an outcome.

spatial models of trade and/or migration to study deforestation (e.g. Araujo et al., 2023; Restrepo and Leite-Mariante, 2023; Gollin and Wolfersberger, 2025; Dominguez-Iino, 2026).[5] These papers tend to emphasize the links between deforestation and international trade, which is most relevant in the South American context. In contrast, we study Sub-Saharan Africa and focus on structural transformation.[6] In a recent paper, Farrokhi et al. (2023) develop a spatial dynamic framework to study global deforestation. They study the effect of international trade on deforestation, including population growth and structural change. They show that decreasing agricultural trade costs can reduce deforestation if cheaper agricultural products increase demand for non-agricultural goods, which generates structural change. The focus of their analysis is on international trade rather than internal trade and migration, but their takeaway resonates with ours: structural change in developing countries can help curb deforestation.[7]

Finally, our work shows that the effects of structural transformation on natural resources are spatial general equilibrium outcomes. The reallocation of labor out of agriculture and the increase in demand for agricultural goods are connected consequences of productivity-driven development, but have opposite implications for land use. In this respect, we connect the literature on productivity-driven labor reallocation and non-homothetic demand (Matsuyama, 2009; Ngai and Pissarides, 2007; Herrendorf et al., 2014; Kongsamut et al., 2001; Boppart, 2014; Comin et al., 2021) to the study of deforestation. We also provide evidence of the environmental consequences of structural transformation, one key aspect of growth and development that has received little attention so far (Herrendorf et al., 2014).

The remainder of the paper is structured as follows. Section 2 describes the data. Section 3 presents the reduced-form evidence. Section 4 develops the spatial model, which we quantify in Section 5. Section 6 presents the counterfactual exercises, and Section 7 concludes.

2 Data

Our study combines data on deforestation, agriculture, migration, and employment. We use satellite data, which has global coverage, for deforestation and agriculture; however, migration data is not available for every country in sub-Saharan Africa (SSA), which limits our geographical coverage. In this section, we first describe our sample, before presenting each data source in turn.

[5] A related strand of the environmental economics literature use spatial models to quantify the effects of climate change on productivity with changes in trade and/or migration (Desmet and Rossi-Hansberg, 2015; Costinot et al., 2016; Burzynski et al., 2022; Conte, 2022). See Balboni and Shapiro (2025) for a recent review.

[6] Several papers have estimated the effect of structural transformation on pollution (Grossman and Krueger, 1995; Herrendorf et al., 2014).

[7] In a non-developing country context, Coeurdacier et al. (2026) build a multi-sector spatial general equilibrium model similar to ours to study land use and carbon balance in France since the 1950s. They emphasize two margins that are important in their context: change in land use through urban sprawl and forest plantations. We do not model these two aspects which are less relevant in the Sub-Saharan African context, where the carbon footprint of economic development is concentrated in tropical forests cleared for agricultural use (Curtis et al., 2018).

2.1 Sample

Our sample consists of 28 countries in SSA for which we were able to collect migration information in the 1990s, 2000s and 2010s. For 23 countries, we collect migration data from Integrated Public Use Microdata Series-International (IPUMS-I).[8] For each of these, we use the most disaggregated administrative unit available on IPUMS-I for which migration data is reported.[9] We supplement this data with census data for Angola and Madagascar, which we collect directly from the countries' statistical agencies, aggregate tables for two censuses in Gabon, and nationally representative Living Standards Measurement Study (LSMS) data for Ethiopia and Nigeria. Together, this yields 1324 administrative units in our main sample.

Figure 1 shows the countries in our sample, divided into their respective administrative units, superimposed on a map of tree cover in 2000 and deforestation between 2001 and 2020 from the Hansen et al. (2013) data (described below). Our sample includes West-African, East-African, and Southern African countries. Together, our sample covers 43.8% of total forest cover in SSA in 2000, and 59.7% of deforestation between 2001 and 2020, meaning that the active frontier of deforestation is over-represented in our sample. Notably, much of the Congo Basin is absent from our sample, which bears on the external validity of our results to all of SSA. Selection into our sample is not random: it requires a country to have conducted a census or a representative household survey, deposited harmonized microdata, and asked about migration history. These conditions are correlated with stronger institutions, which may affect the rate of structural transformation and forest management. The countries in the Congo Basin that are not in our sample have lower population density and more sparse road networks. We expect the trade channel in particular, which works through market access, to be weaker there, and we caution against extrapolating our counterfactuals to those countries.

2.2 Satellite Data on Land Use

Deforestation We construct our measure of forest loss using the Hansen et al. (2013) global forest cover change dataset. The data divides the world into $30m \times 30m$ pixels and reports, for each pixel, the share of area covered by tree canopy in 2000, as well as the year in which the tree cover loss occurred, defined as a clearing of at least half of the tree cover within a pixel. Once a pixel is classified as having experienced loss, it is masked from further updates in the loss year, so subsequent dynamics such as regrowth are not recorded. To help disentangle tree cover in forested regions from tree plantations for commercial purposes, we consider that pixels with tree canopy cover of 30% or more are likely to

[8] Of the 49 *present-day* countries in sub-Saharan Africa, 28 have census data on IPUMS-I, and 24 collect migration information. For the main analysis we focus on the period 2000–2010. Because South Sudan separated from Sudan in 2011, we treat it as a single country, which reduces the number of countries in our dataset from IPUMS-I by one.

[9] For 17 of the countries in our sample we use the second administrative unit (more disaggregated), and for the remaining 11, we use the first administrative unit. We report the administrative level and number of units in each country in Appendix Table A.1.

be forested. Then, we measure the area of tree cover lost across forested pixels located within each administrative unit each year and define the share of administrative unit o that was deforested annually as:

$$Deforestation_o = \sum_{t=T_1+1}^{T_2-1} \frac{forest\,loss_{ot}}{total\,area_o(T_2 - T_1 + 1)}$$

where T_1 and T_2 are the years of the first and second census, respectively, and $total\,area_o$ is the total area of the spatially harmonized administrative unit. Thus, $Deforestation_o$ measures the share of area in o that was deforested in the period between the two censuses.

Deforestation in Tropical Moist Forests While the Global Forest Change dataset produced by Hansen et al. (2013) has the advantage of being global, this inevitably leads to inaccuracies when compared to datasets focused on particular forest types. We therefore also use data on deforestation of Tropical Moist Forests produced by Vancutsem et al. (2021). While some forested areas in our data do not fall into this categorization, tropical moist forests represent the most important kind of forest in our countries of interest. The advantage of using this dataset is that this captures deforestation often missed by Hansen et al. (2013), in particular deforestation which occurs through a gradual degradation process (Vancutsem et al., 2021). Such a gradual process is likely to be particularly important when areas are deforested due to the gradual spread of small-scale agriculture.

Deforestation prior to 2000 Information on deforestation in the 1990s is not available in the Hansen et al. (2013) data. To test for differential pre-trends, we use land class data from ESA CCI (Copernicus Climate Change Service, 2019), and define forested areas as those with tree-cover classifications.[10] We construct two measures of forest loss, one that only uses ESA CCI, and another that combines it with Hansen data from 2001 onward.

Crop Land To document the extent to which forest cover change is driven by changes in agricultural activities, we estimated the change in cropland area for each administrative unit over time. The estimates were produced using a global consistent map of $30m$ spatial resolution that identifies the pixels with cropland over 4-year periods: 2000-2003, 2004-2007, 2008-2011, 2012-2015, and 2016-2019. Pixels are classified as cropland if a growing crop was detected during any of year of the 4 year epoch (see Potapov et al., 2022). Use of 4-year long interval increases the number of clear-sky satellite observations and hence the accuracy of cropland detection. Areas that are not used to grow crops for human consumption, forage, or bio-fuel for more than 4 years are not considered as cropland. When the years do not align exactly with the census year, we use the year that is closest to each census

[10]The tree cover classifications that we use are broadleaved evergreen, broadleaved deciduous, needleleaved evergreen, needleleaved deciduous, and mixed leaf type, with 15% cover or higher.

and define the change in the log share of area that is cropped ($\Delta \log S^A$) as:

$$\Delta \log S_o^A = \log\left(\frac{area\,cropped_{o,t_2}}{total\,area_o}\right) - \log\left(\frac{area\,cropped_{o,t_1}}{total\,area_o}\right)$$

where t_1 and t_2 are the boundaries of the epochs closest to the census years, T_1 and T_2. Cropped area is measured in hectares; to avoid losing units with no area cropped in one round, we replace zeros with 100 m^2 (0.01 hectares). This leaves areas with no cropped land in either period unchanged.

Agricultural Output and Potential Yields We collect data on agricultural output and potential yields from GAEZ (Fischer et al., 2021). For potential yields, we use the values under under low-inputs and rain-fed irrigation. Following Adamopoulos and Restuccia (2022), we compute the potential agricultural yield across 22 major crops, valuing the yield with international prices from 2000, and weighting each crop by its share of local land use in 2000.

Nighttime Lights We collect data on the value of output in manufacturing and services from national GDP. We allocate total output across sub-national units using nighttime lights. When there are minimal nighttime lights in a location, we estimate output using the share of workers in manufacturing and mining and services (Elvidge et al., 2022).

2.3 Census Data on Employment and Migration

Migration We use census data, where available, and LSMS data to measure migration flows between locations o (origin) and d (destination). We define the set of migrants from o in d as the set of working aged (15-64) individuals who are located in d at the time of the census and report that their previous location was o. This definition is intended to exclude family moves (under 15) and return migration (65 and over). There are several difficulties with creating a consistent measure of migration flows across countries which we address. First, individual migration history is asked differently across countries: some ask for the year of migration, some ask where the respondent lived X years ago, where X varies from one year to ten years, and some ask where the respondent was born. When we know the year of migration, we include all individuals who migrated since the previous census, except those who migrated in the past year to limit over-counting seasonal migrants.[11] We construct the annual rate of migration as: $m_{od} = M_{od}/X$, where M_{od} are measured migration flows and X is the migration reporting window length, which we set to 15 yeqrs for migration since birth. From this, we construct the annual emigration rate from o in country C as: $e_o = \sum_{d \in C \setminus o} m_{od}/N_o^{pa}$, where N_o^{pa} is the working-aged population in o. Second, the geo-political boundaries of administrative units are

[11] See Appendix Section C.1 for details on how migration is measured in each country.

changing over time. We use spatially harmonized geographical units from IPUMS-I as our unit of analysis. While this reduces the number of observations, it prevents us from mislabeling as migrants people who changed locations because a border was redrawn. Third, for international migrants, we observe the country of origin, but not the location within that country. For the main analysis, we exclude international migrants, and, as a robustness check, we impute their location of origin within the sending country.[12]

Employment We use data from the census, where available, to calculate the employment shares. We include all persons who are employed (including self-employment), and define three sectors of work: agriculture, manufacturing and mining, and services. Our preferred measure is the industry of employment. When this is not available we classify occupation codes into each category. We use this data to compute non-agricultural employment in each census year N_{dt}^{non-ag} and the change between two censuses $\Delta N_{dt}^{nonag} = N_{dT_2}^{nonag} - N_{dT_1}^{nonag}$.

In the model, we calculate the population in each location in each period using data from the gridded population of the world (CIESIN, 2018). We then allocate labor into each sector using the employment shares from the nearest census.

3 Motivating Evidence

In this section we use data from 18 countries with two censuses to provide evidence on the effect of structural change on deforestation. We first provide correlations between deforestation and structural change in the same location, before investigating the effect of structural change in one location on deforestation in other locations via two channels, migration and trade.

3.1 Deforestation and Local Structural Change

We begin by comparing the geography of non-agricultural employment (including manufacturing, mining and services) and forest cover within the countries of our sample, by running OLS regressions with country fixed effects. In Table 1 Column 1, we show that there is a strong negative correlation between the share of land area forested and the share of non-agricultural employment at the administrative unit level. This reflects the fact that most non-agricultural employment is concentrated in urban and industrial clusters located far from the forests. In Column 2, we find that non-agricultural employment growth is not correlated with baseline forest cover. This suggests that despite the small geographical overlap between non-agricultural activities and forests some forested places do experience fast

[12]For the robustness exercise we impute the origin of migrants from country A to country B probabilistically, based on the fraction of migrants from country B in each location of country A. This imputation is based on two empirical observations: migration is often circular and follows ethnic proximity across international borders.

non-agricultural employment growth, e.g. because of mining activity (Morton et al., 2026).

In Column 3, we find no correlation between the rate of deforestation and the share of non-agricultural employment. This result resonates with Curtis et al. (2018)'s argument that deforestation in Sub-Saharan Africa is primarily linked with subsistence agriculture. In Column 4, we find no correlation either between non-agricultural employment growth and deforestation. This evidence is only correlational, and there is more causal evidence that growing incomes (Ferraro and Simorangkir, 2020) or mining booms (Morton et al., 2026) can lead to local deforestation. But taken together results in Table 1 do suggest that the main effects of structural change on deforestation may not be local. Indeed, in our sample, more than 50 % of the total growth in non-agricultural employment takes place in areas that together contain less than 10% of the total forested area. It is therefore important to focus on how growth in non-agricultural employment in one location might affect deforestation in other locations, through spatial links such as migration or trade.

3.2 Deforestation and Structural Change in Other Locations

Migration Shift-Share To capture the effect of structural change via the migration channel, we use a shift-share based on previous migration patterns, similar to the "Card" instrument in the migration literature (Card, 2001; Jaeger et al., 2018; Imbert et al., 2022). Let o and $d \in C$ be the generic indices for origin and destination administrative units in country C. We define the migration shift-share instrument as

$$\text{Migration S-S} = \sum_{d \in C \setminus o} \frac{m_{od,t-1}}{total\ area_o} \frac{\Delta N_{dt}^{nonag}}{N_{d,t-1}^{nonag}(T_2 - T_1)} \tag{1}$$

Here we can see that the 'shifts' are the annualized growth rates in non-agricultural employment, i.e. ΔN_d^{nonag} is the change in non-agricultural employment at destination between the first and second census, which is divided by the baseline employment $N_{d,t-1}^{nonag}$ and the time between the two censuses $T_2 - T_1$. The 'shares' composed of the total number of migrants out of origin o into destination d, $m_{od,t-1}$, divided by the origin land area $total\ area_o$. The shares are normalized by the land area to estimate the change in population density, which is the relevant variable when thinking about change in land use shares. Because the shares do not sum to one, we include as a control the sum of the migration shares W_o^m:

$$W_o^m = \sum_{d \in C \setminus o} \frac{m_{od,t-1}}{total\ area_o}$$

Trade Shift-Share To capture the effect of structural change via the trade channel, we use another shift-share based on inverse distance, similar to the "market access" measures in trade (Donaldson and Hornbeck, 2016). We define a similar set of shift-share instruments using inverse distance weights.

These instruments capture the effects of trade. We define the distance shift-share using non-agricultural employment as shifts:

$$\text{Trade S-S} = \sum_{d \in C \setminus o} \frac{1}{d_{od}} \frac{N_{d,t-1}}{total\ area_o} \frac{\Delta N_{dt}^{nonag}}{N_{d,t-1}^{nonag}(T_2 - T_1)} \tag{2}$$

Where (d_{od}) is the travel distance in km between the centroids of o and d based on 2000 Michelin road maps from Müller-Crepon et al. (2021). The shares are made up of the inverse of this distance multiplied by the destination population (to approximate 'market access') and, as with the migration share, divided by the land area at origin. In this way, we estimate the intensity of demand from trade on a given area - this reflects the fact that a similar demand for agricultural goods will have a much larger effect on land use on a small region than on a large one. To control for the fact that the shares do not sum to one, we additionally define the sum of the distance weights W_o^d as

$$W_o^d = \sum_{d \in C \setminus o} \frac{1}{d_{od}} \frac{N_{d,t-1}}{total\ area_o}$$

As a robustness check, we also compute a trade shift-share where the distance d_{od} is computed as the travel duration (in minutes) between o and d.

Estimating Equation Our main estimating equation is the following:

$$y_o = \beta_m \textit{Migration S-S}_o + \beta_t \textit{Trade S-S}_o + \mathbf{X_o} + \varepsilon_o \tag{3}$$

where y_o is the outcome in administrative unit o (deforestation or change in land use). We include in the estimation only origin administrative units with positive forest cover at baseline (779 out of 831). *Migration S-S*$_o$ is the migration shift-share, *Trade S-S*$_o$ the trade shift-share, $\mathbf{X_o}$ is a vector of controls which includes the sum of migration shares W_o^m, and the sum of trade shares W_o^d. These controls are important for identification, since origins which are better connected to the rest of the country in terms of migration (high W_o^m) or in terms of trade (high W_o^d), may be on different trends. We also include country fixed effects in $\mathbf{X_o}$ to account for national economic and policy shocks and use robust standard errors.

Identification In order for β_m to capture migration linkages and β_t to capture trade linkages, it is crucial to include both shift-shares simultaneously. Most papers in the economics literature use either migration shift-shares (Card, 2001) or trade shift-shares (Donaldson and Hornbeck, 2016) but not both. Since migration patterns follow gravity, they are correlated with distance and population at

destination.[13] Our claim is that, conditional on the migration shift-share, the trade shift-share will capture the effect of trade. The correlation between migration and trade shares at the origin-destination pair level in our context is 0.3. The correlation between the two shift-shares at the origin level is 0.5.

Also crucial for identification is the assumption that the destination shifts in the shift-shares are exogenous to deforestation trends at origin (Borusyak et al., 2022). We believe that this assumption is reasonable, since the shift we use is the change in non-agricultural employment in another part of the country. We check the robustness of our results to controlling for the local shift, i.e. the change in non-agricultural employment locally, for weather and geographical variables (rainfall, temperature and ruggedness) as well as conflict. We also conduct a placebo test, in which we regress deforestation prior to the first census on the shift-shares.

3.3 Shift-Share Results

This section presents empirical results on the effects of structural change on deforestation through migration and trade channels.

Emigration We first use specification (3) and regress emigration on the migration and trade shift shares. The dependent variable is total emigration out of the origin region divided by the land area - in other words, the reduction in population density due to emigration. Columns 1-3 in Table 2 present the results. Reassuringly, the coefficient on the migration shift-share is positive and highly significant. As Column 2 shows, on its own the trade shift-share is also positively correlated with emigration. This is expected, since migration also follows gravity forces. When we put both shift-shares in the regression (Column 3), however, the coefficient on the migration shift-share becomes bigger, and the coefficient on the trade shift-share flips sign and becomes negative. We interpret these coefficients as reflecting the opposing effect of the migration and trade channels: an increase in structural change in one location will draw more migrants from other locations, but also increase demand for agricultural products from these locations, which reduces incentives to migrate.

Deforestation We then regress the share of area deforested from Hansen et al. (2013) on the two shift-shares (Columns 4-6 in Table 2). Column 4 suggests that there is less deforestation in regions that are migration-close to growing non-agricultural centers. In contrast, there is more deforestation in regions that are trade-close to growing non-agricultural centers (Column 5). When we put the the two shift-shares together in the same specification, the migration shift-share has a larger negative coefficient, and the trade shift-share a larger positive one (Column 6). These results suggest that structural change in one location decreases deforestation in areas with which it has migration links,

[13] Sometimes distance-based shares are even used to capture the effect of migration when reliable information on migration shares is not available (Madhok et al., 2022)

but decreases it in areas to which it is closer in a market access sense, i.e. from which it is likely to import food.

Controls One potential issue with our shift-share specification is the spatial correlation of non-agricultural employment growth, so that we may confound the effect of the destination-level shocks with that of correlated origin-level ones. To alleviate this concern, we control for local shocks to non-agricultural employment. We also control for other geographical characteristics, weather shocks, and conflict which may predict migration, trade and deforestation: the results are virtually unchanged (Appendix Table A.4).

Placebo We also follow Borusyak et al. (2022)'s recommendation and test for differential pre-trends by regressing deforestation prior to the first census on the two shift-shares that are based on non-agricultural employment growth between the two censuses. Table 3 column 1 first tests the correlation between non-agricultural employment growth and prior deforestation in the same administrative unit. Then columns 2 to 4 show that there is no correlation between shift-shares and prior deforestation, which suggests that our main estimates capture the effect of contemporaneous structural change rather than pre-existing trends. [14]

Different measures of land use Our analysis relies on one measure of deforestation based on satellite data from Hansen et al. (2013). In Appendix Table A.6 Column 1 we check that we find similar results with an alternative measure of deforestation from Vancutsem et al. (2021), which is more accurate for tropical moist forests. Since our interpretation relies on deforestation driven by agricultural land, we also check that we find a corresponding decrease (migration channel) and increase (trade channel) in crop land based on Potapov et al. (2022)'s land use classification (Column 2 in Appendix Table A.6). Hansen et al. (2013) data offers multiple definition of deforestation based on different thresholds of forest cover. Our main specification uses 30%, but in Columns 3 to 6 in Table A.6 we present results based on alternative thresholds (10, 50, 70, and 90%). The coefficients on the migration and trade shift-shares are qualitatively similar for all thresholds, although they are largest for 50% and smaller for very low or very high thresholds of forest cover.

Alternative shift-shares We proxy trade proximity using shares based on population and inverse distance, but we check that we obtain similar results if we use inverse of travel duration instead (Appendix Table A.7 Columns 1 and 2). Our main analysis focuses on the effects of structural change in one location of a country on other locations in the same country via internal migration and trade. We

[14] Depending on the date of the first census, we can measure prior deforestation either with ESA CCI data only or combine it with Hansen et al. (2013) data from 2001 onward. Appendix Table A.5 presents the results with the second measure and finds similar results

can also build shift-shares that include shifts in other countries to include international migration and trade. The migration shares between cross-country pairs are imputed using information on the location of international migrants at destination (see Section 2) We show that our results are qualitatively and quantitatively similar (Columns 3 and 4 in Appendix Table A.7).

Inference The literature suggests that standard errors in shift-share designs are spatially correlated (Adão et al., 2019; Borusyak et al., 2022). We follow Borusyak et al. (2022) and present results inverted at the level of the shock, i.e. destination-level regressions. Since there is no guidance on how to do this with two shift-shares, we apply the procedure for one shift-share at a time. Appendix Table A.8 presents the results. The relationship between the trade shift-share and emigration becomes insignificant (Column 2). All other coefficients remain highly significant, including the positive relationship between the migration shift-share and emigration, the negative effect of structural change on deforestation via migration, and the positive effect via trade (Column 1, 3 and 4).

Heterogeneity based on trade openness One would expect the effects of structural change in one location on land use in other locations to be smaller if workers in growing non-agricultural centers can import food from abroad or if the agricultural sector mostly produces for the world market. To test this, we divide the sample between countries that are more or less open to international trade, based on the ratio of agricultural imports plus exports over agricultural GDP between 2000 and 2020 (World Bank, 2026). Appendix Table 4 shows that the effects of structural change on emigration and deforestation via the trade channel are very different depending on international trade openness. In more closed economies, structural change in one location means less emigration and more deforestation from trade-close locations which respond to its growing demand for food. In contrast, in more open economies the coefficient on the trade shift-share is positive for emigration and small and insignificant for deforestation. The coefficients on the migration shift-shares are similar in magnitude across the two sets of countries but insignificant in more open economies.

Overall, our shift-share results point to an ambiguous relationship between structural change and deforestation. We find evidence that structural change is not correlated with deforestation locally, but related to less deforestation in places that are likely to send migrants to growing non-agricultural centers, and to more deforestation that are likely to export food to these centers. Thanks to an extensive list of placebo and robustness checks, we think that these results do capture the effects of structural change and not of other correlated economic and geographical factors. But even if they do reflect the effect of structural change, our shift-share estimates are only suggestive and do not have a quantitative interpretation, because they do not capture equilibrium effects, i.e. effects in other locations that are connected with origin and destination through migration and trade. For this, the next sections will develop and quantify a spatial equilibrium model.

4 The Model

In this section, we develop a spatial general equilibrium model of structural transformation and deforestation with migration and trade.

4.1 Model Set-up

Environment. The economy is composed of a measure of individuals of a discrete set of domestic locations with generic index $i, j \in \mathcal{D}$. Each location is endowed with a fixed land area, $\bar{L}_i$, which captures the heterogeneity in size of administrative units. Some land is forested (L_i^F) and some land is cleared already (L_i^C). Land is composed of a continuum of plots $\omega \in \bar{L}_i$. The economy produces a set of crops indexed by $g \in \{1, \ldots, G\}$. Each plot is endowed with crop-specific agricultural productivity $Z_{ig}(\omega) \sim F(\theta, \Lambda_{ig})$ and a clearing cost $Z_i^0(\omega) \sim F(\theta, \Lambda_i^0)$, where $F(\cdot, \cdot)$ are mean-normalized Fréchet distributions. These assumptions imply that the distribution of clearing costs and crop-specific productivities vary across locations. Markets for goods also include a foreign location F, and we denote $\mathcal{L} = \mathcal{D} \cup \{F\}$ the set of locations that enter goods markets.

Technology. There are three sectors: agriculture (A), manufacturing (M), and services (S).[15] Agricultural production in each location i crop g combines land $L_{ig}(\omega) \in \{0, 1\}$ and labor N_{ig}. Output of crop g on plot ω is given by:

$$Q_{ig}(\omega) = Z_{ig}^A \left(N_{ig}(\omega)\right)^{\alpha_g} \left(Z_{ig}(\omega) L_{ig}(\omega)\right)^{1-\alpha_g}, \tag{4}$$

where $\alpha_g \in (0, 1)$ is the crop-specific labor share and Z_{ig}^A is crop-specific agricultural TFP. Each plot is allocated to the crop that maximizes profit, and forested plots are cleared only when the operating profit exceeds the cost $Z_i^0(\omega)$ of clearing the land for agriculture.

The non-agricultural sectors use labor only, via decreasing returns to scale technologies:

$$Q_i^k = Z_i^k (N_i^k)^{\alpha_k}, \quad k \in \{M, S\}, \tag{5}$$

where $\alpha_k \in (0, 1)$ and Z_i^k is TFP in sector k, location i.

Trade. Shipping agricultural goods between locations is subject to iceberg costs: for one unit of agricultural good from i to arrive in j, $\tau_{ij} \geq 1$ units must be shipped. We assume that costs are symmetric, $\tau_{ij} = \tau_{ji}$ and that $\tau_{ii} = 1$. Trade with the foreign location F is governed by crop-specific export and import wedges. Denoting τ_{ijg} the resulting iceberg cost: $\tau_{ijg} = \tau_{ij}$ for domestic bilateral flows, while flows involving F include a crop-specific foreign wedge, which can be different for

[15] The manufacturing sector includes all non-agricultural goods that can be traded (e.g., mining).

imports and exports: $\tau_{Fjg} = \tau^M_{Fg} \times \tau_{Fj}$ and $\tau_{iFg} = \tau^X_{Fg} \times \tau_{iF}$. Trade in manufacturing goods is not subject to trade costs, whereas services are consumed locally and are not traded.

Preferences. Individuals are endowed with a birth location i, and a vector of idiosyncratic preferences for location-sector pairs $\{\chi_{jk}\}$. Individuals choose a location $j \in \mathcal{D}$ and a sector $k \in \{A, M, S\}$ in which they supply one unit of labor inelastically to maximize their utility of consumption c, given sector-location amenities (b_{jk}) and migration costs $(\delta_{ij} \geq 1)$:

$$U_i = \max_{jk} \{\zeta_{jk} b_{jk} c_{jk} \delta_{ij}^{-1}\}. \tag{6}$$

As with trade costs, we assume that migration costs are symmetric, such that $\delta_{ij} = \delta_{ji}$, and that there is no cost to not migrating, $\delta_{ii} = 1$. The idiosyncratic taste shocks $\{\zeta_{jk}\}$ are distributed via a nested generalized extreme value distribution. In the upper nest, individuals choose between locations with elasticity κ, and in the lower nest they choose between sectors with elasticity γ. We assume that $b_{jA} = b_j$, and $b_{jk \neq A} = \beta^k_j b_j$, where β^k_j is the non-monetary penalty from working outside of agriculture in sector $k \neq A$ in location j.

Individuals have non-homothetic preferences over consumption of the agricultural and manufacturing goods and services. Following Comin et al. (2021), we define the utility of consumption for an individual in location i, working in sector k, (c_{ik}) via the implicit function:

$$(\Omega_A)^{\frac{1}{\sigma}} \left(\frac{c^A_{ik}}{c^{\varepsilon_A}_{ik}} \right)^{\frac{\sigma-1}{\sigma}} + (\Omega_M)^{\frac{1}{\sigma}} \left(\frac{c^M_{ik}}{c^{\varepsilon_M}_{ik}} \right)^{\frac{\sigma-1}{\sigma}} + (\Omega_S)^{\frac{1}{\sigma}} \left(\frac{c^S_{ik}}{c^{\varepsilon_S}_{ik}} \right)^{\frac{\sigma-1}{\sigma}} = 1, \tag{7}$$

where c^A_{ik}, c^M_{ik}, and c^S_{ik} are the composite agricultural good, the manufacturing good, and services, respectively, for an individual located in i and working in sector k. The parameter σ governs the elasticity of substitution across broad consumption categories. The parameters ε_k, $k \in \{A, M, S\}$, control the degree of non-homotheticity in preferences: when $\varepsilon_k = 1$, preferences are homothetic in good k, while deviations from unity imply that expenditure shares on that good vary with real income. Finally, Ω_k is a good-specific scale parameter.

As in Comin et al. (2021) we use the manufacturing good as base good, and set $\Omega_M = \varepsilon_M = 1$:

$$(\Omega_A)^{\frac{1}{\sigma}} \left(\frac{c^A_{ik}}{c^{\varepsilon_A}_{ik}} \right)^{\frac{\sigma-1}{\sigma}} + \left(\frac{c^M_{ik}}{c_{ik}} \right)^{\frac{\sigma-1}{\sigma}} + (\Omega_S)^{\frac{1}{\sigma}} \left(\frac{c^S_{ik}}{c^{\varepsilon_S}_{ik}} \right)^{\frac{\sigma-1}{\sigma}} = 1 \tag{8}$$

We assume $\sigma, \varepsilon_A, \varepsilon_S, \Omega_A$, and Ω_S are all strictly positive for each i. If $\varepsilon_A \neq 1$ or $\varepsilon_S \neq 1$, preferences are non-homothetic. When $\varepsilon_A < 1$ and $\varepsilon_S > 1$, the consumption shares of agricultural goods and services decline and increase, respectively, as income increases relative to the manufacturing good. These preferences also nest the CES case when $\varepsilon_A = \varepsilon_S = 1$, thus we assume further that at

least one of these parameters differs from unity. This preference structure allows that the preference weights on each good may vary across locations and as structural transformation occurs.

The composite agricultural good is a CES aggregator over crops. For each crop, consumption is itself a CES aggregator over origins:

$$c_{jk}^{A} = \left[\sum_{g} a_g^{1/\zeta} \left(C_{jkg}\right)^{\frac{\zeta-1}{\zeta}}\right]^{\frac{\zeta}{\zeta-1}}, \text{where} \quad C_{jkg} = \left[\sum_{i\in\mathcal{L}} \left(c_{ijkg}\right)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}. \tag{9}$$

The parameter $\zeta > 0$ governs substitution across crops, $\eta > 0$ governs substitution between origins within a crop, and a_g are crop preference weights with $\sum_{g=1}^{\mathcal{G}} a_g = 1$.

Total income for an individual living in j and working in sector k is:

$$v_{jk} = w_{jk} + \frac{R_j + \sum_{k\neq A} \Pi_j^k}{N_j}.$$

It is composed of wage w_{jk} from sector k, land rents R_j, and profits Π_j^k from the manufacturing and services sectors, which are redistributed lump-sum to local residents.

4.2 Equilibrium

Production. A forested plot is cleared whenever it is profitable to do so, that is:

$$\max_{\{g\}} \{Z_{ig}(\omega)\psi_{ig}\} > p^M Z_i^0(\omega),$$

where

$$\psi_{ig} = \left(\tilde{\alpha}_g Z_{ig}^A p_{ig} (w_{iA})^{-\alpha_g}\right)^{\frac{1}{1-\alpha_g}}, \tag{10}$$

and with $\tilde{\alpha}_g = \alpha_g^{\alpha_g}(1-\alpha_g)^{1-\alpha_g}$. Under the distributional assumptions on land productivity and clearing costs, integrating over ω yields the share of forest plots converted to agriculture in i:

$$S_i = \frac{\Phi_i^\theta}{(p^M \Lambda_i^0)^\theta + \Phi_i^\theta}, \tag{11}$$

with $\Phi_i = \left[\sum_g (\Lambda_{ig}\psi_{ig})^\theta\right]^{\frac{1}{\theta}}$. Equation (11) states that the share of forested land used for agriculture is increasing in aggregate agricultural profitability and decreasing in the average local land clearing cost. The parameter θ measures the heterogeneity in costs and productivity across plots within a location. While the cost of clearing a plot of land is fixed, whether it is profitable to do so depends on the productivity of that plot as well as the price of agricultural goods from i and the local wage. Higher values of θ mean that agricultural production will be more sensitive to changes in prices.

Conditional on land being cultivated, the share allocated to crop g is:

$$\eta_{ig} = \frac{(\Lambda_{ig}\psi_{ig})^{\theta}}{\Phi_i^{\theta}}. \tag{12}$$

A crop receives a larger land share when it has higher average suitability or higher profitability. Let $\Xi_i = S_i^{\frac{\theta-1}{\theta}} L_i^F + L_i^C$ denote effective cultivated land. Total labor demand in agriculture in region i is:

$$N_i^A = \sum_g \frac{\alpha_g}{1-\alpha_g} \eta_{ig} \frac{\Phi_i}{w_{iA}} \Xi_i. \tag{13}$$

Equation (13) states that the measure of workers engaged in agriculture is higher in locations with stronger agricultural profitability, Φ_i, and a larger amount of cultivated land, Ξ_i. Outside of agriculture, labor demand is given by:

$$N_i^k = \alpha_k \frac{p_i^k Q_i^k}{w_{ik}}, \forall k \neq A. \tag{14}$$

Total agricultural production in location i is:

$$Q_i^A = \sum_g \frac{1}{1-\alpha_g} \eta_{ig} \frac{\Phi_i}{p_{ig}} \Xi_i, \tag{15}$$

and land rents are given by:

$$R_i = \sum_g (1-\alpha_g) p_{ig} Q_{ig} - p^M \Lambda_i^0 L_i^F \left[1 - (1-S_i)^{\frac{\theta-1}{\theta}}\right]. \tag{16}$$

Consumption. Iceberg costs imply that the price of crop g produced in i and consumed in j is $p_{ijg} = \tau_{ijg} p_{ig}$. Since the manufacturing good is freely traded, we have $p_{ij}^M = p^M$ for all i, j. We further use manufacturing as numeraire, and set $p^M = 1$. In equilibrium, the quantity of crop g that an individual living in location j and working in sector k buys from location i is given by:

$$c_{ijkg} = (\tau_{ijg} p_{ig})^{-\eta} (P_{jg})^{\eta-1} e_{jkg}, \tag{17}$$

where $P_{jg} \equiv \left[\sum_{i \in \mathcal{L}} (\tau_{ijg} p_{ig})^{1-\eta}\right]^{\frac{1}{1-\eta}}$ designates the price index of crop g. The share of agricultural spending on crop g realized by a worker of type k in location j is then:

$$e_{jkg} = a_g \left(\frac{P_{jg}}{P_j^A}\right)^{1-\zeta} e_{jk}^A,$$

with $P_j^A \equiv \left[\sum_g a_g (P_{jg})^{1-\zeta}\right]^{\frac{1}{1-\zeta}}$.

Let e_{ik} denote total expenditures of an individual living in location i and working in sector k, and

μ_{ik}^A and μ_{ik}^S denote its expenditure shares for agriculture and services, respectively. The expenditure share on manufacturing is defined residually as $\mu_{ik}^M = 1 - \mu_{ik}^A - \mu_{ik}^S$. Appendix B.2 expresses equilibrium expenditure shares as a function of prices and real consumption. A fixed point problem determines equilibrium expenditure shares and price indices for given prices.

Sector Choice. We first consider the choice of sector k by an individual born in i conditional on migrating to a given location j. The indirect utility of working in sector $k \in \{A, M, S\}$ is $V_{ijk} = b_{jk}/\delta_{ij} \times v_j^k/P_j^k \times \zeta_{ijk}$. Using the properties of the extreme value distribution, the fraction of workers in j that work in agriculture is:

$$n_{jA} = \frac{(v_j^A/P_j^A)^\gamma}{\sum_k \left(\beta_j^k v_j^k/P_j^k\right)^\gamma}. \tag{18}$$

Migration. We next consider the choice of migrating to j for an individual born in i. The migration rate from i to j is:

$$m_{ij} = \frac{\left(V_{ij}\right)^\kappa}{\sum_{j'} \left(V_{ij'}\right)^\kappa}, \tag{19}$$

where $V_{ij} = b_j/\delta_{ij}\big((v_j^A/P_j^A)^\gamma + (\beta_j^M v_j^M/P_j^M)^\gamma + (\beta_j^S v_j^S/P_j^S)^\gamma\big)^{1/\gamma}$ is the expected indirect utility from migrating to j.

Equilibrium Definition. An equilibrium is given by prices $\{w_i^A, w_i^M, w_i^S, p_{ig}, p_i^S, P_{jg}, P_j^A, P_j^M, P_j^S\}$ and quantities $\{c_{ijkg}, c_{jk}^M, c_{jk}^S, S_i, \eta_{ig}, Q_{ig}, Q_i^M, Q_i^S, N_i^A, N_i^M, N_i^S, N_i, m_{ij}\}$ for all i, j, k, g, such that:

- Labor markets clear in each domestic location i: $N_i^A + N_i^M + N_i^S = N_i$.
- Labor supply is consistent with migration decisions: $\sum_i m_{ij}\bar{N}_i = N_j$.
- In each location j markets clear for:
  - Each crop g:
$$Q_{ig} = \sum_{j \in \mathcal{L}} \sum_{k \in \{A,M,S\}} c_{ijkg}\tau_{ijg}N_j^k.$$
  - Services: $Q_i^S = c_i^S N_i$
  - Manufacturing by Walras' Law
- Manufacturing profits, services profits, and land rents are consistent with profit maximization in each sector.
- Consumption is consistent with the budget constraint, $\sum_k e_{ik}N_i^k = \sum_k w_{ik}N_i^k + R_i + \Pi_i^M + \Pi_i^S$.

4.3 Model Discussion

Since Kuznets (1973), economic development has been understood to involve the reallocation of economic inputs from agriculture to manufacturing, and eventually to services. The relationship between economic development and economic activity — structural transformation — has been documented both historically and cross-sectionally (Herrendorf et al., 2014). In the context of our model, structural transformation comprises both productivity growth in the three sectors — agriculture, manufacturing and mining, and services — and rising income. Before quantifying the model, we first describe, qualitatively, how these components of structural transformation affect deforestation through the lens of our model.

Manufacturing and Mining Productivity Growth In our model, an increase in manufacturing and mining productivity in one location raises local labor demand in that sector. This draws workers out of agriculture and services locally, and from other locations via migration. Locally, this drives up wages in the other sectors, and the fall in local agricultural employment reduces land use which reduces local deforestation.[16] At the same time, rising local incomes cause all workers to demand more agricultural goods, which increases local land use and deforestation.

Productivity growth also affects other locations through both migration and trade. Local productivity growth in manufacturing and mining drives down the relative price of manufacturing and mining goods everywhere (Matsuyama, 2009). This reduces labor demand in that sector in all other locations, which pushes workers into services and, importantly, agricultural work which increases deforestation elsewhere. However, this affect is offset — and may be reversed — in locations that are more connected via migration. Emigration drives up labor productivity, and hence labor demand, in all sectors, offsetting the trade effect. These are the migration and trade channels highlighted in the motivating evidence.

Services Productivity Growth In our model, the key difference between services and manufacturing and mining, is that services are non-traded. This implies that a rise in local services productivity will have a similar ambiguous effect on local deforestation through the labor reallocation and income channels. Beyond the locality, the increase in service productivity will draw in migrants from other locations and hence reduce deforestation. At the same time, local demand for agricultural goods will increase, and hence both the migration and trade channels will also be present.

[16] In our model, manufacturing and mining does not affect local deforestation directly. This may be true if land or wood are inputs to production or if manufacturing growth causes urbanization and demand for housing (Coeurdacier et al., 2026). While there is some local deforestation caused by manufacturing and mining booms in sub-Saharan Africa (Morton et al., 2026), these effects are quantitatively minor in our setting relative to agricultural land use: tree-cover loss co-located with built-up expansion is 1.3 percent of all loss over 2001–2020 (2.4 percent including pre-existing built-up land in 2000).

Agricultural Productivity Growth We assume Hicks-neutral agricultural productivity growth since we study a diverse set of countries where productivity growth, to the extent it has occurred, has come from different sources (Benin and Nin-Pratt, 2016; Carter et al., 2021; Wollburg et al., 2023). A rise in agricultural productivity therefore has no effect on the land-to-labor ratio, and whether this leads to a rise or fall in land use depends upon the location's openness to trade (Matsuyama, 1992). If trade costs are high, then the location is primarily serving its own market, and it will be able to satisfy demand with less labor and less land. On the other hand, if trade costs are low, the location will now serve other markets at a lower price than before, causing aggregate demand to rise, and leading to an increase in deforestation. Hence, whether agricultural productivity growth leads to a rise or fall in deforestation depends upon the locations agricultural trade costs, both internally and internationally.

Rising Income While heterogeneous productivity growth rates affect structural transformation through supply-side forces, rising income affects structural transformation through demand-side forces. At low levels of income, most expenditure is on food to meet minimal requirements — what is known as the "food problem" (Schultz, 1953; Gollin et al., 2007). As income rises, a larger share of expenditure can be allocated to non-food items, which we model as non-homothetic preferences (Kongsamut et al., 2001; Boppart, 2014). This implies that as incomes rise, demand for manufacturing and mining goods and services rise proportionally more than demand for agriculture. Nonetheless, demand for agriculture rises as well, which generates more deforestation.

An important caveat to our model is that individuals do not have non-homothetic preferences over specific agricultural goods. This may be an important channel if staple crops are less land intensive than cash crops. This channel has been emphasized in the context of Latin America, where cattle ranching is a major driver of deforestation, and the expenditure share on meat is increasing in income (Ferraro and Simorangkir, 2020). In the context of SSA, meat consumption is much lower, and cattle ranching is not a major contributor of deforestation. In Appendix Figure A.1, we report the share of commodity-driven deforestation due to beef consumption across world regions. Overall, across Africa, beef accounted for 18% of deforestation between 2001 and 2010, compared with 72% in South America.

An alternative channel through which rising income may affect deforestation is energy use (Foster and Rosenzweig, 2003; Baland et al., 2018). In particular, if households shift from consuming firewood to charcoal, this may increase pressure on forests (Herzog et al., 2025). In Appendix Figure A.2, we report the primary cooking fuel among SSA households in the Demographic and Health Surveys in the 1990s, 2000s, and 2010s. While charcoal use is increasing over time, it remains low, with only 12 % of households using it as their primary source in the 2000-2009 period. Conceptually, incorporating charcoal into the model would change little to the key mechanisms studied - rising demand for charcoal would act like rising demand agricultural goods, increasing the value of clearing forested land.

5 Quantification

In this section, we describe how we bring our model to data on the 28 sample countries from 2000 and 2010.[17]

5.1 Calibration

To quantify the model, we use data from the year 2000 or 2010. When the data is taken directly from censuses, we choose the census closest to the respective year.

Data on agricultural output, Q_i^A, comes from GAEZ (Fischer et al., 2021). To calibrate our productivity parameter in agriculture, Λ_i, we use the data on potential yields from the same source, under low-inputs and rain-fed irrigation.[18] As in Porteous (2024), to account for potential measurement errors or multiple harvests, we consider a region-level scaling factor where we have $\Lambda_i = a_i \tilde{Z}_i^{\{\text{GAEZ}\}}$. The parameter $\tilde{Z}_i^{\{\text{GAEZ}\}}$ represents the average value of potential yields in region i, and to obtain Λ_i, we calibrate a_i by inverting the equation of agricultural output, given by (15).

Population N_i is calculated from the gridded population of the world (CIESIN, 2018). The amount of land under forest L_i^F is obtained from Hansen (Hansen et al., 2013). To estimate the value of output in manufacturing and services, we use data on national GDP in those two sectors and allocate it across sub-national units using nightlights, making an adjustment based on the number of people working outside of agriculture for places where there are minimal nightlights (Elvidge et al., 2022). We show in appendix that bilateral migration costs can be recovered from migration rates as in Bryan and Morten (2019):[19]

$$\delta_{ij} = \left(\frac{m_{ij} m_{ji}}{m_{ii} m_{jj}} \right)^{-\frac{1}{2\kappa}}. \tag{20}$$

We construct bilateral trade costs τ_{ij} using travel-time between each pair of locations, mapped into iceberg trade costs using the transport-cost elasticity estimated by Sotelo (2020).

[17] We choose these years since they are those where we have estimates of agricultural production and yields at the grid cell level.

[18] We followed Adamopoulos and Restuccia (2022) and computed the potential agricultural yield across 22 major crops, valuing the yield with international prices from the year 2000 and weighting each crop based on its share of local land use in 2000.

[19] For Ethiopia and Nigeria, we use LSMS-ISA data to construct employment, population, and migration variables (World Bank, n.d.). For Ethiopia, employment and population are constructed from wave 2 of the Ethiopia Socioeconomic Survey, collected in 2013–2014, while migration is measured in wave 3, collected in 2015–2016, using information on household members who left since the previous wave (Central Statistical Agency of Ethiopia and World Bank, n.d.). For Nigeria, employment and population are constructed from wave 1 of the General Household Survey-Panel, collected in 2010–2011, and migration is measured similarly to Ethiopia, using information from wave 2, collected in 2012–2013 (National Bureau of Statistics, Federal Republic of Nigeria and World Bank, n.d.). In both countries, we define migrants as prime-age individuals whose reported time since leaving the household is no longer than the median time between survey waves (equivalent to around two years), and whose reported destination differs from the origin region in Ethiopia or state in Nigeria.

Finally, we calibrate a set of parameters that are held fixed across all specifications. Table 5 summarizes these parameters and their sources. We take the shape parameter governing land-use allocation θ and the elasticity of substitution between crops ζ from Sotelo (2020). We adopt the elasticity of substitution across agricultural varieties, η, from Costinot et al. (2016). Sectoral labor shares, α_k, are taken from Gollin (2002) and Dinh and Clarke (2012). We borrow the elasticity of migration, κ, from Bryan and Morten (2019). For the elasticity of sectoral choice, we take the average of the agricultural and non-agricultural productivity dispersion parameters in Lagakos and Waugh (2013). For σ as well as ε_A and ε_S, we use Comin et al. (2021) estimates for non-OECD countries.

5.2 Estimation

The remaining model parameters are estimated through a sequence of nested fixed-point loops. Each loop targets a distinct set of equilibrium conditions and empirical moments covering goods markets, labor markets, land use, and migration. The procedure from the inner to the outer loop is as follows:

Inner loop The inner loop solves the fixed point problem and derives equilibrium expenditure shares μ_i^k and price indices P_{ik} for given agricultural and services prices P_i^A and P_i^S.

Price loop The price loop estimates prices P_i^A and P_i^S, productivity terms (Λ_i, Z_i^M, and Z_i^S), land clearing costs Λ_i^0, and sector-specific amenities β_i^k. It takes as given wages w_i, quantities produced in each sector Q_i^k, employment in each sector N_i^k, and land use (L_i^C, L_i^F and S_i). More precisely, we iterate on P_i^A and P_i^S to find a solution that clears each market. Within each iteration, we set Λ_i to match Q_i^A, we set Z_i^M, Z_i^S to match Q_i^M and Q_i^S, we set Λ_i^0 to match S_i by inverting equation (11), and we set β_i^k to match N_i^k ($k \in \{M, S\}$). Every iteration of the price loop solves the inner loop.

Wage loop We then iterate on wages w_i until labor demand equals labor supply in each sector and location N_i^k, which we take as given. Within each iteration, we run the price loop to ensure that goods markets clear.

Ω **loop** We then iterate on Ω to match the total share of expenditure on agricultural goods at the national level. Within each iteration, we run the wage loop to ensure goods and labor markets clear.

τ_F **loop** We then iterate on τ_{Fg}^M to match crop-level agricultural imports to the rest of the world and iterate on τ_{Fg}^X to match crop-level agricultural exports. Within each iteration, we run the Ω loop so that goods and labor markets clear.

Migration loop Finally, this loop estimates amenities b_i that rationalize observed migration flows from the nearest census given migration costs calibrated from (20).

Three key parameters estimated by this model are productivity in the agricultural sector, Λ_i, productivity in manufacturing, Z_i^M, and productivity in services, Z_i^S. Estimated growth in productivity in these sectors for each country is presented in Figure A.3. We can see that, for most countries, productivity increased in all sectors. However, there exists an important variation across countries: for some countries, productivity declined by more than 1% in at least one sector (Botswana, Gabon, Madagascar, Malawi, Mali, Togo, Uganda, and Zimbabwe). In the next section, we will then explore how these estimated productivity changes impacted structural change and deforestation.

5.3 Validation with External Data Sources

In order to test how well our model performs, we first compare the estimated average consumption in each location in our model to that measured in households surveys, as described in Appendix C.6. Specifically, we compare model estimates for 2010 with estimates based on the LSMS survey closest to 2010 for eight countries: Ethiopia, Ghana, Malawi, Mali, Nigeria, South Africa, Tanzania, and Uganda. This test is demanding, because household surveys measure consumption noisily and are not generally designed to be representative at the level at which we undertake our analysis. Moreover, our model excludes some important sources of income, such as government transfers (including public sector employment), remittances, etc. Nonetheless, we believe the exercise is valuable in testing whether to a first approximation we are predicting which areas are relatively better off than others. In Figure 2a, we plot relative real adult-equivalent consumption as estimated using the model and household surveys. While the fit is not perfect, with the model estimating greater dispersion in consumption than in the survey data, there is a clear positive correlation between the two measures (0.21), significant at the 0.1 % level.

In addition to consumption levels, we are also interested in how well we predict variation in the share of expenditure on agriculture, since this gives us a goodness-of-fit test for the non-homotheticity parameters. For this, we use LSMS survey data from four countries where it is possible to construct such shares: Ethiopia, Malawi, Mali, and Nigeria. Figure 2b presents the results. The survey data has systematically higher expenditure shares on agriculture than the model, since large parts of expenditure on food actually accrue to service sectors (e.g. retail, transport). Importantly for us, however, we observe a similar slope in the survey data and in the model estimates: The correlation between the two measures is 0.46, significant at the 0.1% level. This suggests that the model captures reasonably well the extent to which rising incomes reduce the expenditure share on agricultural goods.

6 Model Results

6.1 Productivity Changes, Structural Transformation, and Deforestation

Within our model, we consider structural transformation as mainly stemming from increases in productivity. Increases in productivity will increase consumption, which, due to non-homothetic preferences, will push relative demand towards non-agricultural goods and services, encouraging movement out of the agricultural sector. Of course, the overall effect will also depend heavily on the source of the productivity changes. In what follows, we therefore aim to isolate the impacts of productivity changes in both the agricultural and the non-agricultural sectors, as well as all sectors together, on structural transformation and on deforestation.

Our counterfactual analysis exploits the fact that we have estimated our model on two years of data, 2000 and 2010, and compares actual outcomes to model estimates of what would have happened if productivity had not changed over the period - i.e. if productivity had stayed at the 2000 levels. An advantage of this strategy is that we consider meaningful counterfactuals, and it also allows to answer the question of how much of observed changes in outcomes over the period can be explained by these productivity changes. However, this means that the counterfactual productivity changes we simulate vary substantially across countries and sectors, as shown in Figure A.3.

Effects on structural change In Figure 3 we first consider the role of productivity changes in explaining structural change, i.e. changes in the share of workers who work outside of agriculture between 2000 and 2010. Looking at the second bar of Figure 3(a), we find that if we run a counterfactual of our model where productivity is as estimated in 2000, but all other parameters are those estimated as estimated in 2010, we obtain a rate of agricultural employment very close to the 2000 level. This suggests that, overall, productivity changes explain most of the structural transformation we actually observe over the period. In appendix Figure A.4, we compare the effect of turning off productivity changes on structural change with and without letting population move in response to them between 2000 and 2010. As expected, we find that migration amplifies the effect of productivity changes on structural change.

Running counterfactuals in which we turn off productivity growth sector-by-sector, we find that the largest individual component is changes in productivity in the mining and manufacturing sectors (fourth bar in Figure 3(a)). We therefore proceed by considering the counterfactual where all productivity values are kept at their 2000 levels as that of most interest when understanding the impacts of structural change. However, agricultural productivity also plays a role in structural transformation (third bar in Figure 3(a)), which in some countries is even dominant.

Effects on deforestation What were the overall impacts of these productivity changes on deforestation? In Figure 3(b), the outcome is the share of total land deforested between between 2010 and 2015. Comparing the first and second bar of Figure 3(b), we find that without productivity changes that took place between 2000 and 2010 deforestation would have been 1.54pp (28%) *higher* by 2015, i.e. that productivity changes over the period substantially reduced deforestation. In appendix Figure A.4, we compare the effect of turning off productivity changes on deforestation with and without letting population move in response to them between 2000 and 2010. We find that the forest-saving impact of productivity changes is magnified by migration, which is consistent with the reduced form evidence.

We now turn to counterfactuals that switch off one sector at a time. The third bar of Figure 3(b) shows that without productivity growth in the agricultural sector, deforestation overall would have been 1.37pp (+25%) *higher* by 2015. This suggests that agricultural productivity growth during this period was land saving, which is consistent with the Borlaug hypothesis. In contrast, turning off manufacturing and mining productivity growth has a small average negative effect on deforestation, it decreases by 0.22pp (–4%)/ Turning off service sector productivity growth has an even smaller effect (0.1pp or +1.8%). This estimate suggests that non-agricultural productivity growth had little effect on deforestation. However, cross-country averages hide wide heterogeneity across countries.

Cross-country Heterogeneity In Figure 4 we show the estimated impact of manufacturing and mining productivity changes on structural change and deforestation across the 28 countries in our sample.[20] We obtain this by computing the difference between structural change and deforestation at endline and in the counterfactual that keeps manufacturing productivity fixed. The cross-country average suggests that manufacturing productivity changes increased non-agricultural employment in most countries (5pp on average). However, there is a dramatic difference between Angola, which experienced manufacturing productivity growth, and Gabon who saw a manufacturing productivity decline (Appendix Figure A.3). We estimate that these changes in manufacturing productivity led to 10pp increase in non-agricultural employment in Angola, and a 5pp decline in Gabon.

Turning to the effects of manufacturing productivity growth on deforestation, the average effect across the 28 countries is close to zero (0.22pp). But once again, the effects are completely different in Angola, where manufacturing productivity growth led to a 1.7pp increase in deforestation, and Gabon, where manufacturing productivity slowdown led to a 0.8pp decline in deforestation. More surprisingly, in Angola and Ghana similar rates of manufacturing productivity growth (Appendix Figure A.3) led to similar rates of structural transformation (10pp and 9.6pp respectively), but had opposite effects on deforestation (+1.7pp in Angola against –0.65pp in Ghana). Hence the vast cross-country heterogeneity we observe is not only linked to different productivity growth paths, but also to how they translate into more or less deforestation. Looking at the maps of manufacturing

[20]We focus on one sector to simplify cross-country comparisons, but document even larger heterogeneity across countries in the impact of all productivity changes on structural change and deforestation in Appendix Figure A.5

productivity growth and forest cover in Angola and Ghana suggests a role of geography (Appendix Figure A.6). In Angola, manufacturing productivity growth happened in the North West, a densely forested area, whereas in Ghana it was strongest in the Northern half, which has little forest cover. We explore the role of internal geography in the next section.

6.2 Deforestation and the Internal Geography of Structural Change

In this final subsection we bring the quantification exercise in line with our motivating facts and study the effect of structural change on deforestation depending on the respective location of productivity changes and forests.

Shift-share estimation with model estimates Beyond simple geographic proximity, the motivating evidence in section 3 highlighted the role of migration and trade links for the effects of structural change on deforestation. We undertake the same regression analysis we did in Section 3, only this time with data generated using model counterfactuals. The dependent variable is the difference between actual deforestation in 2010-15 and deforestation in the counterfactual which holds all productivities at their 2000 levels. We construct shift-shares as in Section 3. The shares are computed in the same way, only the shifts are the difference between the actual growth in non-agricultural employment and the one in the counterfactual where all productivities are held at their 2000 value.

When we regress the model based measure of deforestation on these new shift-shares (column 1 in Table 6) , we obtain estimates that are similar in sign and magnitude with those from our reduced form (column 6 in Table 2). Hence, the model results point in the same direction as the reduced form: structural transformation reduces deforestation via migration and increases it via trade. As in Section 3, we test the mechanisms by splitting our sample between countries that are more or less open to international trade. If most of the food is imported, changes in the internal demand for agricultural goods due to structural transformation should matter less for deforestation. We define more open and more closed countries as below and above median import trade frictions, based on the estimated τ_{Fg}^{M} which we average across crops using import values. As column 2 and 3 in Table 6 shows, the effects of structural change on deforestation are larger in economies more closed to international trade.

Where is structural change most harmful to forests? Our results so far are consistent with the idea that a key determinant of the relationship between structural change and deforestation is the location of the productivity shocks that drive structural change. To test this idea further, we undertake a new set of counterfactuals whereby, instead of changing productivity in the country as a whole, we change only productivity in an individual location. In other words, we consider how much deforestation would have taken place had productivity changed in every location except for one. Since we want to focus on productivity changes that induce local structural change, we again only change

productivity in manufacturing and mining. Appendix Figure A.7 displays the results of these 1324 separate counterfactuals. Since geographic units vary substantially both in size and the direction of the productivity shock, we focus on the amount of deforestation caused by the shock (as a share of the total national area) divided by the share of employees who move out of agriculture (in the whole country). Notably, we see a large amount of heterogeneity within each country, with productivity shocks sometimes causing deforestation and other times reducing it, depending on their precise location.

While no clear pattern is immediately obvious from the map, our earlier results suggest that an important determining factor would be the location of the shock relative to forests. In particular, we would expect that shocks that cause structural change will decrease deforestation in locations that are 'migration-close' to forests - i.e. ones which draw migrants from forested areas - while they will increase deforestation if locations are 'trade-close' to forests - that is, they consume food grown in forested areas. To test this, in Figure 5, we take the results from the unit-specific counterfactuals and plot them against the average forest cover in 'migration-close' and 'trade-close' locations. Consistent with our earlier result, we see that productivity shocks cause less deforestation per worker moved out of agriculture when they take place in areas that draw migrants from forested areas. On the other hand, shocks cause more deforestation when they take place in areas that source food from forested areas.

Where can we foster structural change if we want to save forests? Based on these results, we can now attempt to answer a more forward-looking question - how can structural change continue while minimizing deforestation? While we cannot possibly provide a comprehensive answer to this question, the results of our model suggest that an important component of an answer would be to encourage structural transformation to take place in 'migration-close' locations more than 'trade-close' ones. To test this idea, we undertake two final counterfactuals, which each increase manufacturing and mining productivity by 100% in 5% of the locations in each country. We first rank locations by the difference between forest cover in their 'migration-close' and 'trade-close' locations, and then in one counterfactual shock the top 5% 'migration-close' locations, while in the other we shock the top 5% 'trade-close'. Figure 6 provides a summary of the results. When it comes to non-agricultural employment, both counterfactuals have similar impacts, with an increase of close to 1.3% of the observed growth in the non-agricultural employment share. In terms of deforestation, however, the impacts are very different - while increasing productivity in the 'trade-close' leads to a 1% increase in deforestation, increasing productivity in the 'migration-close' locations decreases deforestation by close to 2%. Hence policies that encourage structural change can have very different impacts on deforestation depending on where they are targeted.

7 Conclusion

The overall impact of structural change on deforestation in sub-Saharan Africa is theoretically ambiguous. More employment in sectors with higher productivity than agriculture will increase domestic demand for agricultural products, thereby increasing the return to clearing land. At the same time, migration from rural areas to urban ones will decrease the return to land clearance. In this paper, we have found that both mechanisms play a significant role, with the dominant mechanism varying across countries, depending on the source and location of the productivity changes.

Our results lead to important policy conclusions. In general, increases in agricultural productivity are likely to put forests least at risk when they occur in relatively closed economies, while increases in manufacturing and mining productivity are likely to be most beneficial for forests when they take place in parts of the country that draw more migrants from forested areas. Hence, when pursuing the dual aims of protecting forests and increasing structural change, policy makers should consider the spatial distribution of where productivity gains should be encouraged.

A number of important caveats should be kept in mind when interpreting our results. Crucially, we are only measuring one of the environmental impacts of structural transformation, and of course issues are likely to arise due to pollution and pressure on other natural resources. Our continent-wide approach also prevents us from looking at heterogeneity in the value of forests or the costs of tree-cover loss. As better data becomes available, an important avenue of future research will be to improve measures of trade-costs and environmental degradation.

References

Abman, Ryan and Clark Lundberg, "Does free trade increase deforestation? The effects of regional trade agreements," *Journal of the Association of Environmental and Resource Economists*, 2020, *7* (1), 35–72.

Adamopoulos, Tasso and Diego Restuccia, "Geography and Agricultural Productivity: Cross-Country Evidence from Micro Plot-Level Data," *The Review of Economic Studies*, July 2022, *89* (4), 1629–1653.

Adão, Rodrigo, Michal Kolesár, and Eduardo Morales, "Shift-share designs: Theory and inference," *The Quarterly Journal of Economics*, 2019, *134* (4), 1949–2010.

Alix-Garcia, Jennifer, Craig McIntosh, Katharine RE Sims, and Jarrod R Welch, "The ecological footprint of poverty alleviation: evidence from Mexico's Oportunidades program," *Review of Economics and Statistics*, 2013, *95* (2), 417–435.

Andersen, Lykke E, *The dynamics of deforestation and economic growth in the Brazilian Amazon*, Cambridge University Press, 2002.

Angelsen, Arild, "Agricultural expansion and deforestation: modelling the impact of population, market forces and property rights," *Journal of Development Economics*, 1999, *58* (1), 185–218.

_ , Mariel Aguilar-Støen, John Herbert Ainembabazi, Edwin Castellanos, and Matthew Taylor, "Migration, Remittances, and Forest Cover Change in Rural Guatemala and Chiapas, Mexico," *Land*, 2020, *9* (3).

Araujo, Rafael, Juliano Assunção, and Arthur Amorim Bragança, "The Effects of Transportation Infrastructure on Deforestation in the Amazon: A General Equilibrium Approach," resreport 10415, World Bank 2023.

Asher, Sam, Teevrat Garg, and Paul Novosad, "The ecological impact of transportation infrastructure," *Economic Journal*, 2020, *130* (629), 1173–1199.

Atkin, David and Dave Donaldson, "Who's Getting Globalized? The Size and Implications of Intra-national Trade Costs," NBER Working Papers 21439, National Bureau of Economic Research, Inc Jul 2015.

Baland, Jean-Marie, François Libois, and Dilip Mookherjee, "Forest Degradation and Economic Growth in Nepal, 2003–2010," *Journal of the Association of Environmental and Resource Economists*, 2018, *5* (2), 401–439.

Balboni, Clare, Aaron Berman, Robin Burgess, and Benjamin A. Olken, "The Economics of Tropical Deforestation," *Annual Review of Economics*, July 2023, *15* (3), 723–54.

_ and Joseph S. Shapiro, "Chapter 9 - Spatial environmental economicsWe thank Isabella Masetto, Meghna Padmanabhan, and Yifan Wang for research assistance, Costas Arkolakis, Dave Donaldson, Lucy Hackett, Stephan Heblich, Matt Kahn, Sarah O'Brien, Stephen Redding, Charles Taylor, Chris Timmins, Daniel Xu, our discussant Scott Taylor, and participants in the Coase Conference

and Princeton handbook conference for useful comments, Chris Knittel for sharing data, and NSF SES-2214949 for generous support.," in Dave Donaldson and Stephen J. Redding, eds., *Handbook of Regional and Urban Economics*, Vol. 6 of *Handbook of Regional and Urban Economics*, Elsevier, 2025, pp. 585–652.

_ **, Robin Burgess, Anton Heil, Jonathan Old, and Benjamin A. Olken**, "Cycles of Fire? Politics and Forest Burning in Indonesia," *AEA Papers and Proceedings*, May 2021, *111*, 415–19.

Barbier, E. B. and J. C. Burgess, "The Economics of Tropical Deforestation," *Journal of Economic Surveys*, July 2001, *15* (3), 413–433.

Benin, Samuel and Alejandro Nin-Pratt, "Intertemporal Trends in Agricultural Productivity," in Samuel Benin, ed., *Agricultural Productivity in Africa: Trends, Patterns, and Determinants*, Washington, DC: International Food Policy Research Institute, 2016, pp. 25–104.

Boppart, Timo, "Structural change and the Kaldor facts in a growth model with relative price effects and non-Gorman preferences," *Econometrica*, 2014, *82* (6), 2167–2196.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "Quasi-Experimental Shift-Share Research Designs," *Review of Economic Studies*, 2022, *89* (1), 181–213.

Brewer, Julia, Ashley Larsen, and Frederik Noack, "The land use consequences of rural to urban migration," *American Journal of Agricultural Economics*, 2022, *n/a* (n/a), 1–29.

Bryan, Gharad and Melanie Morten, "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia," *Journal of Political Economy*, 2019, *127* (5), 2229–2268.

Burgess, Robin, Matthew Hansen, Benjamin A. Olken, Peter Potapov, and Stefanie Sieber, "The Political Economy of Deforestation in the Tropics," *The Quarterly Journal of Economics*, 2012, *127* (4), 1707–1754.

Burzynski, Michal, Christoph Deuster, Frederic Docquier, and Jaime de Melo, "Climate Change, Inequality, and Human Migration," *Journal of the European Economic Association*, 2022, *20* (3), 1145–1197.

Card, David, "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, January 2001, *19* (1), 22–64.

Carter, Michael, Rachid Laajaj, and Dean Yang, "Subsidies and the African Green Revolution: direct effects and social network spillovers of randomized input subsidies in Mozambique," *American Economic Journal: Applied Economics*, 2021, *13* (2), 206–229.

Central Statistical Agency of Ethiopia and World Bank, "Ethiopia Socioeconomic Survey (ESS)," World Bank Microdata Library, LSMS-ISA. Waves used: 2013–2014 and 2015–2016.

Central Statistics Agency (CSA) of Ethiopia and World Bank LSMS-ISA, "Socioeconomic Survey 2013-2014, Ethiopia [dataset]," World Bank Microdata Library 2013-2014.

Chomitz, Kenneth M and David A Gray, "Roads, land use, and deforestation: a spatial model applied to Belize," *World bank economic review*, 1996, *10*, 487–512.

CIESIN, "Gridded Population of the World, Version 4 (GPWv4): Population Count, Revision 11 (Version 4.11)," 2018. [Data set].

Coeurdacier, Nicolas, Cecile Gaubert, and Emeric Henry, "Structural Transformation of Land and the Carbon Balance," 2026.

Comin, Diego, Danial Lashkari, and Martí Mestieri, "Structural Change With Long-Run Income and Price Effects," *Econometrica*, 2021, *89* (1), 311–374.

Conte, Bruno, "Climate Change and Migration: The Case of Africa," CESifo Working Paper Series 9948 2022.

Copernicus Climate Change Service, "Land cover classification gridded maps from 1992 to present derived from satellite observations," 2019.

Costinot, Arnaud, Dave Donaldson, and Cory Smith, "Evolving comparative advantage and the impact of climate change in agricultural markets: Evidence from 1.7 million fields around the world," *Journal of Political Economy*, 2016, *124* (1), 205–248.

Curtis, Philip G., Christy M. Slay, Nancy L. Harris, Alexandra Tyukavina, and Matthew C. Hansen, "Classifying drivers of global forest loss," *Science*, 2018, *361* (6407), 1108–1111.

de Vries, Gaaitzen, Linda Arfelt, Dorothea Drees, Mareike Godemann, Calumn Hamilton, Bente Jessen-Thiesen, Ahmet Ihsan Kaya, Hagen Kruse, Emmanuel Mensah, and Pieter Woltjer, "The Economic Transformation Database (ETD): content, sources, and methods," February 2021.

Desmet, Klaus and Esteban Rossi-Hansberg, "On the spatial economic impact of global warming," *Journal of Urban Economics*, 2015, *88* (C), 16–37.

Dijkstra, Edsger W., "A note on two problems in connexion with graphs," *Numerische Mathematik*, 1959, *1* (1), 269–271.

Dinh, Hinh T and George RG Clarke, *Performance of manufacturing firms in Africa: An empirical analysis*, World Bank Publications, 2012.

Direction Générale de la Statistique (DGS), Gabon, "Recensement général de la population et de l'habitat, du 1er au 31 juillet 1993 : Principaux résultats [PDF]," Libreville, Gabon: Institut National de la Statistique (INSTAT Gabon) 1993.

_ , "Résultats globaux du Recensement Général de la Population et des Logements de 2013 du Gabon (RGPL-2013) [PDF]," Libreville, Gabon: Direction Générale de la Statistique (DGS) 2015.

Dominguez-Iino, Tomas, "Efficiency and Redistribution in Environmental Policy: An Equilibrium Analysis of Agricultural Supply Chains," *Journal of Political Economy*, 2026, *134* (4), 1075–1109.

Donaldson, Dave and Richard Hornbeck, "Railroads and American economic growth: A "market access" approach," *The Quarterly Journal of Economics*, 2016, *131* (2), 799–858.

Elvidge, Christopher D., Kimberly Baugh, Tilottama Ghosh, Mikhail Zhizhin, Feng Chi Hsu, Tamara Sparks, Morgan Bazilian, Paul C. Sutton, Kenneth Houngbedji, and Ran Goldblatt, "Fifty years of nightly global low-light imaging satellite observations," *Frontiers in Remote Sensing*, 2022, *3*.

FAO, *Global Forest Resources Assessment 2020: Main report*, Rome: FAO, 2020.

Farrokhi, Farid, Elliot Kang, Heitor S. Pellegrina, and Sebastian Sotelo, "Deforestation: A Global and Dynamic Perspective," 2023. Manuscript.

Ferraro, Paul J. and Rhita Simorangkir, "Conditional cash transfers to alleviate poverty also reduced deforestation in Indonesia," *Science Advances*, 2020, *6* (24), eaaz1298.

Ferreira, Susana, "Deforestation, Property Rights, and International Trade," *Land Economics*, 2004, *80* (2), 174–193.

Fetzer, Thiemo and Samuel Marden, "Take What You Can: Property Rights, Contestability and Conflict," *Economic Journal*, 2017, *127* (601), 757–783.

Fischer, Günther, Freddy O. Nachtergaele, Harrij T. van Velthuizen, Federico Chiozza, Giuliano Franceschini, Michel Henry, David Muchoney, and Sylvia Tramberend, "Global Agro-Ecological Zones v4 – Model Documentation," techreport, Food and Agriculture Organization of the United Nations (FAO), Rome, Italy 2021.

Food and Agriculture Organization (FAO), "Food Price Monitoring and Analysis (FPMA) Tool, Version 4," https://fpma.fao.org/giews/fpmat4/ 2026.

Food and Agriculture Organization of the United Nations, "Food Balance Sheets: A Handbook," 2001.

_ , "GIEWS Country Brief: Mauritius," Global Information and Early Warning System 2021.

_ , "FAOSTAT Agricultural Producer Prices," FAOSTAT Statistical Database 2023.

_ , "FAOSTAT Food Balance Sheets: Historic Series (old methodology, up to 2013)," FAOSTAT Statistical Database 2023.

Foster, Andrew D. and Mark R. Rosenzweig, "Economic Growth and the Rise of Forests*," *The Quarterly Journal of Economics*, 05 2003, *118* (2), 601–637.

Funk, Chris, Pete Peterson, Martin Landsfeld, Diego Pedreros, James Verdin, Shraddhanand Shukla, Gregory Husak, James Rowland, Laura Harrison, Andrew Hoell, and Joel Michaelsen, "The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes," *Scientific Data*, 2015, *2*, 150066.

Gaskin, T. and G. J. Abel, "Deep learning four decades of human migration," *Nature*, 2026, *655*, 148–157.

Gollin, Douglas, "Getting income shares right," *Journal of political Economy*, 2002, *110* (2), 458–474.

_ **and Julien Wolfersberger**, "Agricultural trade and deforestation: the role of new roads," *The Economic Journal*, 2025, p. ueaf125.

_ **, Stephen L. Parente, and Richard Rogerson**, "The food problem and the evolution of international income levels," *Journal of Monetary Economics*, 2007, *54* (4), 1230–1255.

Grossman, Gene M and Alan B Krueger, "Economic Growth and the Environment," *The Quarterly Journal of Economics*, 1995, *110* (2), 353–377.

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend, "High-Resolution Global Maps of 21st-Century Forest Cover Change," *Science*, 2013, *342* (6160), 850–853.

Herrendorf, Berthold, Richard Rogerson, and Ákos Valentinyi, "Chapter 6 - Growth and Structural Transformation," in Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth*, Vol. 2 of *Handbook of Economic Growth*, Elsevier, 2014, pp. 855–941.

Hersbach, H., B. Bell, P. Berrisford, G. Biavati, A. Horányi, J. Muñoz Sabater, J. Nicolas, C. Peubey, R. Radu, I. Rozum, D. Schepers, A. Simmons, C. Soci, D. Dee, and J.-N. Thépaut, "ERA5 hourly data on single levels from 1940 to present," 2023. Accessed on DD-MMM-YYYY.

Herzog, Ian, Yue Yu, and Clark Lundberg, "Transportation Infrastructure, Market Access, and Deforestation," SSRN Working Paper 5355483, Social Science Research Network 2025. Conditionally accepted at American Economic Journal: Economic Policy.

Hoang, Nguyen Tien and Keiichiro Kanemoto, "Mapping the deforestation footprint of nations reveals growing threat to tropical forests," *Nature Ecology and Evolution*, 2021, *5* (6), 845–853.

ILO, "ILO modelled estimates database (accessed via World Development Indicators)," Technical Report 2021.

Imbert, Clement, Marlon Seror, Yifan Zhang, and Yanos Zylberberg, "Migrants and Firms: Evidence from China," *American Economic Review*, June 2022, *112* (6), 1885–1914.

Institut National de la Statistique (INSTAT), Madagascar, "Recensement Général de la Population et de l'Habitation (RGPH-3) de 2018 [dataset]," 2018.

Jack, B. Kelsey and Seema Jayachandran, "Self-selection into payments for ecosystem services programs," *The National Academy of Sciences*, March 2019, *116* (12), 5326–5333.

Jaeger, David A., Joakim Ruist, and Jan Stuhler, "Shift-Share Instruments and the Impact of Immigration," resreport, NBER 2018.

Jarvis, A., H. I. Reuter, A. Nelson, and E. Guevara, "Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database," 2008. https://srtm.csi.cgiar.org.

Kaczan, David J., "Can roads contribute to forest transitions?," *World Development*, 2020, *129*, 104898.

Kongsamut, Piyabha, Sergio Rebelo, and Danyang Xie, "Beyond balanced growth," *The Review of Economic Studies*, 2001, *68* (4), 869–882.

Kuznets, Simon, "Modern Economic Growth: Findings and Reflections," *The American Economic Review*, 1973, *63* (3), 247–258.

Lagakos, David and Michael E. Waugh, "Selection, Agriculture, and Cross-Country Productivity Differences," *American Economic Review*, April 2013, *103* (2), 948–980.

Leblois, Antoine, Olivier Damette, and Julien Wolfersberger, "What has Driven Deforestation in Developing Countries Since the 2000s? Evidence from New Remote-Sensing Data," *World Development*, 2017, *92*, 82–102.

Madhok, Raahil, Frederik Noack, Ahmed Mushfiq Mobarak, and Olivier Deschenes, "Internal Migration and the Re-organization of Agriculture," Mimeo 2022.

Mano, Rui and Marola Castillo, "The Level of Productivity in Traded and Non-Traded Sectors for a Large Panel of Countries," *IMF Working Papers*, 2015, *15* (48), 1.

Massicotte, Philippe and Andy South, *rnaturalearth: World Map Data from Natural Earth* 2026. R package version 1.2.0.9000.

Matsuyama, Kiminori, "Agricultural productivity, comparative advantage, and economic growth," *Journal of economic theory*, 1992, *58* (2), 317–334.

_ , "Structural Change in an Interdependent World: A Global View of Manufacturing Decline," *Journal of the European Economic Association*, 2009, *7* (2-3), 478–486.

Morton, Oscar, Christopher G Bousfield, Prince Dégny Valé, Ieuan Lamb, Victor Maus, Robert G Bryant, and David P Edwards, "Mining triggers extensive additional deforestation in sub-Saharan Africa," *Nature*, 2026, *654* (8120), 971–977.

Müller-Crepon, Carl, Philipp Hunziker, and Lars-Erik Cederman, "Roads to Rule, Roads to Rebel: Relational State Capacity and Conflict in Africa," *Journal of Conflict Resolution*, 2021, *65* (2-3), 563–590.

National Bureau of Statistics, Federal Republic of Nigeria, "General Household Survey, Panel 2010-2011, Wave 1, Nigeria [dataset]," World Bank Microdata Library 2010-2011.

_ **and World Bank**, "Nigeria General Household Survey-Panel (GHS-Panel)," World Bank Microdata Library, LSMS-ISA. Waves used: 2010–2011 and 2012–2013.

Ngai, L Rachel and Christopher A Pissarides, "Structural change in a multisector model of growth," *American economic review*, 2007, *97* (1), 429–443.

Nunn, Nathan and Diego Puga, "Ruggedness: The Blessing of Bad Geography in Africa," *The Review of Economics and Statistics*, 2012, *94* (1), 20–36.

OECD/SWAC, "Africapolis Database," 2024.

Pellegrina, Heitor S., "Trade, productivity, and the spatial organization of agriculture: Evidence from Brazil," *Journal of Development Economics*, 2022, *156* (July 2020), 102816.

Pfaff, Alexander, Juan Robalino, Robert Walker, Steven Aldrich, Marcellus Caldas, Eustaquio Reis, Stephen Perz, Claudio Bohrer, Eugenio Arima, William Laurance et al., "Road investments, spatial spillovers, and deforestation in the Brazilian Amazon," *Journal of Regional Science*, 2007, *47* (1), 109–123.

Pfaff, Alexander S.P., "What drives deforestation in the Brazilian Amazon? Evidence from satellite and socioeconomic data," *Journal of Environmental Economics and Management*, 1999, *37* (1), 26–43.

Porteous, Obie, "Agricultural trade and adaptation to climate change in Sub-Saharan Africa," *Available at SSRN 4777962*, 2024.

Potapov, Peter, Svetlana Turubanova, Matthew C Hansen, Alexandra Tyukavina, Viviana Zalles, Ahmad Khan, Xiao-Peng Song, Amy Pickens, Quan Shen, and Jocelyn Cortez, "Global maps of cropland extent and change show accelerated cropland expansion in the twenty-first century," *Nature Food*, 2022, *3* (1), 19–28.

Redding, Stephen J. and Esteban Rossi-Hansberg, "Quantitative Spatial Economics," *Annual Review of Economics*, 2017, *9* (1), 21–58.

Restrepo, Veronica Salazar and Gabriel Leite-Mariante, "Does Conservation Work in General Equilibrium," 2023. Manuscript.

Schiavina, Marcello, Sergio Freire, Alessandra Carioli, and Kytt MacManus, "GHS-POP R2023A - GHS population grid multitemporal (1975–2030)," 2023.

Schultz, Theodore W, *The Economic Organization of Agriculture*, McGraw Hill, 1953.

Schwarzmueller, Florian and Thomas Kastner, "Agricultural trade and its impact on cropland use and the global loss of species' habitats," *Sustainability Science*, 2022, *17* (6), 2363–2377.

Singh, Chandrakant and U. Martin Persson, "Global patterns of commodity-driven deforestation and associated carbon emissions," *Nature Food*, 2026, *7* (2), 138–151.

Sotelo, Sebastian, "Domestic trade frictions and agriculture," *Journal of Political Economy*, 2020, *128* (7), 2690–2738.

Souza-Rodrigues, Eduardo, "Deforestation in the Amazon: A unified framework for estimation and policy analysis," *The Review of Economic Studies*, 2019, *86* (6), 2713–2744.

Sundberg, Ralph and Erik Melander, "Introducing the UCDP Georeferenced Event Dataset," *Journal of Peace Research*, 2013, *50* (4), 523–532.

Tombe, Trevor, "The Missing Food Problem: Trade, Agriculture, and International Productivity Differences," *American Economic Journal: Macroeconomics*, July 2015, *7* (3), 226–58.

United Nations Statistics Division, "UN Comtrade Database: International Trade Statistics," https://comtradeplus.un.org/ 2026. Accessed: 2026-08-07.

Vancutsem, C., F. Achard, J.-F. Pekel, G. Vieilledent, S. Carboni, D. Simonetti, J. Gallego, L. E. O. C. Aragão, and R. Nasi, "Long-term (1990–2019) monitoring of forest cover changes in the humid tropics," *Science Advances*, 2021, *7* (10), eabe1603.

Vonderschmidt, Alexander, Bartlomiej Arendarczyk, Lindsay M. Jaacks, Alexandra L. Bellows, and Peter Alexander, "Analysis combining the multiple FAO food balance sheet datasets needs careful treatment," *The Lancet Planetary Health*, 2024, *8* (2), e69–e71.

Wollburg, Philip, Thomas Bentze, Yuchen Lu, Christopher Udry, and Douglas Gollin, "Agricultural productivity growth in Africa: new evidence from microdata," *World Bank Economic Review*, 2023, *37* (1), 123–147.

World Bank, "Commodity Markets Outlook: Pink Sheet Data," World Bank Commodity Price Data 2024.

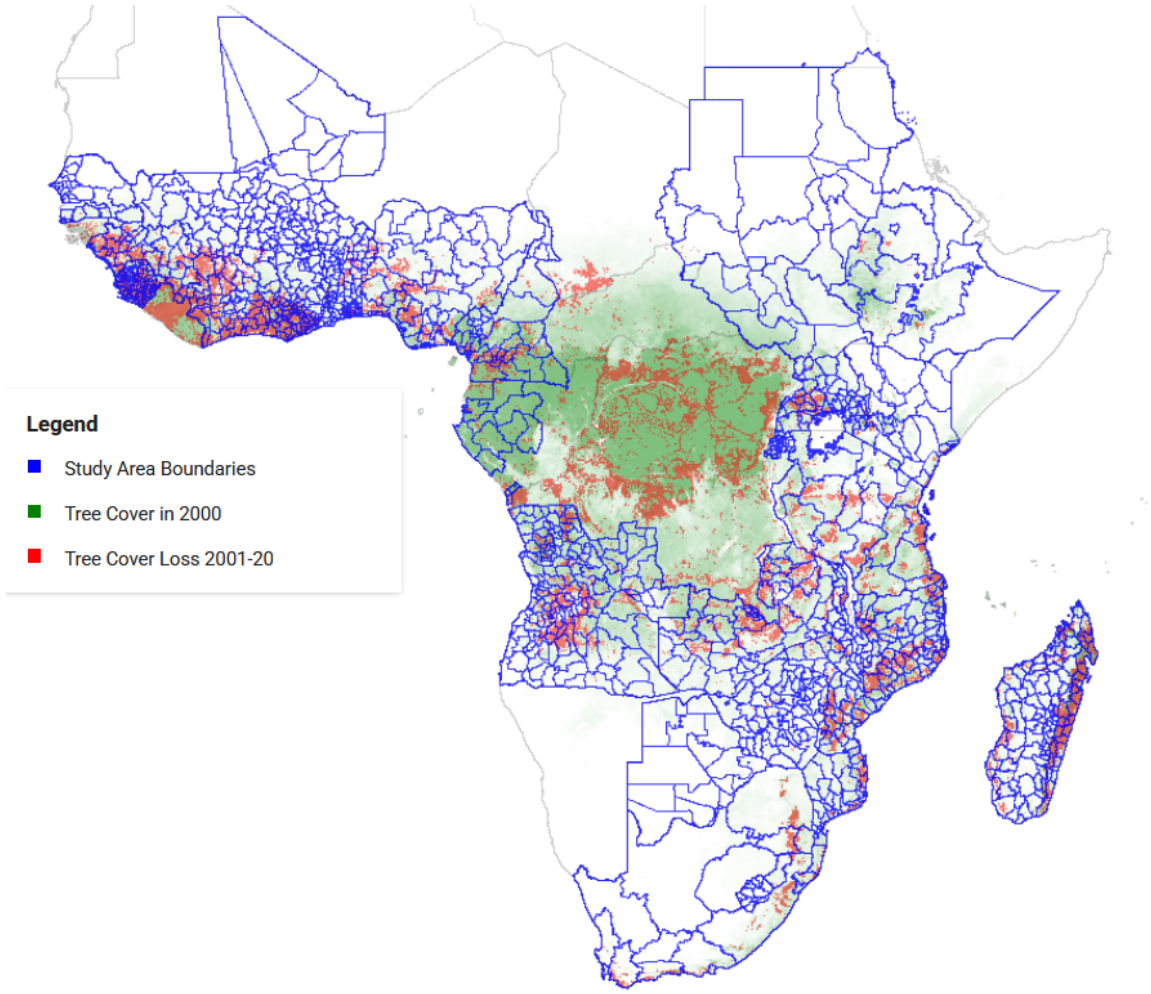
_ , "World Development Indicators," 2026. Accessed: 2026-08-07.

_ , "Living Standards Measurement Study – Integrated Surveys on Agriculture (LSMS-ISA)," World Bank.

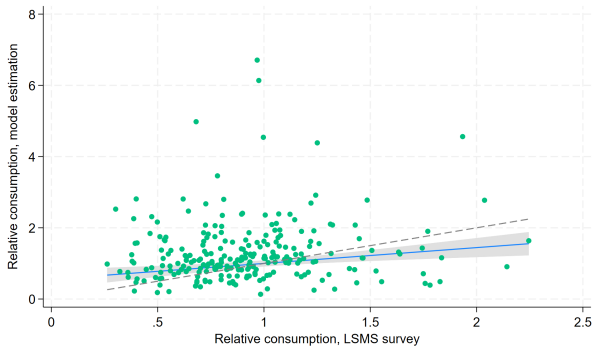
Wren-Lewis, Liam, Luis Becerra-Valbuena, and Kenneth Houngbedji, "Formalizing land rights can reduce forest loss: Experimental evidence from Benin," *Science Advances*, 2020, *6* (26), 1–9.

Figures and Tables

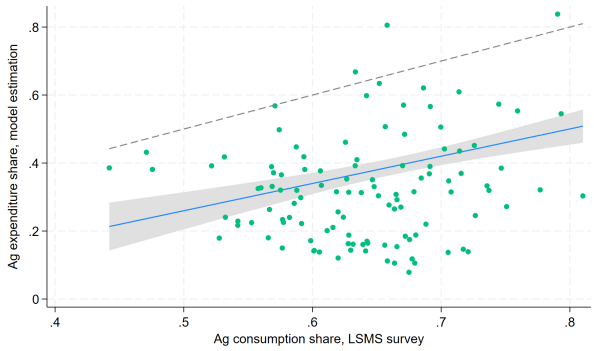
Figure 1: Tree Cover and Deforestation in Sub-Saharan Africa and in our Sample



Notes: The map represents in green tree cover in 2000, in red tree cover loss between 2001 and 2020, and in blue the boundaries of the countries that meet the data requirement for the study.



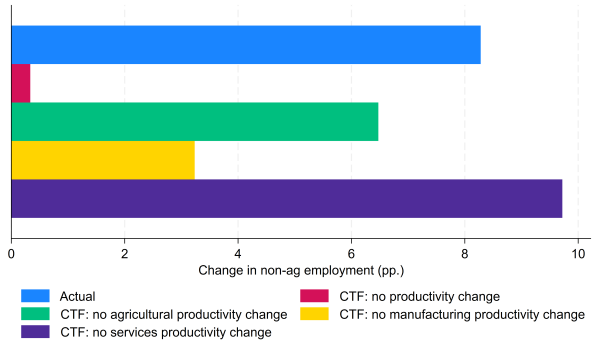
(a) Comparing estimated to surveyed real consumption



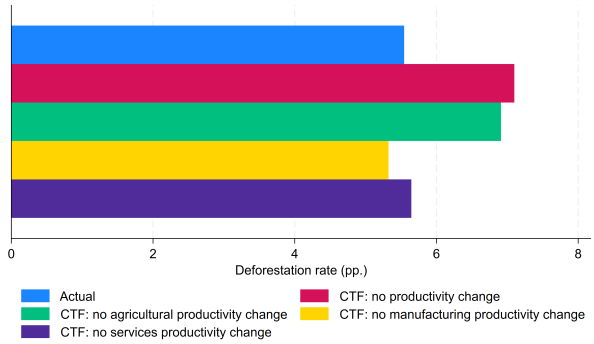
(b) Comparing estimated agricultural expenditure shares to surveyed consumption shares

Figure 2: Comparing estimates to household survey data

Notes: Panel (a) plots real consumption as estimated in the model against real adult-equivalent consumption as measured using household surveys, taking the mean consumption in each geographical unit and dividing by the mean across units. Panel (b) plots the expenditure share on agriculture as estimated in the model against the expenditure share on agriculture as measured using household surveys. In both panels, we plot one point for each region, along with a line of best fit weighted by population, and include a 95 % confidence interval. The LSMS survey closest to 2010 is used and compared to our estimates for 2010: panel (a) uses eight countries (Ethiopia, Ghana, Malawi, Mali, Nigeria, South Africa, Tanzania, and Uganda), while panel (b) uses four countries (Ethiopia, Malawi, Mali, and Nigeria).



(a) Change in non-agricultural employment

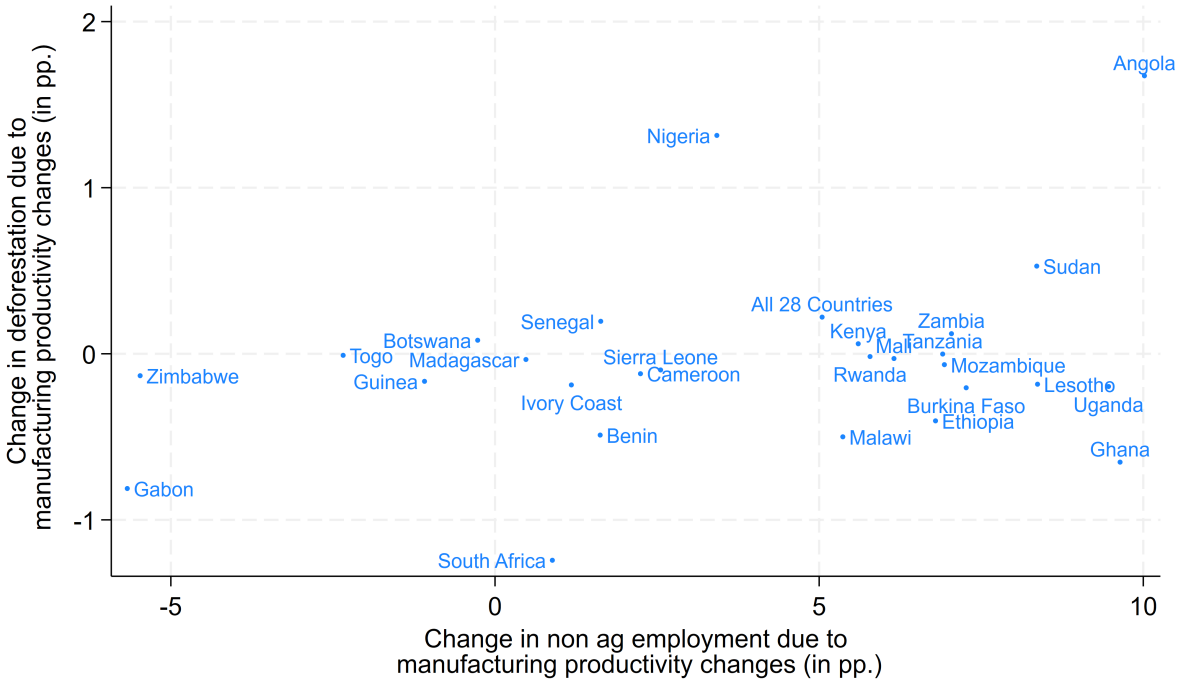


(b) Deforestation

Figure 3: Actual and counterfactual rates of structural transformation and deforestation

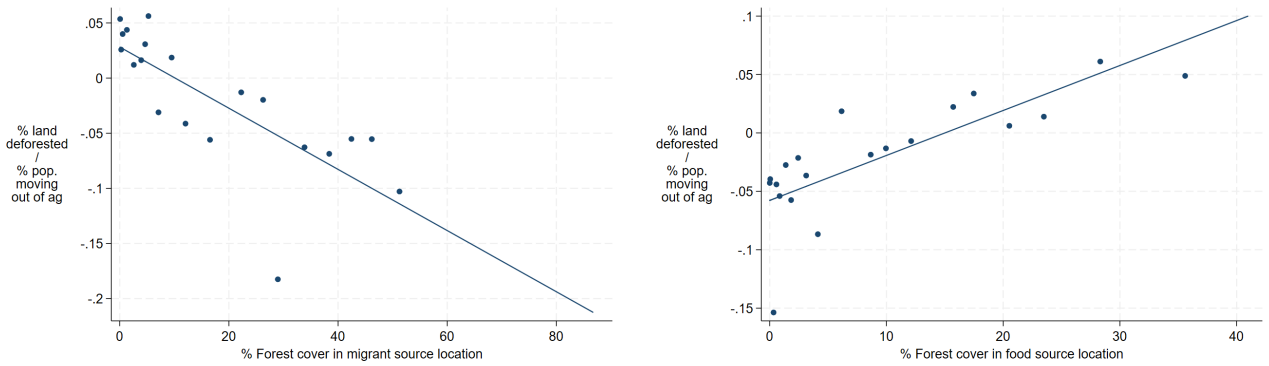
Notes: Panel (a) shows the change in non-agricultural employment rates between 2000 and 2010 under alternative scenarios, while panel (b) shows cumulative deforestation between 2010 and 2015 as a share of total land area. The first bar (in blue) is as measured in the data, while the other bars are estimates of what would have occurred under varying counterfactuals. In the second bar (in red), we estimate what would have happened had productivity in all sectors remained at levels estimated in 2000. Remaining bars then each estimate what would have happened had productivity in one sector remained at 2000 levels, while other sectors had moved to their estimated 2010 levels.

Figure 4: Heterogeneous effects of manufacturing productivity growth on deforestation and structural change



Notes: This figure is based on the model counterfactual in which manufacturing / mining productivity is moved to its 2000 level, but all other parameters are kept at their 2010 levels. Each point shows the difference between the outcome at endline and in this counterfactual - i.e. we can interpret points as the impact of manufacturing / mining productivity moving from 2000 to 2010 levels. Mauritius is excluded because manufacturing growth led to a very large decrease in deforestation (-3pp). Mauritius only represents a very small share of the forest cover in our sample (0.02%).

Figure 5: Impacts of local shocks depending on forest closeness

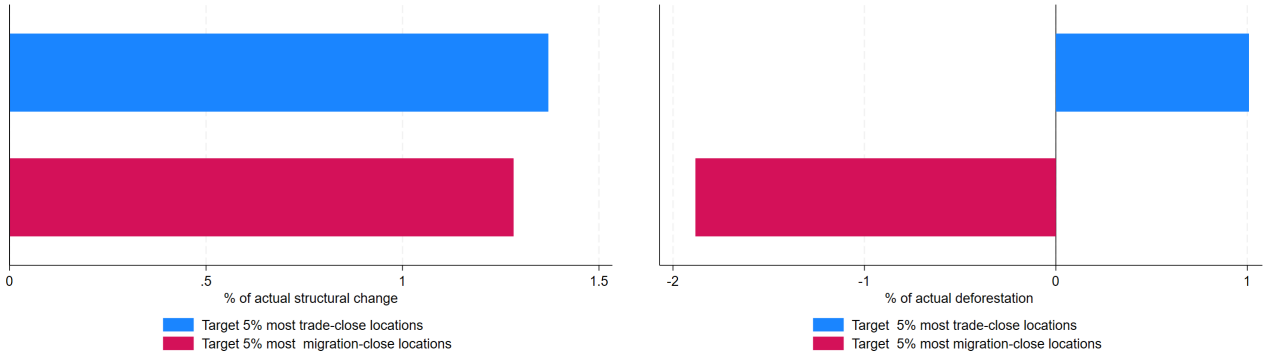


(a) Impacts by 'migration-closeness'

(b) Impacts by 'trade-closeness'

Notes: This figure shows binned scatter plots of the impacts of location specific shocks against the shock location in relation to forests. Each counterfactual sets manufacturing and mining productivity in a single location to the value estimated in 2000, while all remaining parameters are kept at their 2010 values. The change in land deforested (nationally) divided by the change in the non-agricultural population share (nationally) is the variable plotted on the y-axis. In subfigure (a), the x-axis plots the average forest cover in all units in the country, weighted by the share of migrants that the shocked area sources from each unit. In subfigure (b), the x-axis plots the average forest cover in all units in the country, weighted by the share of consumed agricultural goods that the shocked area sources from each unit.

Figure 6: Results of hypothetical productivity increase



(a) Impacts on non-agricultural employment share

(b) Impacts on share of land deforested

Notes: This figure reports results from two different counterfactuals which each increase manufacturing and mining productivity by 100 % in 5% of the locations in each country. Locations are first ranked by the difference between forest cover in their 'migration-close' and 'trade-close' locations, and the productivity increase is then applied separately to top and the bottom 5% of this ranking. Outcomes are measured relative to the endline equilibrium, allowing prices, employment and land allocation to adjust endogenously. Panel (a) reports the change in non-agricultural employment as a share of total employment, while panel (b) reports the change in deforested land as a share of total land area. Results are aggregated across all countries in the estimation sample.

Table 1: Non-agricultural employment growth and deforestation at the local level

	Baseline share of area forested		Share of area deforested	
	(1)	(2)	(3)	(4)
Share of non-ag employment	-0.380***		0.002	
	(0.090)		(0.001)	
Growth in non-ag employment		0.027		-0.000
		(0.038)		(0.001)
Observations	831	831	779	779
Outcome Mean in %	34.04	34.04	.19	.19

The unit of observation is an administrative unit. The sample includes all 831 units across the 18 countries that have two censuses. In Column 1 and 2 the dependent variable is the share of land area covered with forests. In Columns 3 and 4 it is the annual rate of deforestation, computed as a share of total land area. Regressions in Columns 3 and 4 are weighted by baseline forest cover. Robust standard errors in parentheses. All regressions include country fixed effects. $^{*}p < .1,^{**}p < .05,^{***}p < .01$

Table 2: Effect of Non-Agricultural Employment Growth on Deforestation via Migration and Trade

	Emigration over land area			Share of area deforested		
	(1)	(2)	(3)	(4)	(5)	(6)
Migration S-S	1.001***		1.250***	-0.034**		-0.075***
	(0.369)		(0.407)	(0.016)		(0.021)
Trade S-S		0.614***	-0.350***		0.044***	0.053***
		(0.194)	(0.128)		(0.013)	(0.015)
Observations	680	680	680	779	779	779
Outcome Mean in %	.44	.44	.44	.19	.19	.19

The unit of observation is an administrative unit. The sample includes 779 units with non-zero forest cover in the first census year across the 18 countries that have two censuses. In Columns 1-3 the dependent variable is the average annual emigration rate divided by total land area. In Columns 3-5 it is the annual rate of deforestation as a share of total land area. In Columns 1-3 the sample is reduced since we cannot compute the emigration rate in Ghana. All regressions are weighted by baseline forest cover. Robust standard errors in parentheses. Regressions include the sum of migration shares across all destinations and the sum of trade shares across all destinations for each origin as controls. All models include country fixed effects. $^{*}p < .1,^{**}p < .05,^{***}p < .01$

Table 3: Placebo Effect of Non-Agricultural Employment Growth on Prior Deforestation

	Prior deforestation			
	(1)	(2)	(3)	(4)
Growth in non-ag employment	0.003			
	(0.002)			
Migration S-S		0.009		0.022
		(0.026)		(0.037)
Trade S-S			-0.003	-0.012
			(0.004)	(0.011)
Observations	741	741	741	741
Outcome Mean in %	.19	.19	.19	.19

The unit of observation is an administrative unit. The dependent variable is annual deforestation between 1992 and the first census year based on ESA CCI data. The sample includes 741 units with non-zero forest cover in the first census year across the 18 countries that have two censuses minus Cameroon whose first census predates ESA CCI deforestation data. All regressions are weighted by baseline forest cover. Robust standard errors in parentheses. Regressions include the sum of migration shares across all destinations and the sum of trade shares across all destinations for each origin as controls. All models include country fixed effects. $^{*}p < .1,^{**} p < .05,^{***} p < .01$

Table 4: Effect of non-agricultural employment growth on Deforestation via migration and internal trade in countries more or less open to international trade

	Emigration		Deforestation	
	(1)	(2)	(3)	(4)
	Less Open	More Open	Less Open	More Open
Migration S-S	1.210***	2.091	-0.080***	-0.105
	(0.414)	(1.624)	(0.022)	(0.093)
Trade S-S	-0.305***	3.702*	0.064***	0.018
	(0.117)	(1.938)	(0.018)	(0.020)
Observations	356	324	356	423
P-value Trade S-S Low=High		.04		.06
Outcome Mean in %	.4	.62	.17	.24

The unit of observation is an administrative unit. The sample includes all 779 units with non-zero forest cover in the first census year across the 18 countries that have two censuses. In Columns 1 and 2 the dependent variable is the average annual emigration rate divided by total land area. In Columns 3 and 4 it is the annual rate of deforestation as a share of total land area. Columns 1 and 3 include countries with below median openness to international trade, measured as agricultural imports plus exports divided by agricultural GDP. Columns 2 and 4 include countries with above median openness to international trade. In Column 2 the sample is reduced since we cannot compute the emigration rate in Ghana. Regressions weighted by baseline tree cover. Robust standard errors in parentheses. Regressions include the sum of migration shares across all destinations and the sum of trade shares across all destinations for each origin as controls. All models include country fixed effects. $^{*}p < .1,^{**} p < .05,^{***} p < .01$

Table 5: Parameters from the literature

Parameter	Value	Source
θ	1.7	Sotelo (2020)
ζ	2.5	Sotelo (2020)
η	5.4	Costinot et al. (2016)
α_A, α_S	0.7	Gollin (2002)
α_M	0.34	Dinh and Clarke (2012)
κ	3.2	Bryan and Morten (2019)
γ	4	Lagakos and Waugh (2013)
σ	0.48	Comin et al. (2021)
ε_A	0.2	Comin et al. (2021)
ε_S	1.37	Comin et al. (2021)

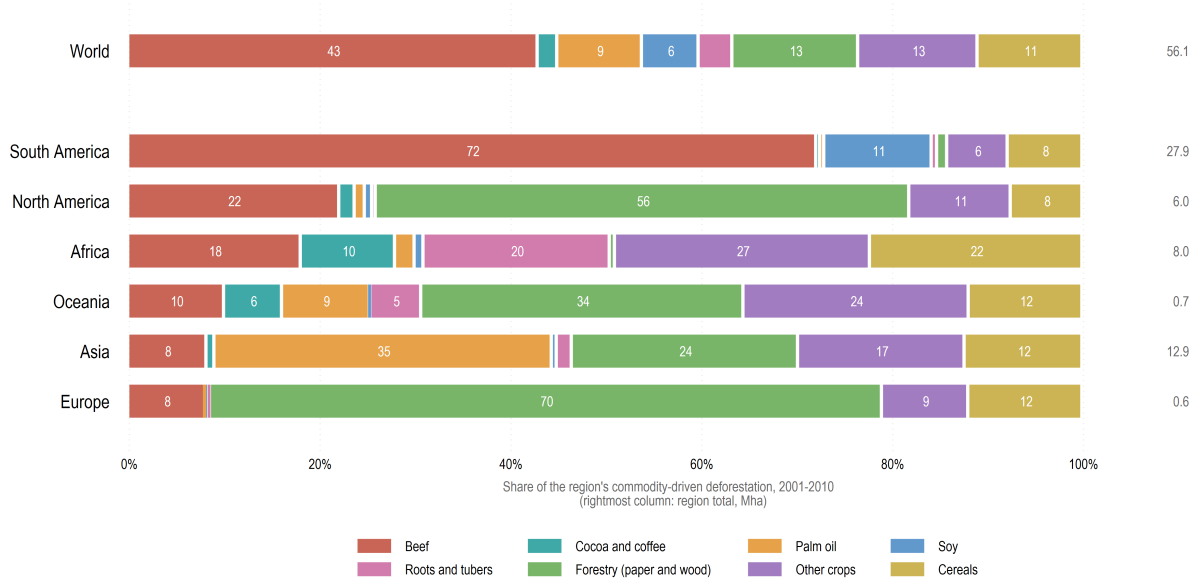
Table 6: Effect of Non-Agricultural Employment Growth on Deforestation via Migration and Trade based on the Model Counterfactuals

	Share of area deforested		
	(1) All	(2) Less Open	(3) More Open
Migration S-S	-0.048**	-0.064	-0.022
	(0.023)	(0.041)	(0.016)
Trade S-S	0.050**	0.052**	0.019
	(0.022)	(0.024)	(0.012)
Observations	1186	713	473
Outcome Mean in %	-.16	-.18	-.13

The unit of observation is an administrative unit. The sample includes all 1186 units with non-zero forest cover in the first census year across all countries. Regressions weighted by tree cover . Robust standard errors in parentheses. *Migration S-S* is the migration shift-share, and *Trade S-S* is the trade shift share. The shift is the difference in the share of non-agricultural agricultural employment between 2010 data and the model counterfactual where all sectoral productivities are held at their 2000 values. The migration shares and the trade shares are defined as in the reduced form. Regressions that include the migration shift-share also include the sum of migration shares for that origin across all destinations. Regressions that include the trade shift-share also include the sum of the trade shares. All models include country fixed effects. $^{*}p < .1,^{**}p < .05,^{***}p < .01$

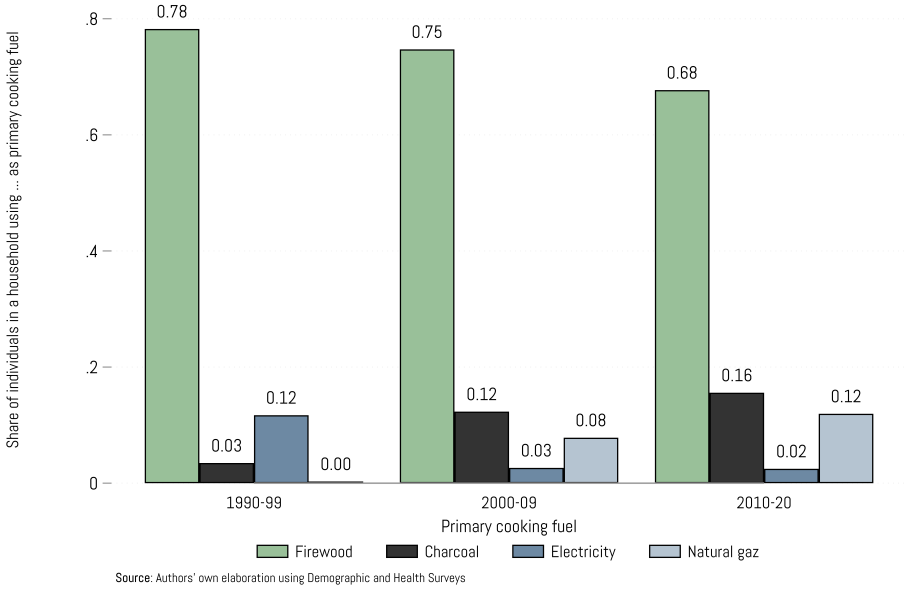
A Appendix Figures and Tables

Figure A.1: Composition of commodity-driven deforestation, by World region, 2001–2010



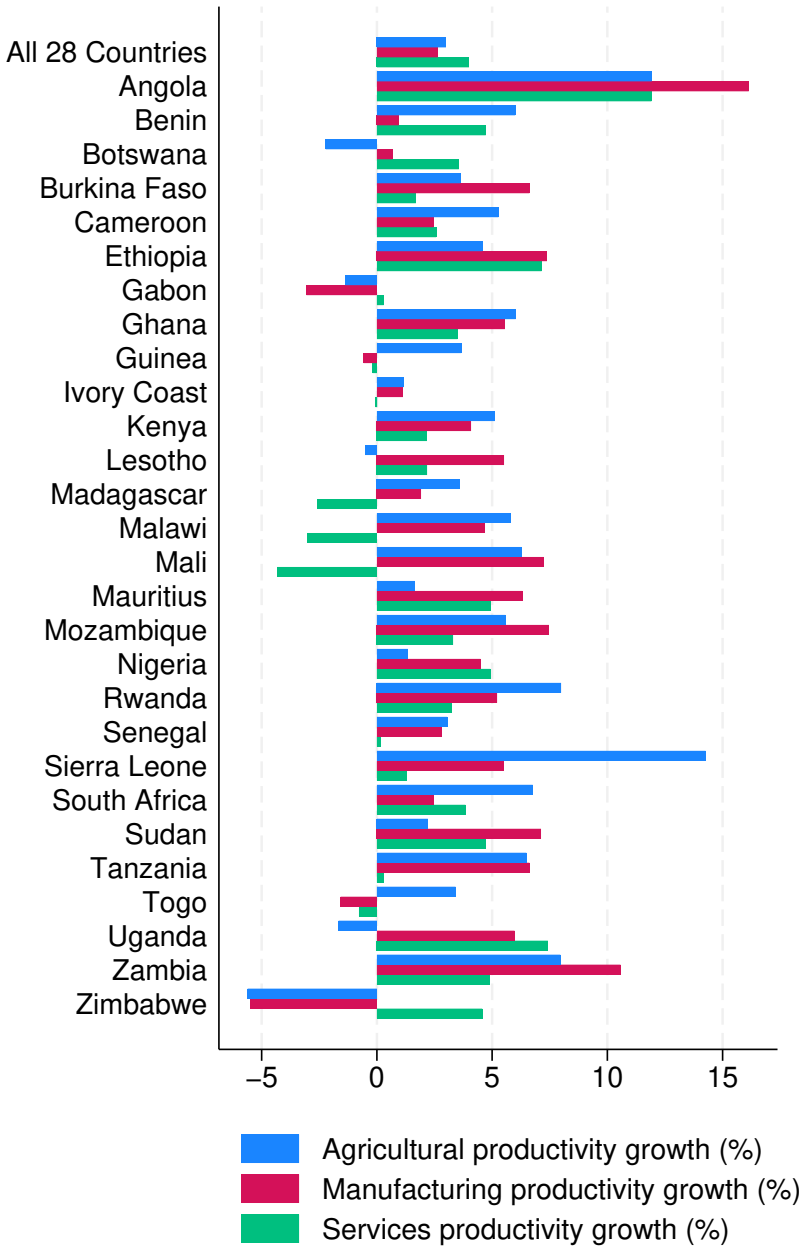
Notes: Each bar decomposes a region's commodity-driven deforestation over 2001–2010 into the eight commodity groups distinguished by the source, normalized to 100 percent of the region's total. The number to the right of each bar reports the region's total commodity-driven deforestation in millions of hectares. Regions are ordered by the share attributed to beef. The six regions are mutually exclusive and exhaustive and, commodity by commodity, sum to the World total. The World row is shown for reference and represents the sum of the six regions, not a seventh region. Segment shares below 4 percent are not labeled. Hectares are attributed to the commodity produced on the cleared land following deforestation; indirect and displaced land-use change is not captured. Source: Authors' calculations from Singh and Persson (2026), using the DeDuCE database as distributed by Our World in Data.

Figure A.2: Change in cooking fuel in Sub-Saharan Africa, 2010 to 2020



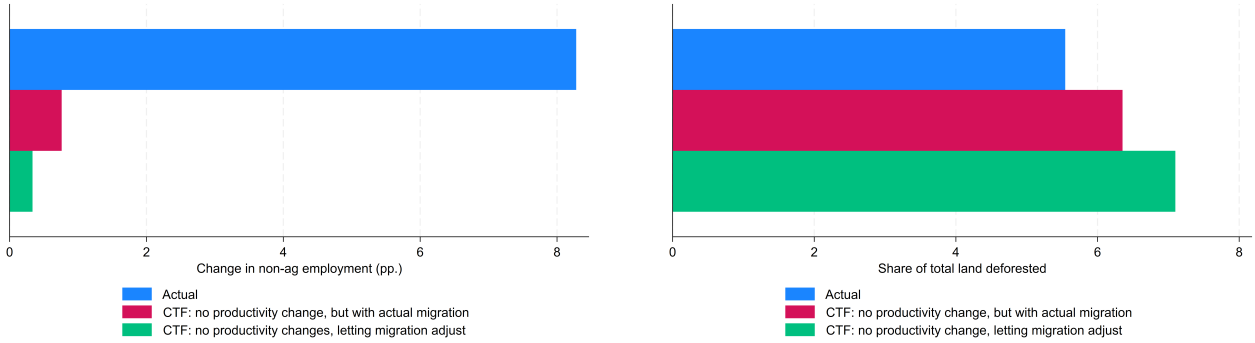
Notes: This figure shows the share of Sub-Saharan African respondents in the Demographic and Health Surveys that declare using firewood, charcoal, natural gaz or other fuel as their primary cooking fuel by decade.

Figure A.3: Estimated Average Annual Productivity Changes, 2000 to 2010



Notes: This figure plots the estimated average annual changes in productivity in agriculture (Z_i^A), manufacturing (Z_i^M) and services (Z_i^S) in each country, with values below 1 representing a decrease and values above 1 an increase. The changes are calculated by estimating the structural model first with data from the year 2000 values and then with data from 2010. For agricultural productivity, averages are first taken across crop weighted by baseline production. Averages are then taken across districts weighted by baseline production in the respective sector.

Figure A.4: Actual and counterfactual structural change and deforestation, with and without migration

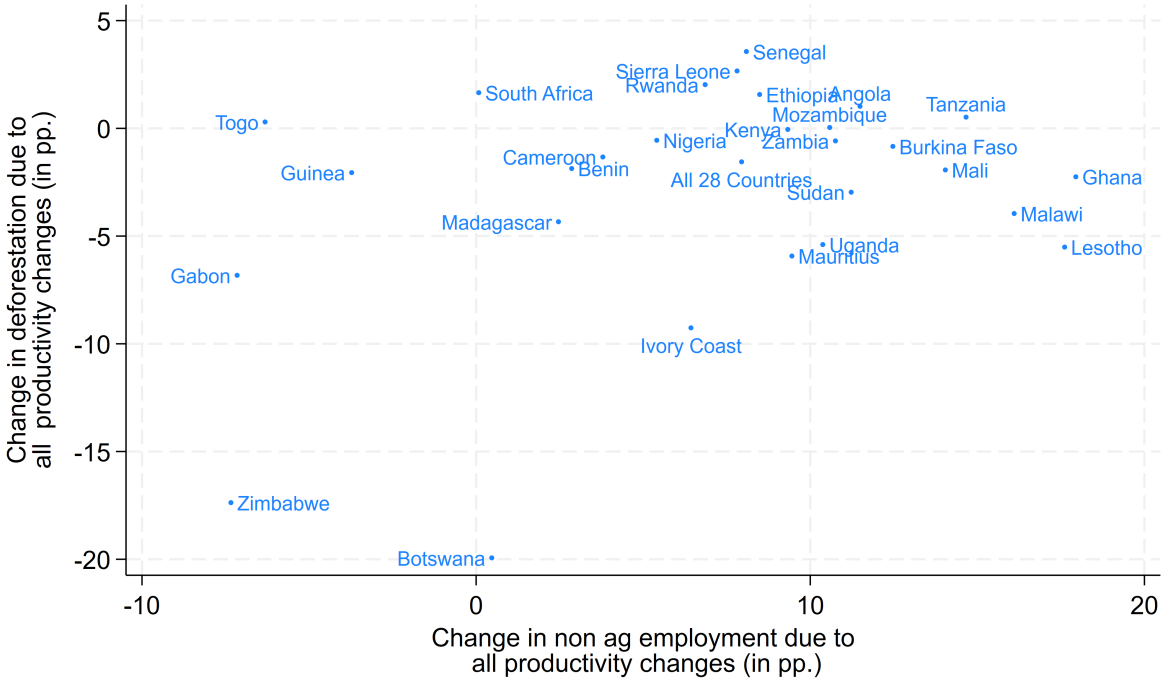


(a) Change in non-agricultural employment

(b) Deforestation

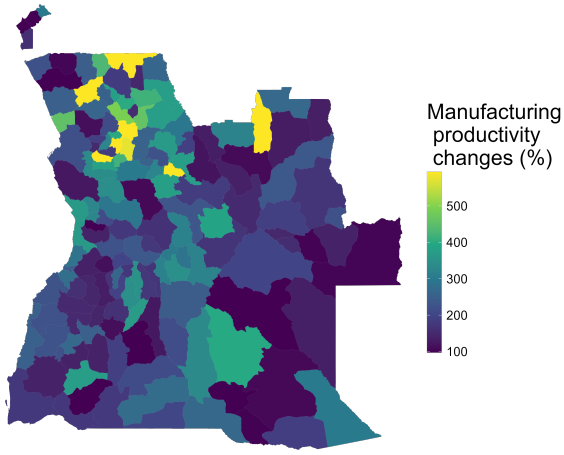
Notes: This figure shows the effect of productivity changes and migration on structural change (a) and deforestation (b). The blue bar is the actual outcome. The red bar shows the outcome when all productivities are at 2000 levels, but the total population in each district is assumed to be at the 2010 level. The green bar shows the outcome when all productivities are at 2000 levels and migration is allowed to adjust to them

Figure A.5: Heterogeneous effects of productivity growth in all sectors on deforestation and structural change

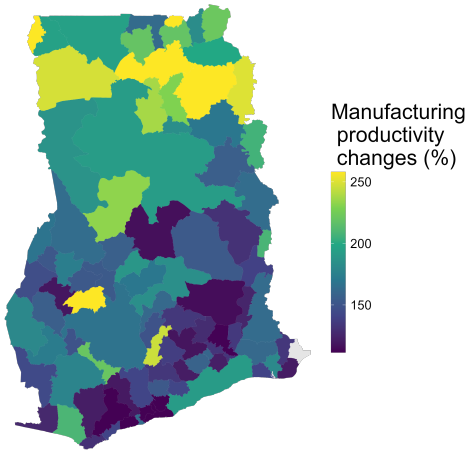


Notes: This figure is based on the model counterfactual in which productivity in the three sectors is moved to its 2000 level, but all other parameters are kept at their 2010 levels. Each point shows the difference between the outcome at endline and in this counterfactual - i.e. we can interpret points as the impact of manufacturing / mining productivity moving from 2000 to 2010 levels.

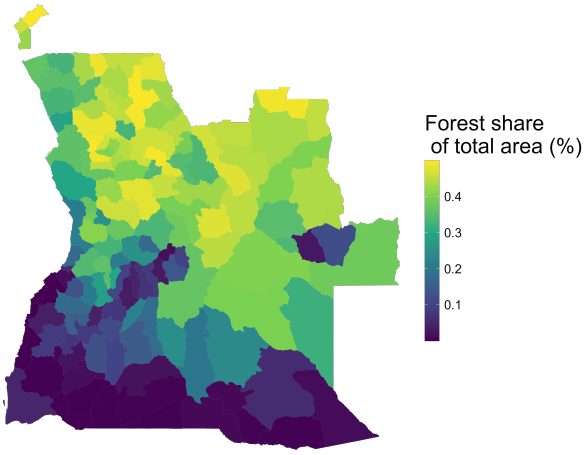
Figure A.6: The geography of forest cover and manufacturing and mining productivity growth in two sample countries: Angola and Ghana



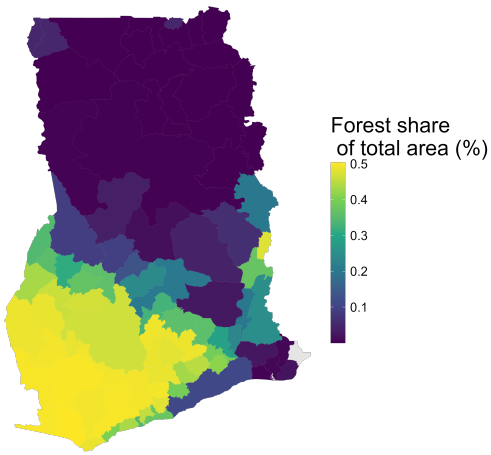
(a) Manufacturing and mining productivity growth in Angola



(b) Manufacturing and mining productivity growth in Ghana



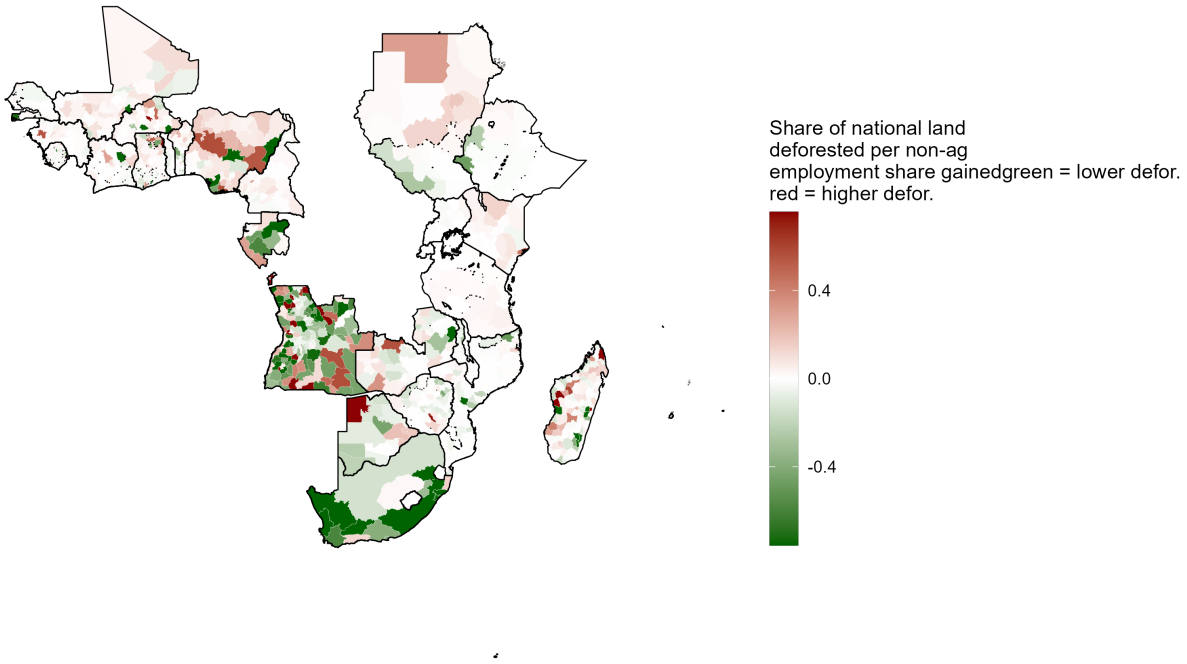
(c) Forest cover in Angola



(d) Forest cover in Ghana

Notes: This figure shows model estimates of manufacturing and mining productivity growth between 2000 and 2010 in Angola (a) and Ghana (b). It also shows forest cover in 2000 in Angola (a) and Ghana (b).

Figure A.7: Impacts of local productivity shocks on national deforestation



Notes: This map shows the estimated impact on national outcomes from changes in manufacturing and mining productivity in individual locations. In particular, each geographic unit is shaded corresponding to the result of a separate counterfactual estimation, whereby manufacturing and mining productivity productivity in that area is set to the value estimated in 2000, while all remaining parameters are kept at their 2010 values. Units are shaded according to the amount of extra deforestation per extra non-agricultural worker induced by the productivity being at its 2010 level rather then the 2000 level. A unit where the change in productivity growth increased non-agricultural employment while decreasing deforestation will therefore be shaded in green, while one with an increase in both non-agricultural employment and deforestation is shaded in red.

Table A.1: Sample Administrative Unit level

Country	Administrative Unit Level	Number of Units	Census Year	
			First	Second
Angola	2	161	2014	
Benin	2	77	2002	2013
Botswana	1	21	2001	2011
Burkina Faso	2	45	1996	2006
Cameroon	2	39	1987	2005
Ethiopia	1	11	2014	
Gabon	1	9	1993	2013
Ghana	2	102	2000	2010
Guinea	1	33	1996	2014
Ivory Coast	2	50	1998	
Kenya	2	35	1999	2009
Lesotho	1	10	1996	2006
Madagascar	2	114	2018	
Malawi	1	26	2008	
Mali	2	47	1998	2009
Mauritius	1	10	2000	2011
Mozambique	2	137	1997	2007
Nigeria	1	37	2012	
Rwanda	2	30	2012	
Senegal	2	27	2002	2013
Sierra Leone	2	126	2004	2015
South Africa	2	18	2001	2011
Sudan	1	25	2008	
Tanzania	1	23	2002	2012
Togo	2	37	2010	
Uganda	1	36	2002	2014
Zambia	2	55	2000	2010
Zimbabwe	2	88	2012	

Notes: The table reports the administrative unit level used in our analysis and the total number of units in the country at that level. The first and second census years are those that we used in the analysis. Although data is available at the admin 2 level for Mauritius, we use admin level 1 because most of the migration at the admin 2 level is occurring between districts within Port Louis, so the more aggregated data on the island is more representative of the type of migration that we are interested in.

Table A.2: Descriptive statistics for the main outcome variables of the reduced form analysis (sample of countries with two censuses)

Country	Emigration over land area	Share of area deforested	Share of area deforested (TMF)	Δ share cropland	N
BEN	3.73	0.64	0.21	-0.94	71
BWA	0.87	0.00	.	0.04	17
CMR	3.53	0.74	0.34	.	38
GAB	0.31	0.68	0.12	0.00	9
GHA	.	4.34	1.22	0.39	99
GIN	3.29	2.45	1.09	0.11	33
KEN	34.27	1.09	0.51	2.79	35
LSO	3.76	0.00	.	-0.52	10
MLI	1.03	0.02	0.00	0.96	20
MOZ	1.61	2.10	0.01	0.51	137
MUS	30.97	1.10	0.96	0.00	9
MWI	33.69	2.62	1.16	3.28	26
SEN	7.25	0.34	0.22	0.17	19
SLE	9.79	10.00	5.49	0.06	124
TZA	7.72	1.80	0.08	1.78	23
UGA	8.69	2.74	3.03	6.86	36
ZAF	2.34	1.66	.	0.14	18
ZMB	2.11	1.34	0.03	0.62	55
All Countries	4.45	1.86	0.54	1.03	779

Notes: The table reports the average of the outcomes considered in the reduced form analysis for the sample of countries with two censuses. Ghana has missing migration information in the second census. Botswana, Lesotho and South Africa are not in the tropical zone.

Table A.3: Annual Migration by Measure and Census

measure	outcome	census year	population (millions) (1)	migration (thousands) (2)	migration rate (3)
Angola					
5 years: 2009 - 2014	emigration	2014	11.992	146.02	1.22
Benin					
Duration: 1992 - 2001	migration shares	2002	3.388	57.71	1.70
Duration: 2003 - 2012	emigration	2013	5.095	63.10	1.24
Birth	none	2002	3.388	139.01	4.10
Birth	none	2013	5.095	186.43	3.66
Botswana					
5 years: 1996 - 2001	migration shares	2001	0.981	51.21	5.22
5 years: 2006 - 2011	emigration	2011	1.260	66.71	5.29
Birth	none	2001	0.981	49.58	5.06
Birth	none	2011	1.260	64.78	5.14
Burkina Faso					
1 year: 1995 - 1996	none	1996	5.346	368.73	6.90
1 year: 2005 - 2006	emigration	2006	7.075	190.68	2.70
Birth	migration shares	1996	5.346	172.15	3.22
Birth	none	2006	7.075	219.92	3.11
Ghana					
5 years: 1996 - 2000	migration shares	2000	10.087	138.59	1.37
Guinea					
Duration: 1986 - 1995	migration shares	1996	3.646	102.22	2.80
Duration: 2004 - 2013	emigration	2014	5.393	83.05	1.54
Birth	none	1996	3.646	132.32	3.63
Birth	none	2014	5.393	163.64	3.03
Kenya					
Duration: 1989 - 1998	migration shares	1999	14.979	23.02	0.15
1 year: 2008 - 2009	emigration	2009	20.543	758.13	3.69
Birth	none	1999	14.848	482.24	3.25
Birth	none	2009	20.396	626.88	3.07
Lesotho					
10 years: 1987 - 1996	migration shares	1996	1.078	11.51	1.07
10 years: 1997 - 2006	emigration	2006	1.088	12.39	1.14

...Continued on next page

Annual Migration by Measure and Census – *continued from previous page*

measure	outcome	census year	population (millions) (1)	migration (thousands) (2)	migration rate (3)
Birth	none	1996	1.078	20.06	1.86
Birth	none	2006	1.088	19.31	1.77
Madagascar					
10 years: 2008 - 2018	emigration	2018	14.391	115.70	0.80
Birth	none	2018	14.381	278.98	1.94
Mali					
Duration: 1988 - 1997	migration shares	1998	4.958	30.68	0.62
Duration: 1999 - 2008	emigration	2009	7.021	90.10	1.28
Birth	none	1998	4.958	119.92	2.42
Birth	none	2009	6.941	297.66	4.29
Mauritius					
5 years: 1996 - 2000	migration shares	2000	0.820	12.61	1.54
5 years: 2007 - 2011	emigration	2011	0.871	9.18	1.05
Mozambique					
5 years: 1992 - 1997	none	1997	8.171	347.60	4.25
5 years: 2002 - 2007	emigration	2007	10.240	221.60	2.16
Birth	migration shares	1997	8.171	336.07	4.11
Birth	none	2007	10.240	390.98	3.82
Senegal					
5 years: 1998 - 2002	migration shares	2002	5.356	101.05	1.89
5 years: 2009 - 2013	emigration	2013	7.175	233.29	3.25
Birth	none	2002	5.356	217.58	4.06
Birth	none	2013	7.175	271.21	3.78
Sierra Leone					
14 years: 1990 - 2004	none	2004	2.652	56.45	2.13
5 years: 2011 - 2015	emigration	2015	3.966	132.76	3.35
Birth	migration shares	2004	2.652	145.31	5.48
Birth	none	2015	3.966	207.50	5.23
South Africa					
Duration: 1996 - 2000	migration shares	2001	28.220	193.52	0.69
Duration: 2001 - 2010	emigration	2011	33.904	147.47	0.43
Tanzania					

...Continued on next page

Annual Migration by Measure and Census – *continued from previous page*

measure	outcome	census year	population (millions) (1)	migration (thousands) (2)	migration rate (3)
1 year: 2001 - 2002	none	2002	17.178	635.85	3.70
1 year: 2011 - 2012	emigration	2012	22.898	933.78	4.08
Birth	migration shares	2002	17.178	383.65	2.23
Birth	none	2012	22.898	569.20	2.49
Uganda					
10 years: 1993 - 2002	migration shares	2002	12.137	116.95	0.96
10 years: 2005 - 2014	emigration	2014	17.540	175.21	1.00
Birth	none	2002	12.137	344.70	2.84
Zambia					
1 year: 1999 - 2000	none	2000	5.189	335.46	6.46
1 year: 2009 - 2010	emigration	2010	6.880	693.58	10.08
Birth	migration shares	2000	5.189	267.44	5.15
Birth	none	2010	6.880	374.78	5.45

Notes: Reporting the prime-aged population in millions, average annual number of migrants and the implied annual migration rate for each reported measure. Migration since birth is annualized by dividing each observed migration by the individual's age - 14. That is, for each person who has migrated since birth, we assume that the migration happened once the individual became prime-aged and that there was a single migration episode in that period. This transformation is for expositional purposes only. All migration since birth episodes are used as shares in the analysis.

Table A.4: Effect of Non-Agricultural Employment Growth on Deforestation via Migration and Trade including Controls

	Share of area deforested			
	(1)	(2)	(3)	(4)
Migration S-S	-0.075***	-0.077***	-0.078***	-0.080***
	(0.021)	(0.021)	(0.021)	(0.021)
Trade S-S	0.054***	0.054***	0.052***	0.053***
	(0.015)	(0.015)	(0.015)	(0.014)
Local Shift Control	Yes	No	No	Yes
Geographical Controls	No	Yes	No	Yes
Conflict Control	No	No	Yes	Yes
Outcome Mean in %	.19	.19	.19	.19
Observations	779	779	779	779

The unit of observation is an administrative unit. The sample includes all 779 units with non-zero forest cover in the first census year across the 18 countries that have two censuses. In Column 1, we control for the change in non-agricultural employment locally. In Column 2 we control for average temperatures, average rainfall and ruggedness. In Column 3 we control for the number of conflict events (ACLED). In Column 4 we include all controls together. All regressions are weighted by baseline forest cover. Robust standard errors in parentheses. Regressions include the sum of migration shares across all destinations and the sum of trade shares across all destinations for each origin as controls. All models include country fixed effects. All models include country fixed effects. $^{*}p < .1,^{**} p < .05,^{***} p < .01$

Table A.5: Placebo Effect of Non-Agricultural Employment Growth on Prior Deforestation measured using ESA CCI and Hansen et al. (2013) data

	Prior deforestation			
	(1)	(2)	(3)	(4)
Growth in non-ag employment	0.002			
	(0.002)			
Migration S-S		0.005		0.017
		(0.026)		(0.035)
Trade S-S			-0.003	-0.011
			(0.004)	(0.010)
Observations	741	741	741	741
Outcome Mean in %	.19	.19	.19	.19

The unit of observation is an administrative unit. The dependent variable is annual deforestation between 1992 and the first census year based on ESA CCI data until 2000 and, for countries where the first census is after 2000, we use Hansen et al. (2013) data after that. The sample includes 741 units with non-zero forest cover in the first census year across the 18 countries that have two censuses minus Cameroon whose first census predates ESA CCI deforestation data. All regressions are weighted by baseline forest cover. Robust standard errors in parentheses. Regressions include the sum of migration shares across all destinations and the sum of trade shares across all destinations for each origin as controls. All models include country fixed effects. $^{*}p < .1,^{**}p < .05,^{***}p < .01$

Table A.6: Effect of Non-Agricultural Employment Growth on Deforestation via Migration and Trade using Alternative Measures of Land Use

	Deforestation (TMF)	Δ cropland (MODIS)	Deforestation (Hansen)			
	(1)	(2)	(3) 10%	(4) 50%	(5) 70%	(6) 90%
Migration S-S	-0.040**	-0.107*	-0.040**	-0.079***	-0.038*	-0.037
	(0.016)	(0.055)	(0.019)	(0.023)	(0.023)	(0.023)
Trade S-S	0.030**	0.054**	0.026**	0.069***	0.052***	0.051***
	(0.012)	(0.023)	(0.012)	(0.016)	(0.019)	(0.019)
Outcome Mean in %	.06	.37	.21	.13	.05	.02
Observations	593	735	820	744	690	638

The unit of observation is an administrative unit. In Column 1, the dependent variable is deforestation according to Vancutsem et al. (2021)'s measure, which focuses on regions with tropical moist forest cover at baseline. In Column 2, the dependent variable is the change in cropland share from Potapov et al. (2022)'s land use classification which is not available for Cameroon, Madagascar, and Mauritius. In Columns 3-6 the dependent variables are alternative measures of deforestation based on Hansen et al. (2013) using different percentages of forest cover per pixel to define forested area. We use 30% in our main specification, whereas Columns 3, 4, 5, 6 use 10%, 50%, 70% and 90% respectively. The number of observations with non-zero forest cover changes based on the threshold used. All regressions are weighted by baseline forest cover. Robust standard errors in parentheses. Regressions include the sum of migration shares across all destinations and the sum of trade shares across all destinations for each origin as controls. All models include country fixed effects. All models include country fixed effects. $^{*}p < .1,^{**}p < .05,^{***}p < .01$

Table A.7: Effect of Non-Agricultural Employment Growth on Deforestation via Migration and Trade using alternative shift-shares.

	Duration in Trade Share		International Shares	
	(1) Emigration	(2) Deforestation	(3) Emigration	(4) Deforestation
Migration S-S	1.390***	-0.072***		
	(0.407)	(0.020)		
Trade S-S	-0.027**	0.002***		
	(0.011)	(0.001)		
Migration S-S			1.108***	-0.064***
			(0.407)	(0.019)
Trade S-S			-0.128*	0.030***
			(0.069)	(0.009)
Observations	680	779	680	779
Outcome Mean in %	.44	.19	.44	.19

The unit of observation is an administrative unit. The sample includes all 779 units with non-zero forest cover in the first census year across the 18 countries that have two censuses. In Columns 1 and 2 the shifts are the same as in the main specification but the trade shift-share uses travel duration rather than distance to compute market access. In Column 3 and 4 the shifts include non-agricultural employment growth not only within the same country but also in other sample countries. All regressions are weighted by baseline forest cover. Robust standard errors in parentheses. Regressions include the sum of migration shares across all destinations and the sum of trade shares across all destinations for each origin as controls. All models include country fixed effects. All models include country fixed effects. $^{*}p < .1,^{**}p < .05,^{***}p < .01$

Table A.8: Effect of Non-Agricultural Employment Growth on Deforestation via Migration and Trade using Borusyak et al. (2022) inversion

	Emigration over land area		Share of area deforested	
	(1)	(2)	(3)	(4)
Migration S-S	1.001***		-0.034**	
	(0.220)		(0.016)	
Trade S-S		0.614		0.044***
		(0.411)		(0.013)
Observations	729	729	831	831

The unit of observation is an administrative unit. In Columns 1 and 2 the sample includes all 729 administrative units across 17 countries, since emigration information is missing for Ghana. In Columns 3 and 4 the sample includes all 831 administrative units in the 18 countries with two censuses. The sample sizes are larger than in Table 2 because they include all units (some with zero forest coverage). The inverted specification is weighted by baseline forest cover, it controls for the sum of migration shares across all destinations (Columns 1 and 3) and the sum of trade shares across all destinations (Columns 2 and 4). It also includes country fixed effects. $^{*}p < .1,^{**}p < .05,^{***}p < .01$

Table A.9: Impacts of shift-shares on migration and trade using international data

	Emigration over land area			Exports over land area		
	(1)	(2)	(3)	(4)	(5)	(6)
Migration S-S	0.370***		0.097	0.000		-0.001
	(0.082)		(0.085)	(0.001)		(0.001)
Trade S-S		0.310	0.133		0.003	0.004*
		(0.204)	(0.251)		(0.002)	(0.002)
Observations	26	26	26	26	26	26

The unit of observation is a country. $^{*}p < .1,^{**}p < .05,^{***}p < .01$

B Mathematical Appendix

B.1 Bartik instruments

Jaeger et al. (2018) define the immigration shift-share instrument as

$$\tilde{m}_{dt} = \sum_{o} \frac{m_{od,t^0}}{m_{o,t^0}} \frac{\Delta m_{ot}}{N_{d,t-1}}$$

Where $m_{od,t^0}/m_{o,t^0}$ is the share of immigrants from o in d at reference date t^0 which predates t. Δm_{ot} is the total number of immigrants across all locations at time t and $N_{d,t-1}$ is the local population in d in the previous period. In this context, we do not necessarily have the stock of immigrants in each location, so we replace with the flow from the previous period and estimate the Bartik instrument for immigration as

$$B_{dt} = \sum_{o} \frac{m_{od,t-1}}{m_{o,t-1}} \frac{m_{ot}}{N_{d,t-1}}$$

Where $m_{od,t-1}/m_{o,t-1}$ is the share of migration out of o into d in the previous period, m_{ot} is total migration out of o in the current period, and $N_{d,t-1}$ is the total population in d at the start of the period. We use the reverse of this to define the emigration shift-share instrument.

B.2 Model Equilibrium

Agricultural Production Labor demand on a plot that is cleared equates marginal cost of that plot with marginal revenue

$$w_i = \alpha p_i^A (N_i^A(\omega))^{\alpha-1} \left(Z_i^A(\omega)\right)^{1-\alpha} \tag{21}$$

$$N_i^A(\omega) = \left(\frac{\alpha p_i^A}{w_i}\right)^{\frac{1}{1-\alpha}} Z_i^A(\omega) \tag{22}$$

Then profit on a plot abstracting from the cost of opening the plot is given by

$$\begin{aligned}
p_i^A \left(\frac{\alpha p_i^A}{w_i}\right)^{\frac{\alpha}{1-\alpha}} Z_i^A(\omega) - w_i \left(\frac{\alpha p_i^A}{w_i}\right)^{\frac{1}{1-\alpha}} Z_i^A(\omega) &= Z_i^A(\omega)\left((p_i^A)^{\frac{1}{1-\alpha}} \alpha^{\frac{\alpha}{1-\alpha}} w_i^{\frac{-\alpha}{1-\alpha}} - (p_i^A)^{\frac{1}{1-\alpha}} \alpha^{\frac{1}{1-\alpha}} w_i^{\frac{-\alpha}{1-\alpha}}\right) \\
&= Z_i^A(\omega)(p_i^A)^{\frac{1}{1-\alpha}} w_i^{\frac{-\alpha}{1-\alpha}} \left(\alpha^{\frac{\alpha}{1-\alpha}} - \alpha^{\frac{1}{1-\alpha}}\right) \\
&= Z_i^A(\omega)\left(p_i^A w_i^{-\alpha} \tilde{\alpha}\right)^{\frac{1}{1-\alpha}} \\
&= \psi_i Z_i^A(\omega)
\end{aligned}$$

Forested plot ω is cleared whenever it is profitable to do so

$$\psi_i Z_i^A(\omega) > Z_i^0(\omega)$$

Where $\tilde{\alpha} = \alpha^\alpha (1-\alpha)^{1-\alpha}$. Integrating over ω the share of forested land for which it is profitable

to clear is then

$$S_i^A = Pr\Big(\psi_i Z_i^A(\omega) > Z_i^0(\omega)\Big) = \frac{\big(\Lambda_i\psi_i\big)^\theta}{\big(p^M\Lambda_i^0\big)^\theta + \big(\Lambda_i\psi_i\big)^\theta}$$

Total labor demand in j is found by summing over demand on each plot, forested and deforested:

$$\begin{aligned}
N_i^A =& \left(\frac{\alpha p_i^A}{w_i}\right)^{\frac{1}{1-\alpha}} \Big[\bar{L}_i^F S_i E[Z_i^A|\psi_i Z_i^A > Z_i^0] + \bar{L}_i^C E[Z_i^A]\Big] \\
=& \left(\frac{\alpha p_i^A}{w_i}\right)^{\frac{1}{1-\alpha}} \Big[\frac{1}{\psi_i}\bar{L}_i S_i E\Big[Z_i^A\psi_i|Z_i^A\psi_i > Z_i^0\Big] + \bar{L}_i^C E[Z_i^A]\Big] \\
=& \left(\frac{\alpha p_i^A}{w_i}\right)^{\frac{1}{1-\alpha}} \Big[\frac{1}{\psi_i}\bar{L}_i S_i\Big(\big(p^M\Lambda_i^0\big)^\theta + \big(\Lambda_i\psi_i\big)^\theta\Big)^{1/\theta} + \bar{L}_i^C\Lambda_i\Big]\Gamma\Big(1-\frac{1}{\theta}\Big) \\
=& \left(\frac{\alpha p_i^A}{w_i}\right)^{\frac{1}{1-\alpha}} \Big[\frac{1}{\psi_i}\bar{L}_i S_i\frac{\Lambda_i\psi_i}{(S_i)^{1/\theta}} + \bar{L}_i^C\Lambda_i\Big]\Gamma\Big(1-\frac{1}{\theta}\Big) \\
=& \left(\frac{\alpha p_i^A}{w_i}\right)^{\frac{1}{1-\alpha}} \Lambda_i\Big[\bar{L}_i S_i (S_i)^{-1/\theta} + \bar{L}_i^C\Big]\Gamma\Big(1-\frac{1}{\theta}\Big) \\
=& \frac{\alpha}{1-\alpha}\Big(\frac{\psi_i}{p_i^A}\Big)\Big[\bar{L}_i (S_i)^{\frac{\theta-1}{\theta}} + \bar{L}_i^C\Big]\Gamma\Big(1-\frac{1}{\theta}\Big)
\end{aligned}$$

The final expression is found by normalizing the Frechet distributions by $\bar{\Gamma} = \Gamma\Big(1-\frac{1}{\theta}\Big)^{-1/\theta}$ so that $E[Z_i^A] = \Lambda_i$ and $E[A_i^0] = \Lambda_i^0$.

Note that on each plot $Q_i^A(\omega) = (N_i^A(\omega))^\alpha (Z_i^A(\omega))^{(1-\alpha)}$. Substituting out for $Z_i^A(\omega)$ using (22) then gives $Q_i^A(\omega) = N_i^A(\omega)\frac{w_i}{\alpha p_i^A}$. Integrating over ω then gives

$$Q_i^A = N_i^A\frac{w_i}{\alpha p_i^A} \tag{23}$$

$$= \frac{1}{1-\alpha}\Big(\frac{\psi_i}{p_i^A}\Big)\Lambda_i\Big[\bar{L}_i (S_i)^{\frac{\theta-1}{\theta}} + \bar{L}_i^C\Big] \tag{24}$$

Land rents, which are returned to local workers are equal to agricultural profits minus the fixed cost of opening the land. The expected value of Z_i^0 conditional on the land being used for agriculture is given by

$$\begin{aligned}
E[Z_i^0|Z_i^A\psi_i > Z_i^0] * Pr(Z_i^A\psi_i > Z_i^0) =& E[Z_i^0] - E[Z_i^0|Z_i^0 > Z_i^A\psi_i] * Pr(Z_i^0 > Z_i^A\psi_i) \\
E[Z_i^0|Z_i^A\psi_i > Z_i^0] * Pr(Z_i^A\psi_i > Z_i^0) =& p^M\Lambda_i^0 - \Big(\big(p^M\Lambda_i^0\big)^\theta + \big(\Lambda_i\psi_i\big)^\theta\Big)^{1/\theta} * (1-S_i)
\end{aligned}$$

We can then substitute using the expression for the share of land not used for agriculture

$$1 - S_i = \frac{\left(p^M \Lambda_i^0\right)^\theta}{\left(p^M \Lambda_i^0\right)^\theta + \left(\Lambda_i \psi_i\right)^\theta}$$

So that

$$\begin{aligned} E[Z_i^0 | Z_i^A \psi_i > Z_i^0] * Pr(Z_i^A \psi_i > Z_i^0) =& p^M \Lambda_i^0 - p^M \Lambda_i^0 (1 - S_i)^{1/\theta} * (1 - S_i) \\ E[Z_i^0 | Z_i^A \psi_i > Z_i^0] * Pr(Z_i^A \psi_i > Z_i^0) =& p^M \Lambda_i^0 \left(1 - (1 - S_i)^{\frac{\theta-1}{\theta}}\right) \end{aligned}$$

Total land rents are then

$$\begin{aligned} R_i =& p_i^A Q_i^A - w_i N_i^A - \bar{L}_i E[Z_i^0 | Z_i^A \psi_i > Z_i^0] * Pr(Z_i^A \psi_i > Z_i^0) \\ R_i =& (1 - \alpha_A) p_i^A Q_i^A - \bar{L}_i^F p^M \Lambda_i^0 \left(1 - (1 - S_i)^{\frac{\theta-1}{\theta}}\right) \end{aligned}$$

Manufacturing Profits Profit maximization in the manufacturing sector implies that labor demand in manufacturing is given by

$$N_d^M = \left(\frac{\alpha Z_d}{w_d}\right)^{\frac{1}{1-\alpha}}$$

Then profit in the manufacturing sector is given by

$$\begin{aligned} \Pi_d^M =& Z_d \left(\frac{\alpha Z_d}{w_d}\right)^{\frac{\alpha}{1-\alpha}} - w_d \left(\frac{\alpha Z_d}{w_d}\right)^{\frac{1}{1-\alpha}} \\ \Pi_d^M =& Z_d^{\frac{1}{1-\alpha}} w_d^{\frac{-\alpha}{1-\alpha}} \left(\alpha^{\frac{\alpha}{1-\alpha}} - \alpha^{\frac{1}{1-\alpha}}\right) \\ \Pi_d^M =& Z_d^{\frac{1}{1-\alpha}} w_d^{\frac{-\alpha}{1-\alpha}} \alpha^{\frac{\alpha}{1-\alpha}} \left(1 - \alpha\right) \\ \Pi_d^M =& \left(\frac{Z_d \tilde{\alpha}}{w_d^\alpha}\right)^{\frac{1}{1-\alpha}} \end{aligned}$$

Consumption Decisions Fixing d, let $U(C^M, C^A)$ be the utility value of C_d when consuming quantities C^M and C^A. The individuals problem can be expressed as choosing quantities to maximize the constrained optimization problem

$$\begin{aligned} & \max_{C_d^M, \{C_{od}^A\}} U(C_d^M, C_d^A) \\ \text{subject to} \qquad & C_d^A = \left(\sum_o (\beta_{od})^{\frac{1}{\eta}} (C_{od}^A)^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}} \\ & p_M C_d^M + \sum_o p_o^A \tau_{od} C_{od}^A = w_d + R_d / N_d \end{aligned}$$

The first order condition for C_{od}^A is

$$U_2 \beta_{od}^{\frac{1}{\eta}} {C_{od}^A}^{\frac{-1}{\eta}} {C_d^A}^{\frac{1}{\eta}} - \lambda p_o^A \tau_{od} = 0$$

Rearranging and summing over o yields an expression for λ / U_2

$$\begin{aligned}
\beta_{od}^{\frac{1}{\eta}} (C_{od}^A)^{\frac{\eta-1}{\eta}} =& \frac{\lambda p_o^A \tau_{od} C_{od}^A}{U_2 (C_d^A)^{\frac{1}{\eta}}} \\
\sum_o \beta_{od}^{\frac{1}{\eta}} (C_{od}^A)^{\frac{\eta-1}{\eta}} =& \frac{\lambda}{U_2 (C_d^A)^{\frac{1}{\eta}}} \Big(\sum_o p_o^A \tau_{od} C_{od}^A \Big) \\
\left(\sum_o \beta_{od}^{\frac{1}{\eta}} (C_{od}^A)^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}} =& C_d^A = \left(\frac{\lambda}{U_2 (C_d^A)^{\frac{1}{\eta}}} \right)^{\frac{\eta}{\eta-1}} \Big(\sum_o p_o^A \tau_{od} C_{od}^A \Big)^{\frac{\eta}{\eta-1}} \\
C_d^A =& \left(\frac{\lambda}{U_2 (C_d^A)^{\frac{1}{\eta}}} \right)^{\frac{\eta}{\eta-1}} (X_d^A)^{\frac{\eta}{\eta-1}} \\
\frac{C_d^A}{X_d^A} =& \frac{\lambda}{U_2}
\end{aligned}$$

Where X_d^A is total expenditure on agricultural goods in d. Returning to the FOC, we can express demand for C_{od}^A as

$$C_{od}^A = \beta_{od} \left(\frac{X_d^A}{p_o^A \tau_{od}} \right)^{\eta} (C_d^A)^{1-\eta}$$

This expression still depends upon C_d^A. We can solve for it by substituting this expression into the definition

$$\begin{aligned}
C_d^A =& \Big(\sum_o (\beta_{od})^{\frac{1}{\eta}} (C_{od}^A)^{\frac{\eta-1}{\eta}} \Big)^{\frac{\eta}{\eta-1}} \\
C_d^A =& \Big(\sum_o (\beta_{od})^{\frac{1}{\eta}} \left(\beta_{od} \left(\frac{X_d^A}{p_o^A \tau_{od}} \right)^{\eta} (C_d^A)^{1-\eta} \right)^{\frac{\eta-1}{\eta}} \Big)^{\frac{\eta}{\eta-1}} \\
(C_d^A)^{\eta} =& \left(\sum_o \beta_{od} \Big(p_o^A \tau_{od} \Big)^{1-\eta} \right)^{\frac{\eta}{\eta-1}} (X_d^A)^{\eta} \\
P_d^A C_d^A =& X_d^A
\end{aligned}$$

Where $P_d^A = \left(\sum_o \beta_{od} \Big(p_o^A \tau_{od} \Big)^{1-\eta} \right)^{\frac{1}{1-\eta}}$ is the price index that equates consumption of the composite good with total expenditure on agricultural goods. Demand for C_{od}^A can then be expressed as

$$C_{od}^A = \beta_{od} \big(p_o^A \tau_{od} \big)^{-\eta} (P_d^A)^{\eta-1} X_d^A$$

Non-Homothetic Demand Functions To make the exposition simpler, we will consider the general case in which there are non-homotheticities in both goods which I will denote by ε_M and ε_A. Recall that C_d is the utility function that the individual seeks to maximize. By rearranging (7), we can express this quantity as:

$$C_d = \left[(\Omega_d^M)^{\frac{1}{\sigma}} \left(\frac{C_d^M}{(C_d)^{\varepsilon_M - 1}} \right)^{\frac{\sigma-1}{\sigma}} + (\Omega_d^A)^{\frac{1}{\sigma}} \left(\frac{C_d^A}{(C_d)^{\varepsilon_A - 1}} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}$$

Then we can express the maximization problem in the standard way

$$C_d = \max_{C_d^M, C_d^A} \left[(\Omega_d^M)^{\frac{1}{\sigma}} \left(\frac{C_d^M}{(C_d)^{\varepsilon_M - 1}} \right)^{\frac{\sigma-1}{\sigma}} + (\Omega_d^A)^{\frac{1}{\sigma}} \left(\frac{C_d^A}{(C_d)^{\varepsilon_A - 1}} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} + \lambda(E - p_M C_d^M - p_A C_d^A)$$

Implicit differentiation with respect to C_d^M yields

$$\begin{aligned}
\frac{\partial C_d}{\partial C_d^M} =& \frac{\sigma}{\sigma - 1} \left[(\Omega_d^M)^{\frac{1}{\sigma}} \left(\frac{C_d^M}{(C_d)^{\varepsilon_M - 1}} \right)^{\frac{\sigma-1}{\sigma}} + (\Omega_d^A)^{\frac{1}{\sigma}} \left(\frac{C_d^A}{(C_d)^{\varepsilon_A - 1}} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{1}{\sigma-1}} \times \\
& \left(\frac{\sigma - 1}{\sigma} (\Omega_d^M)^{\frac{1}{\sigma}} \left(\frac{C_d^M}{(C_d)^{\varepsilon_M - 1}} \right)^{\frac{-1}{\sigma}} \left(\frac{(C_d)^{\varepsilon_M - 1} - C_d^M (\varepsilon_M - 1)(C_d)^{\varepsilon_M - 2} \frac{\partial C_d}{\partial C_d^M}}{((C_d)^{\varepsilon_M - 1})^2} \right) + \right. \\
& \left. \frac{\sigma - 1}{\sigma} (\Omega_d^A)^{\frac{1}{\sigma}} \left(\frac{C_d^A}{(C_d)^{\varepsilon_A - 1}} \right)^{\frac{-1}{\sigma}} \left(\frac{-C_d^A (\varepsilon_A - 1)(C_d)^{\varepsilon_A - 2} \frac{\partial C_d}{\partial C_d^M}}{((C_d)^{\varepsilon_A - 1})^2} \right) \right) - \lambda p_M
\end{aligned}$$

Under the assumption that we are at an interior solution, at an optimum, $\frac{\partial C_d}{\partial C_d^M} = 0$. Thus the equation simplifies to

$$\begin{aligned}
\lambda p_M =& \frac{\sigma}{\sigma - 1} \left[(\Omega_d^M)^{\frac{1}{\sigma}} \left(\frac{C_d^M}{(C_d)^{\varepsilon_M - 1}} \right)^{\frac{\sigma-1}{\sigma}} + (\Omega_d^A)^{\frac{1}{\sigma}} \left(\frac{C_d^A}{(C_d)^{\varepsilon_A - 1}} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{1}{\sigma-1}} \times \\
& \left(\frac{\sigma - 1}{\sigma} (\Omega_d^M)^{\frac{1}{\sigma}} \left(\frac{C_d^M}{(C_d)^{\varepsilon_M - 1}} \right)^{\frac{-1}{\sigma}} \left(\frac{(C_d)^{\varepsilon_M - 1}}{((C_d)^{\varepsilon_M - 1})^2} \right) \right) \\
\lambda p_M =& \left[(\Omega_d^M)^{\frac{1}{\sigma}} \left(\frac{C_d^M}{(C_d)^{\varepsilon_M - 1}} \right)^{\frac{\sigma-1}{\sigma}} + (\Omega_d^A)^{\frac{1}{\sigma}} \left(\frac{C_d^A}{(C_d)^{\varepsilon_A - 1}} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{1}{\sigma-1}} \left((\Omega_d^M)^{\frac{1}{\sigma}} \left(\frac{C_d^M}{(C_d)^{\varepsilon_M - 1}} \right)^{\frac{-1}{\sigma}} \left(\frac{1}{(C_d)^{\varepsilon_M - 1}} \right) \right) \\
\lambda p_M =& (C_d)^{\frac{1}{\sigma}} \left((\Omega_d^M)^{\frac{1}{\sigma}} \left(\frac{C_d^M}{(C_d)^{\varepsilon_M - 1}} \right)^{\frac{-1}{\sigma}} \left(\frac{1}{(C_d)^{\varepsilon_M - 1}} \right) \right) \\
\lambda p_M C_d^M =& (C_d)^{\frac{1}{\sigma}} \left((\Omega_d^M)^{\frac{1}{\sigma}} \left(\frac{C_d^M}{(C_d)^{\varepsilon_M - 1}} \right)^{\frac{\sigma-1}{\sigma}} \right)
\end{aligned}$$

Using the equivalent logic for C_d^A we solve for λ by

$$\begin{aligned}
\lambda\left(p_M C_d^M + p_A C_d^A\right) =& (C_d)^{\frac{1}{\sigma}}\left((\Omega_d^M)^{\frac{1}{\sigma}}\left(\frac{C_d^M}{(C_d)^{\varepsilon_M - 1}}\right)^{\frac{\sigma-1}{\sigma}}\right) + (C_d)^{\frac{1}{\sigma}}\left((\Omega_d^A)^{\frac{1}{\sigma}}\left(\frac{C_d^A}{(C_d)^{\varepsilon_A - 1}}\right)^{\frac{\sigma-1}{\sigma}}\right) \\
\lambda X_d =& (C_d)^{\frac{1}{\sigma}}\left((\Omega_d^M)^{\frac{1}{\sigma}}\left(\frac{C_d^M}{(C_d)^{\varepsilon_M - 1}}\right)^{\frac{\sigma-1}{\sigma}} + (\Omega_d^A)^{\frac{1}{\sigma}}\left(\frac{C_d^A}{(C_d)^{\varepsilon_A - 1}}\right)^{\frac{\sigma-1}{\sigma}}\right) \\
\lambda X_d =& (C_d)^{\frac{1}{\sigma}}(C_d)^{\frac{\sigma-1}{\sigma}} \\
\lambda =& \frac{C_d}{X_d}
\end{aligned}$$

Where X_d is total expenditure in d. Returning to the expression for C_d^M

$$\begin{aligned}
(C_d^M)^{\frac{1}{\sigma}} =& (C_d)^{\frac{1}{\sigma}}\left((\Omega_d^M)^{\frac{1}{\sigma}}\left(\frac{1}{(C_d)^{\varepsilon_M - 1}}\right)^{\frac{\sigma-1}{\sigma}}\right)\frac{X_d}{P_d^M C_d} \\
C_d^M =& C_d \Omega_d^M \left(\frac{1}{(C_d)^{\varepsilon_M - 1}}\right)^{\sigma-1}\left(\frac{X_d}{P_d^M C_d}\right)^{\sigma} \\
C_d^M =& \Omega_d^M \left(\frac{1}{(C_d)^{\varepsilon_M}}\right)^{\sigma-1}\left(\frac{X_d}{P_d^M}\right)^{\sigma} = \Omega_d^M \left(\frac{X_d}{P_d^M}\right)^{\sigma}(C_d)^{\varepsilon_M(1-\sigma)} \\
C_d^A =& \Omega_d^A \left(\frac{1}{(C_d)^{\varepsilon_A}}\right)^{\sigma-1}\left(\frac{X_d}{P_d^A}\right)^{\sigma} = \Omega_d^A \left(\frac{X_d}{P_d^A}\right)^{\sigma}(C_d)^{\varepsilon_A(1-\sigma)}
\end{aligned}$$

The expression for total consumption can be expressed as

$$\begin{aligned}
C_d =& \left[(\Omega_d^M)^{\frac{1}{\sigma}}\left(\Omega_d^M\left(\frac{1}{(C_d)^{\varepsilon_M}}\right)^{\sigma-1}\left(\frac{X_d}{P_d^M}\right)^{\sigma}\frac{1}{(C_d)^{\varepsilon_M - 1}}\right)^{\frac{\sigma-1}{\sigma}}\right. \\
&\left. + (\Omega_d^A)^{\frac{1}{\sigma}}\left(\Omega_d^A\left(\frac{1}{(C_d)^{\varepsilon_A}}\right)^{\sigma-1}\left(\frac{X_d}{P_d^A}\right)^{\sigma}\frac{C_d^A}{(C_d)^{\varepsilon_A - 1}}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \\
C_d =& \left[\Omega_d^M\left(\left(\frac{C_d}{(C_d^{\varepsilon_M})^{\sigma}}\right)\left(\frac{X_d}{P_d^M}\right)^{\sigma}\right)^{\frac{\sigma-1}{\sigma}} + \Omega_d^A\left(\left(\frac{C_d}{(C_d^{\varepsilon_A})^{\sigma}}\right)\left(\frac{X_d}{P_d^A}\right)^{\sigma}\right)^{\frac{\sigma-1}{\sigma}}\right]^{\frac{\sigma}{\sigma-1}} \\
C_d =& \left[\Omega_d^M\left((C_d)^{\varepsilon_M - 1}P_d^M\right)^{1-\sigma} + \Omega_d^A\left((C_d)^{\varepsilon_A - 1}P_d^A\right)^{1-\sigma}\right]^{\frac{\sigma}{\sigma-1}} C_d^{1-\sigma}X_d^{\sigma} \\
C_d P_d =& X_d \\
P_d =& \left[\Omega_d^M\left((C_d)^{\varepsilon_M - 1}P_d^M\right)^{1-\sigma} + \Omega_d^A\left((C_d)^{\varepsilon_A - 1}P_d^A\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}
\end{aligned}$$

In the context of our model, this yields the two Hicksian demand functions

$$
\begin{aligned}
C_d^M =& \Omega_d^M \left(\frac{1}{X_d} \right)^{-\sigma} (C_d)^{(1-\sigma)} \\
C_d^A =& \Omega_d^A \left(\frac{P_d^A}{X_d} \right)^{-\sigma} (C_d)^{\varepsilon(1-\sigma)} \\
C_d =& \left[\Omega_d^M + \Omega_d^A \left((C_d)^{\varepsilon-1} P_d^A \right)^{1-\sigma} \right]^{\frac{1}{\sigma-1}} X_d \\
P_d =& \left[\Omega_d^M + \Omega_d^A \left((C_d)^{\varepsilon-1} P_d^A \right)^{1-\sigma} \right]^{\frac{1}{1-\sigma}}
\end{aligned}
$$

Expenditure shares We now derive an expression of expenditure share of each good as a function of its price and of price indices for each location i and sector of employment k. The ratios of the expenditure shares of agricultural and service goods to the manufacturing good are given by:

$$
\frac{\mu_{ik}^A}{\mu_{ik}^M} = \Omega_A \left(\frac{P_i^A}{p^M} \right)^{(1-\sigma)} c_{ik}^{(\epsilon_A-1)(1-\sigma)}, \tag{25}
$$

$$
\frac{\mu_{ik}^S}{\mu_{ik}^M} = \Omega_S \left(\frac{p_i^S}{p^M} \right)^{(1-\sigma)} c_{ik}^{(\epsilon_S-1)(1-\sigma)}. \tag{26}
$$

Since $\mu_{ik}^M = 1 - \mu_{ik}^A - \mu_{ik}^S$, this implies:

$$
\mu_{ik}^A = \frac{\Omega_A \left(\frac{P_i^A}{p^M} \right)^{1-\sigma} c_{ik}^{(\epsilon_A-1)(1-\sigma)}}{1 + \Omega_A \left(\frac{P_i^A}{p^M} \right)^{1-\sigma} c_{ik}^{(\epsilon_A-1)(1-\sigma)} + \Omega_S \left(\frac{p_i^S}{p^M} \right)^{1-\sigma} c_{ik}^{(\epsilon_S-1)(1-\sigma)}}, \tag{27}
$$

$$
\mu_{ik}^M = \frac{1}{1 + \Omega_A \left(\frac{P_i^A}{p^M} \right)^{1-\sigma} c_{ik}^{(\epsilon_A-1)(1-\sigma)} + \Omega_S \left(\frac{p_i^S}{p^M} \right)^{1-\sigma} c_{ik}^{(\epsilon_S-1)(1-\sigma)}}, \tag{28}
$$

$$
\mu_{ik}^S = \frac{\Omega_S \left(\frac{p_i^S}{p^M} \right)^{1-\sigma} c_{ik}^{(\epsilon_S-1)(1-\sigma)}}{1 + \Omega_A \left(\frac{P_i^A}{p^M} \right)^{1-\sigma} c_{ik}^{(\epsilon_A-1)(1-\sigma)} + \Omega_S \left(\frac{p_i^S}{p^M} \right)^{1-\sigma} c_{ik}^{(\epsilon_S-1)(1-\sigma)}}. \tag{29}
$$

Defining $P_{ik} = \frac{e_{ik}}{c_{ik}}$ and multiplying numerators and denominators by $(p^M)^{1-\sigma}$, then we can write this as

$$\mu_{ik}^A = \frac{\Omega_A \left(P_i^A\right)^{(1-\sigma)} \left(e_{ik}/P_{ik}\right)^{(\epsilon_A-1)(1-\sigma)}}{(p^M)^{(1-\sigma)} + \Omega_A \left(P_i^A\right)^{(1-\sigma)} \left(e_{ik}/P_{ik}\right)^{(\epsilon_A-1)(1-\sigma)} + \Omega_S \left(p_i^S\right)^{(1-\sigma)} \left(e_{ik}/P_{ik}\right)^{(\epsilon_S-1)(1-\sigma)}}, \quad (30)$$

$$\mu_{ik}^M = \frac{(p^M)^{(1-\sigma)}}{(p^M)^{(1-\sigma)} + \Omega_A \left(P_i^A\right)^{(1-\sigma)} \left(e_{ik}/P_{ik}\right)^{(\epsilon_A-1)(1-\sigma)} + \Omega_S \left(p_i^S\right)^{(1-\sigma)} \left(e_{ik}/P_{ik}\right)^{(\epsilon_S-1)(1-\sigma)}}, \quad (31)$$

$$\mu_{ik}^S = \frac{\Omega_S \left(p_i^S\right)^{(1-\sigma)} \left(e_{ik}/P_{ik}\right)^{(\epsilon_S-1)(1-\sigma)}}{(p^M)^{(1-\sigma)} + \Omega_A \left(P_i^A\right)^{(1-\sigma)} \left(e_{ik}/P_{ik}\right)^{(\epsilon_A-1)(1-\sigma)} + \Omega_S \left(p_i^S\right)^{(1-\sigma)} \left(e_{ik}/P_{ik}\right)^{(\epsilon_S-1)(1-\sigma)}}. \quad (32)$$

Price indices Conversely we can derive an expression of price indices for each location i and sector of employment k as a function of expenditure shares and prices of agricultural goods, manufacturing, and services. To derive an expression for P_{jk} in terms of e_{jk} but not c_{jk}, we use equation (13) of Comin et al. (2021) which gives:

$$\frac{e_{jk}}{P_{jk}} = c_{jk} = \frac{e_{jk}}{p^M}(\mu_{jk}^M)^{\frac{1}{1-\sigma}} \quad (33)$$

Rearranging therefore gives the price indices for each location i and sector of employment k

$$\begin{aligned} P_{ik} &= \frac{p^M}{(\mu_{ik}^M)^{\frac{1}{1-\sigma}}} \\ &= \left((p^M)^{(1-\sigma)} + \Omega_A \left(P_i^A\right)^{(1-\sigma)} \left(e_{ik}/P_{ik}\right)^{(\epsilon_A-1)(1-\sigma)} + \Omega_S \left(p_i^S\right)^{(1-\sigma)} \left(e_{ik}/P_{ik}\right)^{(\epsilon_S-1)(1-\sigma)}\right)^{\frac{1}{1-\sigma}} \\ &= ((p^M)^{(1-\sigma)} + \Omega_A \left(P_i^A\right)^{(1-\sigma)} \left(e_{ik}/p_M\right)^{(\epsilon_A-1)(1-\sigma)} (1-\mu_{ik}^A-\mu_{ik}^S)^{(\epsilon_A-1)} \\ &+ \Omega_S \left(p_i^S\right)^{(1-\sigma)} \left(e_{ik}/p_M\right)^{(\epsilon_S-1)(1-\sigma)} (1-\mu_{ik}^A-\mu_{ik}^S)^{(\epsilon_S-1)})^{\frac{1}{1-\sigma}} \end{aligned} \quad (34)$$

Migration Costs Equation 19 gives the migration rate from i to j :

$$m_{ij} = \frac{\left(V_{ij}\right)^\kappa}{\sum_{j'} \left(V_{ij'}\right)^\kappa}$$

where $V_{ij} = 1/\delta_{ij} \times \left((\beta_{jA} v_j^A/P_j^A)^\gamma + (v_j^M/P_j^M)^\gamma + (\beta_{jS} v_j^S/P_j^S)^\gamma\right)^{1/\gamma}$ is the expected indirect utility from migrating to j.

$$\frac{m_{ij}m_{ji}}{m_{ii}m_{jj}} = \frac{\left(V_{ij}V_{ji}\right)^\kappa}{\left(V_{ii}V_{jj}\right)^\kappa} = \delta_{ij}^{-2\kappa} \quad (35)$$

Which yields the desired expression:

$$\delta_{ij} = \left(\frac{m_{ij}m_{ji}}{m_{ii}m_{jj}}\right)^{-\frac{1}{2\kappa}} \quad (36)$$

C Data appendix

C.1 Census Data

All IPUMS-I censuses are 10% samples unless otherwise indicated. Migration flows are constructed using the most disaggregated level of administrative geography available, except for Mauritius. Our preferred question is, "in which year did you move here and where did you live previously?" When we know the year of migration, we exclude migrants in the previous year to exclude temporary and transitional migrants from our measure. When migration history is asked in more than one way and we do not have the year of migration, we give preference to (in this order), "where did you live 10 years ago?", "where did you live five years ago?", "where were you born?", and "where did you live one year ago?".

Angola: 2014 census. Respondents were asked in which municipality (level 2) and province (level 1) they lived in 5 years ago. Information is also available for the municipality and province of residence 1 year before the census, and for the municipality and province of birth.

Benin: 2002 and 2013 censuses. Migration is asked as, "length of time in current residence," (translated from French to English). For 2013, the values range from 1948-2013, and for 2002, from 1937-2002. Our measures of migration uses flows between 1992-2001 and 2003-2012 for the first and second censuses, respectively.

Botswana: 2001 and 2011 censuses. Respondents were asked where they lived at this time last year and in August of 1996/2006, respectively. This information, as well as the birth location, are collected at the district (level 1 geography) level.

Burkina Faso: 1996 and 2006 censuses. Respondents were asked which province (level 1) and department (level 2) they lived one year ago. The province and department of birth are also available.

Cameroon: 1987 and 2005 censuses. In 1987, respondents were asked in which province (level 1) and department (level 2) they lived in 11 years ago. In 2005, respondents were again asked in which province and department they lived in 11 years ago.

Côte d'Ivoire: 1998 census. Respondents were asked about their region (level 1) and department (level 2) of residence one year ago. Information is also available about the location of birth.

Ethiopia: socioeconomic survey 2013-2014 (Central Statistics Agency , CSA). The survey began in 2011/12 as the Ethiopia Rural Socioeconomic Survey and was expanded in wave 2 to include all urban areas, creating a nationally representative sample of 5,262 households. We use data from wave 2 of the ESS. Respondents were asked whether they had migrated to a different region (level 1) since the timing of the previous wave. As the median time between waves is 2 years, we construct migration variables as moving in the last 2 years.

Gabon: 1993 and 2013 censuses. In the absence of micro data about these censuses, we used aggregated tables that appeared in reports about the results of the censuses. For the 1993 census, the report includes migration matrices showing the number of people living in each province (level 1) and their province of birth (Direction Générale de la Statistique , DGS). Similarly, for the 2013 census, the report includes migration matrices of current province and province of birth (Direction Générale de la Statistique , DGS).

Ghana: 2000 and 2010 censuses. In the 2000 census, respondents were asked what was the district of usual residence five years ago? This question is not repeated in the 2010 census. Instead respondents were asked for how many years have they lived in the current town/village if they haven't lived there since birth. In both censuses, respondents were asked for the region of birth. We use the district of residence five years ago as our measure of migration.

Guinea: 1996 and 2014 censuses. Respondents were asked which prefecture (level 2) they lived in before moving to the current prefecture. In 1996, respondents were asked if they have lived in the current location for less than one year, between 1-4 years, 5-9 years, 10 or more years, or since birth. In 2014, respondents were asked an open-ended question of how many years they have lived in the current prefecture. Information is also available for the location of birth. We construct migration flows using the the periods 1986-1995 and 2004-2013 to construct migration flows in the first and second censuses respectively. The period 1986-1995 will include all migrants from the previous ten years and exclude migrants from the previous year.

Kenya: 1999 and 2009 censuses. Migration history is obtained in two stages. First respondents were asked, "where was [the respondent] living in August of [the previous year]?" Next, respondents who are not living in their district of birth were asked for the year and month that they moved there. We construct flows using the location of previous residence rather than the birth location.

Lesotho: 1996 and 2006 censuses. Respondents were asked where they were living at the time of the previous census (April 1986 and 1996, respectively). If the respondent was not living in the same village they were then asked in what year they moved to their current location.

Malawi: 2008 census. Respondents were asked about their district (level 1) of previous residence, and how long they lived in their current place of residence. We use migrants who lived in a different district within the previous 10 years to construct migration.

Mali: 1998 and 2009 censuses. Data on the year of migration, previous location, and birth location are all available. The exact wording in the migration question is slightly different in the two censuses, but the premise is the same. Respondents were asked if they have lived anywhere else for at least six months. If yes, respondents were asked where they lived immediately preceding their current address. Respondents are then asked for the duration of their current residence in years. We use previous location and duration to construct migration.

Mauritius: 2000 and 2011 censuses. No data is available on the location of birth. We construct migration flows using the location of residence five years ago. Respondents were asked, "what was your usual address five years ago?" This data is available at both the district and region level. However, given the large amount of movement between level 2 administrative units (MVCA/wards) within Port Louis, we use the level 1 administrative units.

Madagascar: 2018 census. This census is not harmonized in IPUMS, and we obtain it from the statistical institute of Madagascar (Institut National de la Statistique , INSTAT). Respondents were asked about their previous region (level 1) and district (level 2) of residence and about the duration of living in the current location. We construct the migration variables based on those.

Mozambique: 1997 and 2007 censuses. In 1997, respondents were asked where they lived in October 1992, when the war ended. In 2007, this text was changed to where was the respondent living five years ago. In both rounds, respondents were also asked where they were living one year ago and where they were born. Migration in the five years after the war will likely contain a lot of moves back home, which is not what we want to measure. For this reason, we choose to estimate migration using the district of birth.

Nigeria: 2010/11 general household survey (National Bureau of Statistics, Federal Republic of Nigeria, 2010-2011). This is the first wave of a nationally representative panel survey, with 5,000 households. Migration variables are constructed using data from the second wave (2012/13), when respondents were asked whether they had migrated to a different state (level 1) since the previous wave. As the median time between waves is 2 years, we define migration as moving within the last 2 years.

Rwanda: 2012 census. Respondents were asked about the district (level 2) of previous residence and about the duration of residence in their current district. We take migration in the last 10 years to construct the migration values.

Senegal: 2002 and 2013 censuses. Although information on the location of residence is available at one, five, and ten year intervals in the 2013 census, it is only available in a five year interval in 2002. Respondents who are residents are asked to write the name of the commune or rural community of residence five years ago, and visitors are asked to record the place where they came from. Birth location is also available, although we use the location of residence five years ago.

Sierra Leone: 2004 and 2015 censuses. In 2004, respondents were asked where they were living in December 1990 (14 years ago), and in 2015 they were asked where they were living in December 2010 (5 years ago). Due to the lack of consistency of these measures, we use the location of birth.

South Africa: The post-apartheid government of South Africa conducted its first census in 1996 and was followed by a second census in 2001. A third census was scheduled for 2006, but Statistics South Africa did not have the capacity to conduct a full census that time and it was replaced with a community survey. The third census took place in 2011 and a second community survey took place in 2016. As a result, the questions in each year varied. In 1996, respondents were asked in which year they moved to their current dwelling and responses were open ended. In 2001, the responses were recorded only for people who moved in the past five years (since the first census). In the 2007 community survey, moves were again only recorded if they happened after 2001. In 2011, the next census year, the question was asked for moves to the current place since 2001. This is different in two dimensions. First, it covers the previous ten years rather than five. Second, the question is less broad. The original question asked about when the respondent moved to their current dwelling (*i.e.* housing unit), while the latter asks when they moved to the current place (*i.e.* municipality). Finally, in 2016, respondents were asked for the year that they moved to the current place again which was top-coded at five-years. Information on the previous location is not available in the 1996 census. The province (level 1 geography) of previous residence and birth is available in all years beginning in 2001. The district of previous residence is only available in 2011 and 2016, and is never available for the birth location. Hence, we ultimately use these latter two surveys (2011 and 2016) in our analysis and use migration flows over the previous five years at the district level.

Sudan: 2008 census. Respondents were asked about their previous region (level 1) of residence, and about the duration of residence in current region. We keep migration that occurred in the 10 years prior to the census. Information about the location of birth is also available, as well as the region of residence 1 year ago. As in 2008, South Sudan has not yet become independent, we treat units in the two modern-day countries, that appear separately in the IPUMS dataset, as part of a single country.

Tanzania: 2002 and 2012 censuses. Respondents were asked where they lived in 2001/2011, respectively, *i.e.* one year prior. The text did not specify in which month. Previous residence information was collected at the region and district level in 2012, but only at the region (level 1 administrative unit) in 2002. We use the location of birth to construct migration flows which is only available at the region level.

Togo: 2010 census. Respondents were asked about their region (level 1) and prefecture (level 2) of previous residence, and about the duration of their residence in the current canton (level 3). We construct migration variables based on migration in the previous 10 years. Information is also available about the location of birth.

Uganda: 2002 and 2014 censuses. Information on the birth location is only available for the 2002 census. Respondents were asked in which district they were living prior to the current district and for

how many years they have been living in the current district. We use this information to construct migration flows during the ten-year windows preceding each census, 1992-2001 and 2004-2013. This excludes all migration that happened in the year just preceding the census.

Zambia: 2000 and 2010 censuses. Respondents were asked where were they living before August of the previous year. This information is available at the first and second level of geographic administrative unit. However, given the short duration, we instead use the location of birth.

Zimbabwe: 2012 census. Respondents were asked about their district (level 2) of resident 10 years ago.

C.2 Crop potential and actual output

We construct measures of agricultural potential and actual output at the sub-national unit level using data from the Global Agro-Ecological Zones version 4 database (Fischer et al., 2021). For each administrative unit, as well as each country not in our sample, we sum pixel-level raster values to produce unit-level aggregates.

Potential Yield For potential yield, we extract for each unit and disaggregated crop the total suitable area in hectares and the potential production quantity implied by the underlying GAEZ yield raster. We then convert raw quantities to kg using GAEZ's crop-specific conversion factors and unit adjustments.[21] Yield per suitable hectare is then computed for each disaggregated crop. Since Fischer et al. (2021) report potential yield at a more disaggregated crop level than the commodity groups used for actual output, each commodity group (e.g., vegetables) may contain multiple sub-crops (e.g., tomato, onion). We assign to each commodity group the maximum yield per hectare across its constituent sub-crops, retaining the best-performing crop as the representative potential. We treat potential yield as time-invariant over our analysis period, as the GAEZ potential-yield layer reflects 1981–2010 conditions.

Actual Output For actual output, we obtain harvested area and production per commodity category in each unit and country in the years 2000 and 2010. We use harvested area in hectares, while production units vary by commodity. For most commodity groups, production is reported in physical quantities, which we convert to kilograms; for several commodity groups that aggregate heterogeneous sub-crops — vegetables, stimulants, fruits, pulses, fodder crops and 'other crops' — GAEZ reports production in millions of Geary-Khamis dollars rather than physical quantities. We therefore retain the production-unit indicator in the crop-level panel. The crop-level dataset is merged at the unit-by-commodity-by-year level, yielding a panel that combines time-invariant potential yield with time-varying actual harvested area and production for each administrative unit in the sample.

C.3 Expenditure and trade shares on crops

We estimate the share of expenditure of each country on each crop in 2000 and 2010 based on the Food Balance Sheets (FBS) of Food and Agriculture Organization of the United Nations (2023b).[22].

[21] For most crops, agro-climatic potential yield is reported in kg dry weight per hectare. Exceptions include sugar beet and sugarcane, reported in kg sugar per hectare; oil palm and olives in kg oil per hectare; and cotton in kg lint per hectare.

[22] We use the historic series covering the period up to 2013. The new series, which covers that period from 2010 onward, does not provide data for the year 2000. Given the changes in methodology between the two series, combining

For each country h, crop c, and year t, expenditure shares are

$$\text{ExpShare}_{h,c,t} = \frac{\text{Supply}_{h,c,t} \times \widetilde{P}_{h,c,t}}{\sum_{c'} \text{Supply}_{h,c',t} \times \widetilde{P}_{h,c',t}}, \tag{37}$$

where $\text{Supply}_{h,c,t}$ is the physical quantity of crop c domestically supplied, and $\widetilde{P}_{h,c,t}$ is its estimated price. We also construct trade shares from the same supply quantities. We describe the construction of supply, price, and trade shares in turn below.

Supply quantities Domestic supply quantity is reported directly by the FAO as a derived variable in the balance sheet, imputed via time-series models and mirror statistics where missing (Food and Agriculture Organization of the United Nations, 2001), and defined as

$$\text{Supply}_{h,c,t} = \text{Production}_{h,c,t} + \text{Imports}_{h,c,t} - \text{Exports}_{h,c,t} + \text{Variation}_{h,c,t}, \tag{38}$$

where $\text{Variation}_{h,c,t}$ is stock at the beginning of the year minus stock at the end, so drawing down inventories is positive and accumulation is negative (Food and Agriculture Organization of the United Nations, 2001). Quantities are in thousands of metric tonnes. We map and aggregate FBS commodity categories to the crop groups used in GAEZ v4 (Fischer et al., 2021), for which we have harmonized area and yield data, retaining only disaggregated items and dropping items with unclear GAEZ mappings or no GAEZ equivalent (e.g., animal products, alcoholic beverages). [23] We use commodity-specific conversion ratios to express processed commodities and prices in raw-crop equivalents, using FAO Food Balance Sheet ratios where available.[24]

Prices We estimate $\widetilde{P}_{h,c,t}$ for sample countries as a weighted average of a local and an international price, weighted by the crop's net import share $\text{ShareImported}_{h,c,t}$:[25]

$$\widetilde{P}_{h,c,t} = \text{ShareImported}_{h,c,t}\, P_{i,c,t} + (1 - \text{ShareImported}_{h,c,t})\, P^{local}_{h,c,t}, \tag{39}$$

The local price $P^{local}_{h,c,t}$ is farm-gate producer prices from FAOSTAT (Food and Agriculture Organization of the United Nations, 2023a), in USD per tonne.[26] The international price $P_{i,c,t}$ is based on three sources: World Bank (2024) where available; average monthly Food and Agriculture Organization (FAO) prices for sunflower and rapeseed; and, for remaining categories (fruits, vegetables, pulses), a production-weighted average of country-level FAOSTAT farm-gate prices.[27]

the two methods would yield inconsistent measures (Vonderschmidt et al., 2024)

[23] For Mauritius, a known sugarcane exporter (Food and Agriculture Organization of the United Nations, 2021), the historic FBS series reports no export, import, or stock-variation quantities for sugarcane, so domestic supply equals total production in the raw data. We impute exports as 95% of recorded production, so 5% of production constitutes domestic supply in each year.

[24] For processed sugar categories, we follow Schwarzmueller and Kastner (2022) by allocating sugar-equivalent quantities to sugar cane and sugar beet using processing shares where available, and production shares otherwise.

[25] $\text{ShareImported}_{h,c,t}$ is conceptually similar to, but constructed separately from, the net import share in Equation 41: it is $(\text{Imports}_{h,c,t} - \text{Exports}_{h,c,t})/\text{Supply}_{h,c,t}$, using FBS-reported domestic supply, and is used only to construct $\widetilde{P}_{h,c,t}$.

[26] These do not incorporate processing or marketing margins, so we underestimate prices — and expenditure shares — for crops where processing adds substantial value (e.g., palm oil, where we use the palm fruit price as a proxy).

[27] Where a local or international price is missing, we fall back to an Africa-wide production-weighted price, then a world average, then a deflated value using average cross-item price growth 2000–2010 where only one year is missing;

Trade shares We also construct trade shares from the same FBS quantities, quantity- rather than value-based, to avoid the price-imputation issues above; the cost is treating units within a GAEZ group as equal, e.g., a tonne of apricots and a tonne of watermelons both count as "Fruits, other" despite different prices. Let $M_{h,c,t}$, $X_{h,c,t}$, $Q_{h,c,t}$ denote imports, exports, and production. Gross shares are

$$\text{ImportShare}^{gross}_{h,c,t} = \frac{M_{h,c,t}}{M_{h,c,t} + Q_{h,c,t} - X_{h,c,t}}, \qquad \text{ExportShare}^{gross}_{h,c,t} = \frac{X_{h,c,t}}{Q_{h,c,t}}, \tag{40}$$

and net shares, bounded between 0 and 1, are

$$\text{ImportShare}^{net}_{h,c,t} = \frac{M_{h,c,t} - X_{h,c,t}}{M_{h,c,t} + Q_{h,c,t} - X_{h,c,t}}, \qquad \text{ExportShare}^{net}_{h,c,t} = \frac{X_{h,c,t} - M_{h,c,t}}{Q_{h,c,t}}. \tag{41}$$

Stock variation is not used in the denominator(it is a residual balancing item in the FAO identity, though reported supply must be positive for import shares to be defined). Trade shares for the rest of the world, treated as a single unit, are set to zero.

C.4 Sectoral GDP and employment

GDP - We use country-level GDP, both total and in the agricultural sector, from the World Development Indicators (World Bank, 2026). For the 17 countries in our sample covered by the sectoral database of de Vries et al. (2021), we estimate tradable and non-tradable GDP.[28] This is based on the mapping shown in Table C.10, which mostly follows Mano and Castillo (2015), except when stated otherwise. We classify mining, manufacturing, business services, and financial services as tradable, and the rest of the non-agricultural sectors as non-tradable. For the remaining countries, we use GDP in industry and services from WDI (World Bank, 2026).[29] Agriculture remains a distinct category in our model.

Employment - We use employment from the census where available. For the 17 countries in our sample covered by the sectoral database of de Vries et al. (2021), we construct tradable and non-tradable employment with the same classification described above for GDP. We estimate that for 13 of the 17 countries for which we have sectoral GDP data.[30] For the remaining countries, and for cases where the census does not provide the split needed for the model, we use WDI employment indicators based on ILO estimates (World Bank, 2026). In these cases, we first use the employment-to-population ratio to match the country-level number of workers, and then use the shares of workers in agriculture, industry, and services to divide total employment across sectors.

We aim to follow the mapping described above, in Table C.10. Both IPUMS and LSMS data largely correspond to the classification of International Standard Industrial Classification (ISIC). In

items with no observation in any country-year are dropped from the category average. RoW uses the international price directly; sample countries with no local price use the international price; items with no recorded production get zero weight in the local aggregation, and countries with no valid local price fall back to the international price where available.

[28] Botswana, Burkina Faso, Cameroon, Ethiopia, Ghana, Kenya, Lesotho, Malawi, Mauritius, Mozambique, Nigeria, Rwanda, Senegal, South Africa, Tanzania, Uganda and Zambia.

[29] When for a country-year, sectoral GDP data are missing but total GDP is observed, we impute it using the country's most recently observed sectoral share of GDP, or its first subsequent observed share when no previous observation is available.

[30] Kenya, Senegal, South Africa and Uganda have inconsistent detailed sectoral employment variables in IPUMS, so we do not include them.

Table C.10: Sectors available in de Vries et al. (2021) with Tradable (T) / Non-tradable (N) classification

Sector	Description	T/N	Notes
Agriculture	Agriculture, forestry, and fishing	-	Separate category
Mining	Mining and quarrying	**T**	
Manufacturing	Manufacturing	**T**	
Utilities	Electricity, gas, water, waste	**N**	
Construction	Construction	**N**	
Trade services	Wholesale, retail, accommodation	**N**	Includes some elements that are classified in Mano and Castillo (2015) as tradable (wholesale, trade), and some that are not (retail, accommodation and food, vehicles repairs).
Transport services	Transportation and storage	**N**	All elements are classified in Mano and Castillo (2015) as tradable. We set as non-tradable due to the importance of local transportation in the countries of our analysis.
Business services	Info, professional, admin	**N**	
Financial services	Financial and insurance	**T**	Includes some local non-tradable components.
Real estate	Real estate activities	**N**	
Government	Public admin, health, education	**N**	
Other services	Arts, household, extraterritorial	**N**	

the main industry variable of IPUMS , 'business services' (which we define as tradable) are grouped together with 'real estate' (which we previously set as non-tradable) as a single industry. We define the following employment industries as tradable: 'Mining and extraction', 'Manufacturing', 'Business services and real estate', and 'Financial services and insurance'. The rest, except agriculture, fishing and forestry, and including 'other industry', are defined as non-tradable. The sectoral employment data for Nigeria and Ethiopia is based on the LSMS, where the relevant variables are largely consistent with the IPUMS definition. However, in Ethiopia's LSMS, there is no data on 'business services', such that 'business services and real estate' there corresponds only to individuals working in real estate. In Nigeria's LSMS, the entire sector of 'business services and real estate' does not exist, such

that only mining, manufacturing and finances are mapped into the tradable sector.

Trade - We use international trade data from UN Comtrade (United Nations Statistics Division, 2026) to construct country-level imports and exports for agricultural and industrial goods in 2000 and 2010. We classify commodities with codes up to 24 as agricultural,[31] and the remaining commodities as industrial, and then aggregate imports and exports separately by sector. We merge these flows with country-level GDP from World Bank (2026), averaged over the corresponding five-year windows, 1996–2000 and 2006–2010, to construct trade-to-GDP ratios. When a sectoral import or export flow is zero or missing, we impute it using the country's mean observed trade-to-GDP ratio for that sector and flow, and recover the corresponding level by multiplying this ratio by GDP.

C.5 Road travel times in 2000 and 2010

We construct bilateral travel times between sample units using historical Michelin African road maps digitized by Müller-Crepon et al. (2021). They extract African road networks from the Michelin map and convert the scanned physical maps into a time-varying digital road atlas. We use the 2000 and 2012 Michelin maps as representative of 2000 and 2010, respectively, with 2012 being the closest year to 2010 available in the data. We keep roads that intersect the countries in our sample and attach road speeds from the Michelin speed file from Müller-Crepon et al. (2021). The speed file maps each Michelin road class to an assumed travel speed. We follow this recoded speed variable.

For each country and map year, we convert the line network into a set of connected graph nodes (road endpoints or intersections) and edges (road segments). We then overlay the 2000 roads on top of the 2010 map. The idea behind this is that we are interested in persistent improvements in road access. Without this step, some 2010 estimated travel times would be higher than the 2000 ones due to missing links, small breaks, or imperfect alignment in the digitized maps rather than from durable road closures. Where the cleaning or overlay process creates duplicate road segments, we keep the fastest speed observed for that segment.

The graph is built separately for each country and year, which implies that when traveling from one unit to another within a country, only road segments in this country can be used. Road segments are treated as undirected graph edges, and the cost of traveling along each road edge is its length divided by the assigned road speed:

$$\text{time}_e = \frac{\text{length}_e}{\text{speed}_e}.$$

We use the highest-population point in the unit as the representative location of each sample unit. This is based on Schiavina et al. (2023) data in 2000, aggregated to a 10km grid before selecting the maximum-population cell. For the model, we aggregate units that are part of the same urban area. We intersect out units with agglomerations from 2015 from OECD/SWAC (2024). Then, for each agglomeration, we find the unit with the highest population point among the units that are located within it, and set its most populated point as the representative point of the agglomeration.

Each admin point can be connected to the road graph through straight-line access edges to nearby road nodes. For each admin unit, we identify road nodes that lie inside its polygon, and keep up to the 20 closest candidates, and connect the admin point to those road nodes with straight lines. These access links are assigned a slow off-road speed of 5 km/h and are included in the total travel time. We

[31] In the HS classification, codes 1–24 cover live animals, animal and vegetable products, foodstuffs, beverages, and tobacco. This definition excludes some agricultural raw materials classified elsewhere, such as cotton.

allow several candidate access points because the closest road node is not necessarily the relevant one for all destinations, as a unit may enter the road network from its northern side when traveling north, and from its southern side when traveling south. [32]

Some units, or groups of units, remain disconnected from the main road graph. This can occur for islands, such as Zanzibar, but also for mainland units with no road data or for disconnected road components. We therefore add a limited set of straight-line gap links between disconnected graph components that contain admin units. For each pair of such components, we add the shortest candidate link between the two components. Mainland gap links are assigned a speed of 10 km/h, while links involving islands are assigned a speed of 30 km/h. For island-mainland links, the mainland endpoint is restricted to coastal mainland units detected using Massicotte and South (2026).

Given this graph, bilateral travel times are computed using Dijkstra's shortest path algorithm (Dijkstra, 1959). For each origin admin node, the algorithm finds the minimum-cost path to every destination admin node, where the cost is the sum of road, access, and gap-link travel times along the path. The output is a country-year origin-destination matrix, with one row for each ordered pair of sample units.

After computing the 2000 and 2010 matrices, we impose a pair-level monotonicity rule, under which the 2010 travel time and distance must not exceed those of 2000. This is done to correct measurement and topology issues in the historical maps.

In the final origin-destination matrix, median travel time is 9.3 hours in 2000, 8.7 hours in 2010, and for comparison 7.7 hours using OSM data. The correlation between the 2000 and 2010 Michelin matrices is 0.98, while the correlation between the 2010 Michelin matrix and OSM is 0.76. OSM is faster in most cases, which is expected because it reflects the contemporary road network and contains more detailed local roads, especially in urban areas. Less than 1 percent of pairs use the 2000 value as the final 2010 travel time, and most of those corrections are only a few minutes.

C.6 Household survey consumption

We use household surveys from the LSMS-ISA (World Bank, n.d.) to compare the model's predicted consumption levels with measured household consumption. We use the eight countries for which we have suitable consumption information in the LSMS data: Ethiopia, Ghana, Malawi, Mali, Nigeria, South Africa, Tanzania, and Uganda. For each country, we use the survey wave closest to 2010.[33] In the model, we use admin units that are not necessarily identical to the geographical units reported in the LSMS data. We match units using the corresponding admin codes or names, when possible, and collapse survey units where needed to match the harmonized model geography.

Consumption - The consumption measure is annual household expenditure per adult equivalent. We aggregate household consumption to the admin units used in the model using survey weights. When the survey files provide real consumption, we use that measure. When they provide nominal consumption and a spatial price deflator, we divide nominal consumption by the deflator before aggregating. For Nigeria and Ghana, we use nominal consumption.

[32] To keep access points comparable across years, we first choose access links using the 2010 road graph. For 2000, we retain the same access links when the corresponding road node also exists in the 2000 graph.

[33] We use Tanzania NPS 2010, Malawi IHS3/IHPS 2010–2011, Uganda UNPS 2013, Mali EAC-I 2014, Ghana 2010, South Africa 2012, Ethiopia ESS 2013, and Nigeria GHS 2010.

Expenditure shares - We also use the household surveys to compare the model's predicted expenditure shares on agricultural goods with data from the LSMS, for four countries for which we can construct those shares: Ethiopia, Malawi, Mali, and Nigeria. Agriculture expenditure is measured from food or agricultural-consumption categories, including food consumed at home, purchased food, and the value of own-produced food consumed by the household, when those are reported separately. We aggregate the household-level shares to the model units using survey weights and compare them to the model-implied agricultural expenditure share.

C.7 Country level matrices

Road distance and travel time - For the country-level analysis, we use the same road layers and graph construction described above, but compute travel times between one representative point per country rather than between all sample units. The representative country point is the most populated 10km cell in the country in 2000, aggregated from Schiavina et al. (2023). Unlike the within-country matrices, the country-level graph uses all cleaned Sub-Saharan African roads, so routes can pass through third countries.[34]

Trade - We construct bilateral trade flows from the import records of UN Comtrade (United Nations Statistics Division, 2026). For each destination country, we sum imports from each origin country by year. We then average annual values over three windows (1996–2000, 2006–2010, and 2016–2020). Domestic trade is defined as GDP measured in current USD (World Bank, 2026), minus total exports. To avoid negative or implausibly small internal trade, we impose that it should be equal to at least one percent of the country's GDP.

Migration - We construct bilateral migration flows based on the estimates of Gaskin and Abel (2026). We use a measure of flows based on previous residence, and sum migration for each pair of countries over five-year windows (1990–1995, 1996–2000, 2000–2005, 2006–2010, 2011–2015, and 2016–2020).

C.8 Control variables

We construct measures of local topography, climate, and armed conflict in each admin unit. For each relevant forest-loss period by country, we get the mean and median of conflict, temperature and precipitation for each unit, whereas topography is constant over time.

Topography For each unit, we extract mean elevation and slope from the SRTM digital elevation model (Jarvis et al., 2008). We compute a terrain ruggedness index (TRI) for each unit from the data of Nunn and Puga (2012), by averaging across cells and weighting by cell area, while excluding water bodies.

[34] Madagascar and Mauritius are connected to the mainland road network using artificial water-gap links at 30 km/h. This means that a route from Madagascar to another country first travels in a straight line from the Madagascar representative point to the nearest mainland road node, and then follows the road network from there.

Climate We construct annual precipitation and temperature for each unit over 1995–2015. Precipitation is computed as the annual sum of daily rainfall estimates from CHIRPS (Funk et al., 2015), while temperature is the annual mean of ERA5 2m air temperature (Hersbach et al., 2023).

Conflict Annual conflict counts are constructed from the UCDP Georeferenced Event Dataset (Sundberg and Melander, 2013). For each unit and year, we count the number of conflict events that takes place inside the unit's territory.